Ribosome Life

From RNA duplication to polypeptide translation & beyond

Edward M. Hedgecock & Rui B. Proença

Department of Biology

Johns Hopkins University

Baltimore MD 21218

United States

PDF 2025 JANUARY 14

ribosomelife.com

Summary

The RNA world has left no clear vestiges of ribozymes for polynucleotide replication (sections 1-4). It is untenable that this central process of RNA life vanished without a trace. Conversely, the RNA world had no clear progenitors of rRNAs and tRNAs of polypeptide translation. It is untenable that this central process of protein life sprang from nowhere. In “An RNA replisome as the ancestor of the ribosome” John Campbell at UCLA explained at once the disappearance of replication ribozymes without descendants, and the appearance of translation ribozymes without ancestors (Campbell 1991). Like the discovery that birds are living descendants of dinosaurs hidden in plain sight, he proposed that rRNAs and tRNAs are surviving descendants of the lost replisome. In this original form of processive RNA copying, the template was read from 5’ to 3’ in nucleotide triplets (*codons*), while identical triplets (*duplicons*) were added from 3’ to 5’ to the nascent copy. Ancestral to transfer RNAs, the adaptors were 64 donor RNAs, each with an *anticodon* triplet near their middle and matching duplicon at their 3’ end.

Campbell’s paper was the first of several schemes for RNA copying based on conjectural ancestors of the ribosome and tRNAs (Weiss & Cherry 1993; Gordon 1995; Poole et al 1998; Yakhnin 2007; Noller 2010/2012). We précis these earlier proposals, introducing some common terms for their description and comparison (sections 5-7). Each scheme appealed to then current knowledge of ribozymes, but those authors little mention, must less build upon one another. In common to all these schemes, the ribosome ancestor, or *duplisome*, used a small adaptor, or *donor RNA* (dRNA), to read templates 5’ to 3’ in one codon steps, and add some form of *duplicon* to nascent polynucleotides. But these schemes otherwise vary widely in the conjectured structure of their dRNA substrates and polynucleotide products.

We propose a simple RNA copying scheme that retains the virtues of Campbell (1991) without its defects (section 8). *Notably, our duplicon comprises the first two nucleotides of the dRNA, not the final three.* As a result, a faithful forward duplicate is made 5’ to 3’ using only 16 dRNAs. Our choice for the size and location of the duplicon constrains the chemistry of polynucleotide elongation, and raises unexpected

questions for the evolution of polymer life. We first discuss the structure of dRNAs and their loading from random oligonucleotides (section 9), the elongation and termination reactions of the duplisome nucleotidyl transfer center (section 9), the energetics and kinetics of the elongation cycle (section 10), and the mechanism of decoding (section 11). We then discuss the origin of the duplisome (section 12), the origin of tRNAs (section 13), the origin of tRNA charging ribozymes (section 14), the breakout of polypeptide translation (section 15), the saga of protein life (sections 16-17), the DNA world (section 18), and beyond (section 19).

Contents

1. The RNA world

The *RNA world* conjectures that life on Earth began as self-replicating communities of RNA molecules, augmented later with coded proteins as more versatile gene products, and later still by duplex DNAs as more durable and reparable genes (Rich 1962; Woese 1967; Crick 1968; Orgel 1968; Pace & Marsh 1985; Gilbert 1986; Lazcano 2016; Fine & Pearlman 2023). Navigating a treacherous strait of natural philosophy, this hypothesis avoids *dualism* insomuch as these first genes were not materially different from prebiotic molecules, nor so improbable in sequence as to be miraculous. It also avoids *reductionism* insomuch as the kinetic principles of RNA life realize the means-ends search of a living system that explores likely means, that is, affordances of its surroundings, for no higher ends than propagating those discoveries. In one early formulation of polymer life, selfish ends and communal means were realized as genes and proteins, respectively, coupled through catalytic cycles of molecular survival and replication (Eigen 1971). The new insight was that polyribonucleotides could play both roles: When unfolded, they were autocatalysts of self-replication called templates, or *ribogenes*. When folded, they were common catalysts called *ribozymes* in analogy to catalytic proteins called *enzymes* (Kruger et al 1982; Guerrier-Takada et al 1983).[1]

Pari passu with the two distinct catalytic roles of RNA molecules, there were two distinct levels of evolutionary selection in the RNA world. As selfish autocatalysts, each RNA molecule competed for common resources of copying, say a pool of precursors, and replicase ribozymes. As common catalysts, each molecule cooperated in promoting the survival and reproduction of the entire RNA community, competing as a whole with similar communities. Thus, at the dawn of life, evolutionary dynamics divided broadly into *vertical gene transmission*, that is, ways of promoting fairness, while limiting selfishness in reproduction of the whole community, and *horizontal gene transmission*, that is, ways of exploring opportunities, while mitigating dangers, from

[1] We reserve *ribozyme* and *enzyme* for RNA- or protein-based catalysts, respectively. Thus, it is inconsistent to say proteinaceous ribozyme, and needless to say proteinaceous enzyme. Names with suffix *-ase*, *viz*. *nuclease*, *polymerase*, or *kinase*, describe a chemical reaction, not composition of the catalyst. Thus, we say polymerase ribozyme, or polymerase enzyme where the type of catalyst matters.

mixis of entire communities, or more limited movements of particular RNAs from one community to another.

In any scheme of RNA life, some form of compartment mediated the coopetition within and between RNA communities contained within aqueous droplets, or upon hydrated surfaces (Mizuuchi & Ichihashi 2021). Likely physical compartments include naked aerosols, interstices of icy brines, and mineral surfaces or pores of rocks and clays (Oberbeck et al 1991; Kanavarioti et al 2001; Ferris 2005, 2006; Attwater et al 2013; Mutscher et al 2015; Zhang et al 2022). According to whether their bulk fluid was air, oil or water, respectively, likely organic compartments include (1) encapsulated aerosols, (2) droplet emulsions in oil, and (3) polyelectrolyte condensates, micelle suspensions, or vesicles with one or more membranes, made of lipids or polypeptides, in water (Hanczyc et al 2003; Jin et al 2018; Poudyal et al 2018; Cohen et al 2022; Ianeselli et al 2023). Two of these organic compartments, lipid bilayer membranes and polyelectrolyte condensates are important in cellular life (refs).

Besides spatial compartments, sequence-based compartments, or cliques, are sophisticated forms of kinetic disproportionation in well-mixed solutions intrinsic to the polynucleotides themselves. Sequence-based compartments arise from positive regulators such as genome tags that identify members of the RNA community for selective aggregation and preferential replication, and negative regulators such as site-specific endonuclease ribozymes that target foreign RNAs for preferential degradation (Weiner & Maizels 1987; Joyce & Orgel 1993). Like regular updating of IDs and passwords to limit counterfeiting, the guide sequences used for primer- or cleavage-site recognition provided running definitions of self for reproduction, and non-self for destruction, respectively.

There are four principal constraints on exhuming the RNA world: (1) physical inferences about planetology and geology of the hadean eon, (2) laboratory experiments on prebiotic chemistry in non-equilibrium environments, (3) rational design and *in vitro* selection of artificial ribozymes, and most clearly, (4) molecular biology and comparative genomics of cells and viruses today. There is some debate whether our planet attained a habitable crust and hydrosphere soon after the great impact that formed the Earth and the Moon, or only later, after a continued period of heavy

meteorite bombardment (Abramov & Mojzsis 2009; Benner et al, 2018; Pearce et al 2018). Thus, prebiotic organic chemistry may have commenced anywhere from 4.5 Ga to 3.9 Ga. Regardless of just when stable preconditions of life were finally attained, there is strong evidence from stromatolites, microfossils, and isotope ratios that cellular life had evolved by about 3.7 Ga (Nutman et al 2016; Javaux 2019).

Whether the period from a habitable planet to cellular life was nearly 800 million years, or barely 200 million, several major evolutionary transitions occurred within this interval. The intermediate stages are characterized by (1) whether RNAs, proteins and DNAs were present, and (2) just how each polymer was made. In one popular scenario, life began with the fitful reproduction of RNA communities by spontaneous copying, leading to faster and more accurate ribozyme-assisted replication (Table 1-1). After the invention of protein translation, these conjectural RNA-directed RNA polymerase (RdRP) ribozymes were retired, replaced by extant polymerase enzymes. Coupling RdDP and DdRP enzymes, duplex DNA stores provided a more stable polymer backbone, as well as allowed evolution of repair pathways that exploit the informational redundancy of complementary strands for error-free repair. The invention of DdDP enzymes, allowing large chromosomes with many genes, consolidated the genome handover to DNA popularized as the *central dogma* of molecular biology (Crick 1955/1958, 1956, 1970; Watson et al 2016). Henceforth, the Janus-faced roles of polyribonucleotides in the RNA world as ribogenes and ribozymes could be delegated to DNAs and proteins, respectively.

Era of life	RNA	Peptides	DNA
prebiotic	random	random	none
early RNA	spontaneous copying	random	none
late RNA	RdRP ribozyme	random	none
polypeptide	RdRP ribozyme	ribosome	none
protein	RdRP enzyme	ribosome	none
early DNA	DdRP enzyme	ribosome	RdDP enzyme
late DNA	DdRP enzyme	ribosome	DdDP enzyme

TABLE 1-1. THE FIRST BILLION YEARS OF POLYMER LIFE ON EARTH

2. Dawn of RNA life from prebiotic molecular clutter

A bold defeasible conjecture about the origin of life on Earth, the RNA world has been used to triage open-ended questions of astrobiology and geochemistry, identifying a handful of major problems to solve (Popper 1959; Joyce & Orgel 1993; Orgel 2004; Szostak 2012, 2017). The first three problems are to demonstrate (1) a prebiotic source of random oligonucleotides, (2) a spontaneous process of RNA copying to multiply chance sequences, and (3) a compatible compartmentation of RNA communities. The last four problems are to explain (4) the (extinct) ribozymatic process of RNA copying, (5) the breakout of polypeptide translation, (6) the new enzymatic process of RNA copying, and lastly (7) the genome handover of longterm storage to DNA. Overlapping both sets of problems, origin of life studies must explain the domestication of energy and metabolism from prebiotic processes, including (8) the sources of amino acids and polypeptides along the way to coded proteins, and (9) the sources of Gibbs energy along the way to the NTP currency used for nucleic acid replication and protein translation in cellular metabolism.[2]

Prebiotic reactions produce a clutter of aliphatic hydroxy and amino acids, including proteinogenic α-amino acids (Miller 1953, 1957; Miller & Urey 1959; Ring et al 1972; Wolman et al 1972; Ferris et al 1978; Cronin 1989; Kasting & Brown 1998). Yields of these prebiotic or *primary amino acids* decline with increased Gibbs energy of formation in rough order G A D E V S I L P T (Higgs & Pudritz 2009). More complex amino acids K F R H N Q C Y M W are rare or absent in abiotic syntheses, as well as carbonaceous chondrites such as the Murchison meteorite, indicating their origin likely

[2] Here we gloss the terms prebiotic, abiotic and biotic: *Prebiotic* refers to the processes, necessarily abiotic, thought to have preceded and been conducive to the origin of life on Earth. Organic and geochemists can model proposed prebiotic environments and their abiotic processes in the laboratory. A *biotic* or living system is any process that explores and exploits affordances of its surroundings for Darwinian purposes of survival and reproduction. For the RNA world, this cashes out as communities of polymers with lineages (*ribogenes*) performing sequence-dependent tests (*riboswitches*) and actions (*ribozymes*). RNA life *tout court* began with one or more ribozymes that modestly assisted *abiotic* non-equilibrium processes, hitherto spontaneous or without evolved catalysis, of RNA copying and sequence amplification. There is a clear conceptual difference between the original prebiotic affordances as found, and new biotic affordances as invented, but all affordances work together seamlessly, and any given environment or invented niche changes over time. Over the great sweep of evolution, abiotic material and energy resources were domesticated, and meanwhile, living species adapted to specialized niches, as well as expanded their range.

required biotic catalysis (Kvenvolden et al 1970; Pizzarello 2006; Martins et al 2007; Cobb & Pudritz 2014; Koga & Naraoka 2017). It is unknown, however, which were *secondary amino acids*, first made in significant amounts by now extinct ribozymes, and which were *tertiary amino acids*, first made by extant enzymes.

Compared to the primary amino acids, prebiotic syntheses of nucleotides are far more challenging (Orgel 2004). In purine nucleotide biosynthesiss, phosphoribosyl transferases add adenine, hypoxanthine and guanine to the activated ribose PRPP (5-phospho-ribose-1-pryophosphate) to form AMP, IMP or GMP, respectively. Additional enzymes convert IMP to AMP or GMP. Another phosphoribosyl transferase adds orotic acid to PRPP to form OMP which is then converted to UMP. Dedicated kinases promote AMP, GMP and UMP to their NDP, while a common kinase promotes these to their NTPs. Completing the four NTPs, *CTP synthetase* uses ATP to transfer the amine from glutamine to UTP. Besides these *salvage* pathways that make all four NMPs starting from the pools of heterocyclic nucleobases and activated ribose, there are pathways for *de novo* synthesis of orotic acid and inosine monophosphate (IMP). Whereas the pyrimidine ring of OMP is essentially complete when orotic acid is added to PRPP, the purine ring of IMP called *hypoxanthine* is built on PRPP by piecemeal additions of carbon and nitrogen from aspartate, glycine, glutamine, and THF. The enzymatic pathways for adding essentially complete pyrimidine or purine rings to PRPP are likely older than pathways for biosynthesis of these rings. All and all, these coupled and cross-regulated pathways provide balanced pools of all four canonical NMPs for RNA and DNA synthesis, as well as various individual roles of NTPs in cellular metabolism.

Remarkably, various abiotic reactions produce ribose, nucleobases, and even nucleosides, amidst a great clutter of other products (Butlerow 1861; Breslow 1959; Oro & Kiball 1960, 1961; Oro 1961; Mueller et al 1990; Anastasi et al 2006; Nam et al 2018; Becker et al 2016, 2019; Sutherland 2016; Yadav et al 2020). One challenge for the synthesis of nucleosides by nucleobase addition is activation of ribose by phosphorylation or other leaving group (Lohrmann & Orgel 1973; Reimann & Zubay 1999). Whereas nucleobases are added to activated ribose in salvage pathways, and hypoxanthine is built upon the ribose carrier in *de novo* purine biosynthesis, there is

one abiotic pathway with no biotic analog, *viz*. formation of nucleosides directly without forming ribose first (Powner et al 2009). Finally, whereas cells creates nucleotide monophosphates from PRPP, *viz*. the activated form of ribose-5-phosphate, prebiotic syntheses make nucleosides, leaving the problem of nucleoside phosphorylation.

No one has demonstrated a one-pot synthesis for anything like a complete feedstock for RNA life. However, there is no reason why sundry prebiotic reactions need to have occurred in one place, nor used common materials and energy sources (Benner et al 2012). Indeed some organic compounds were formed at great removes in space and time in extraterrestrial environments (Chyba et al 1990; Chyba & Sagan 1992; Oba et al 2022). Thus, there were likely several important prebiotic environments with signature materials and energy sources (Stueken et al 2013; Ianeselli et al, 2023). Geological diversity and multi-pot syntheses solve problems of incompatible chemistries, but any such heterogeneity required matching means to stabilize, concentrate and purify oligonucleotides or their precursors, made in modest yield among a clutter of side products (Benner et al 2018; Sasselov et al 2020). Likely geophysical processes to concentrate and enrich these substrates from aqueous solutions include evaporation, freezing, and mineral adsorption (Bernal 1951).

There is yet no consensus on which prebiotic reactions were significant on the hadean Earth, and many gaps must yet be filled, but confidence is high that such pathways can in time be fleshed out, given the compelling evidence of life itself (Orgel quote; Kitadai & Maruyama 2018). Whatever these cyclical processes of concentration, reaction, purification and transport, some niches became in effect commissaries of life, where all of the essential ingredients from the rough-and-tumble of abiotic syntheses were available at once (Wu & Sutherland 2019). *For purposes of this paper, we merely stipulate some prebiotic processes for formation of nucleotides and primary amino acids, as well as compatible processes for their condensation into random oligonucleotides and polypeptides, respectively.*

Still largely unknown, the prebiotic processes of monomer activation and polymer formation cannot be reasonably extrapolated from the pathways and intermediates of cellular metabolism. At the same time, the biological mechanisms of polymer formation frame our understanding of these problems, and illustrate some thermodynamic and

kinetic principals for their solution. Thus, in living cells, polynucleotides and polypeptides are hydrolyzed by nucleases or peptidases, respectively, that increase the number of fragments by one at the expense of one polymer bond. Gibbs energy greatly favors hydrolysis over condensation: The enthalpy reflects the relative stability of the free ends, after any ionization of terminal groups, while the entropy reflects the greater concentration of bulk water (*ca*. 56 M) than free ends (Martin 1998).

In nucleic acid replication, the *water problem*, or unfavorable Gibbs energy of polymer condensation, is solved by coupling bond formation to a high-energy leaving group on mononucleotides, *viz*. the inorganic pyrophosphate of (d)NTP. Polymerase enzymes add these activated monomers to the polynucleotide 3' OH. The greater stability of the phosphoester bond over the phosphoanhydride drives polymerization, while enzymatic hydrolysis of the freed pyrophosphate prevents product inhibition, and makes the transfer reaction effectively irreversible.

Remarkably, in protein translation, the water problem is solved by coupling bond formation to a moderate-energy leaving group, not on the monomer, but the nascent polymer. To wit, an enormous alcohol (the tRNA) is esterified to the carboxy terminus of the polypeptide. Ribosomes add aminoacyl-tRNAs to this activated polypeptide by positioning their α-amine, a strong nucleophile, to attack its acylester bond, meanwhile excluding bulk water from the catalytic center. The proximate source of Gibbs energy is the greater stability of the acylamide bond produced than the acylester bond consumed. More distal sources, notably ATP used by the tRNA charging enzymes and GTP used by translation factors, improve speed and fidelity of the elongation cycle in protein translation.

Inorganic polyphosphates, organic phosphates such as acetyl-phosphate or glycerate-phosphates, and sundry less "biological" molecules have been proposed as common stores of chemical free energy on the hadean Earth (Eakin 1963; Lohrmann & Orgel 1973; Pasek 2020; Pinna et al 2022). For any energy storage intermediate, origin of life studies must demonstrate a likely prebiotic synthesis, as well as how this substrate is coupled to polymer condensation or other organic reactions. A variety of chemical leaving groups have been considered for prebiotic activation of nucleotides and amino acids for condensation. Ideal candidates must be energetically sufficient to

drive condensation, kinetically sluggish to allow evolution of ribozymatic condensation, and of course, have a plausible abiotic synthesis. For prebiotic RNA synthesis, the 5' OH of polyribonucleotides has been proposed to attack the strained 2',3' cyclic phosphates of ribonucleotides or the 3' ends of polyribonucleotides to lengthen the polymer by one nucleotide or ligate two polyribonucleotides together.

Some origin of life studies defer the quixotic search for prebiotic organic reactions in favor of elucidating the probable context of these reactions, that is, Gibbs energy sources and equilibrium compositions of likely hadean environments. There are three broad classifications of energy sources according to cosmological origin, geological distribution, and temporal patterning, respectively (Deamer & Weber 2010). First, the ultimate source may be solar, terrestrial, or extraterrestrial. Thus, solar radiation drove chemical reactions directly, surface warming less directly, and atmospheric (wind, rain, lightning) or oceanic (currents) phenomena least directly. Meanwhile, geological and geochemical processes, as well as extraterrestrial impactors, provided fluxes of both energy and materials.

Second, few if any energy sources were everywhere available: Some only operated at defined planetary latitudes, atmospheric heights, or oceanic depths. Some only operated at the interfaces of air, water, or mineral surfaces. Some only operated in isolated locations (hydrothermal vents, hotsprings, geysers, salt flats). Of these interfaces and locals, some were determined by extrinsic discontinuities (volcanos, coasts), while others arose by disproportionation of initially homogeneous environments (nuclear geysers). Finally, some energy sources were more or less steady (geochemical reactions, radioactive decays), others were entrained to planetary motions, and some fluctuated erratically. Planetary cycles included once-a-day warming and twice-a-day tides, as well as seasonal atmospheric, terrestrial and oceanic phenomena (Gordon & Mikailowsky 2021). Electrical storms and volcanic eruptions occurred more erratically, but on comparable time scales of days or years. Finally, some events occurred quite sporadically, even just once (Moon formation), in the planet's history.

Most likely the water problem of polymer condensation was first solved at physical interfaces that concentrated monomers and oligomers to increase their chemical

activity, and excluded water to decrease its chemical activity (Rodriguez-Garcia et al 2015; Ross & Deamer 2016; Erastova et al 2017; Wei et al 2020; Hao et al 2022; Holden et al 2022). Likely processes include adsorption-elution from clay or mineral surfaces; freezing-thawing of briny eutectics; and drying-wetting of aerosol microdroplets and other evaporites. Extremely complex evaporites form on exposed mineral surfaces that are alternately covered in shallow water from rains, dews, surfs, tides, geysers, etc., and then dried under air. Here repeated cycles of wetting for mixture and drying for condensation can drive polymer elongation. In favorable cases, a regular cycle of wetting and drying, accompanied by cooling and heating, respectively, is driven daily or annually by sunlight. Deliquescent salts likely moderate the humidity swings between wet and dry states (Campbell et al 2019).

The usual distinction between physical and chemical processes of polymer condensation, *viz*. physical concentration and heating of monomers and oligomers, versus their chemical activation and catalysis, is likely too sharp. Thus, air-water or water-mineral interfaces that concentrate monomers and oligomers may also catalyze polymer condensation. In polypeptide formation, the physical principle of dry-down condensation can be combined with chemical activation of amino acids and peptides by esterification to simple alcohols as leaving groups (Griffith & Vaida 2012; Forsythe et al 2015). In polynucleotide formation, activated mononucleotides can form random oligomers in the absence of any template, catalyzed by clay surfaces or metal ions (Ferris 2005, 2006; Gibard et al 2018; Liu et al 2020; Pasek 2020).

Once provisioned with monomers and random oligomers in sufficient concentration, if not purity, the second problem of RNA life was a more-or-less faithful copying of oligonucleotides, allowing Darwinian evolution of the first ribozymes. Considering the genetical implications of the newly discovered structure of DNA, James Watson and Francis Crick first conceived of *template-directed copying* of polynucleotides (Watson & Crick 1953a,b). Extrapolating phenomena of nucleation and growth from 3-dimensional crystals to linear polymers, they suggested wherever free monomers stacked together along an existing chain, adjacent monomers are positioned to polymerize just in case their nucleobases pair (hydrogen bond) correctly to those of the template. In this way, each polynucleotide could direct the synthesis of its own unique

reverse complement. Watson and Crick were agnostic whether the specific template sufficed for replication, or a general replicase was required as well. In either case, a second round of polymerization recreated the original template sequence.

As conjectured, activated mononucleotides and oligonucleotides polymerize faster when correctly paired to a template (Sulston et al 1968a,b; Orgel & Sulston 1971; Inoue & Orgel 1983; Wu & Orgel 1992; Zhou et al 2019a). The kinetics of template-directed copying combines convergence toward fixed-points, familiar from purification through repeated crystallization, with divergence from branch-points seen here as the autocatalytic amplification of sequence variants. Because canonical features of RNA, *viz*. D-ribose sugar, nucleobase alphabet, and 3',5' backbone linkages have modest kinetic advantages over competing clutter, the chemical purity, homochirality, and regioselectivity all can increase, up to a point, through repeated rounds of copying (Giurgiu et al 2017; Kim et al 2021; tailwinds ref; Ross & Deamer 2022, 2023).

All three kinetic principles, perservation of polymer type, nearly faithful copying of polymer sequence, and fair copying of rare sequence variants, defined the genome, or molecular store for hereditary knowledge, of the RNA community. Pure RNA polymers may have emerged early, or there may have been a long *conviviencia* of RNA-like polymers with significant fractions of 2',5' linkages, deoxyribose, non-canonical nucleobases, *viz*. 2,6-diaminopurine or hypoxanthine, or modified nucleosides (Cafferty et al 2018; Fialho et al 2020). Here spontaneous copying performed purifying selection, a form of default repair, for generic features of RNA structure. Meanwhile, occasional changes in nucleotide sequence could be inherited, and importantly, selected upon.

Rather than RNA-like polynucleotides with admixtures of familiar elements, Darwinian evolution may have begun with genomes made of another linear copolymer entirely, perhaps a xenonucleic acid (XNA) with no counterpart in modern life, whose backbone and nucleobases were favored by prebiotic chemistry (Cairns-Smith & Davies 1977; Nelsesteun 1980; Schwartz & Orgel 1985; Weber 1989; Nielsen et al 1991; Joyce & Orgel 1993; Nielsen 1993; Eschenmoser 1999, 2004, 2005; Orgel 2003, 2004; Robertson & Joyce 2012). More radically, this ante-RNA life may have used lattice imperfections on mineral surfaces as 2-dimensional templates (Cairns-Smith 1982).

Xenogenomes introduce three difficult new problems: First, the problem of spontaneous genome copying is kicked back from RNA to XNA or mineral surfaces. Second, for Darwinian selection, xenogenes must act, either more or less directly, on their local surroundings. Whereas folded XNAs might act as xenozymes, the chemical affordances of clay or organopyrite genes are less obvious, and presently, more speculative (Cairns-Smith & Hartman 1986; Wachtershauser 1988). Third, any xenogenome must be systematically translated from the original mineral or XNA medium into familiar nucleic acids, or else essential xenogenes replaced piecemeal by non-orthologous ribogenes, and nonessential ones lost entirely (Hud et al 2013; Zu et al 2022). Genome handover (aka *genetic metamorphosis*) was a bold idea when first proposed for a mineral proto-genome (Cairns-Smith 1965, 1982). Since then, the wildly successful translation of RNA genomes into duplex DNA has been reconstructed from reverse transcriptase and the ancillary enzymes of DNA synthesis (section 18).

Gerald Joyce and colleagues characterized the debate over ante-RNA life thus: "Scientists interested in the origins of life seem to divide neatly into two classes. The first, usually but not always molecular biologists, believe that RNA must have been the first replicating molecule and that chemists are exaggerating the difficulties of nucleotide synthesis. They believe that a few more striking chemical "surprises" will establish that a pool a racemic mononucleotides could have formed on the primitive Earth, and that further experiments with different activating groups, minerals, and chiral amplification processes will solve the enantiomeric cross-inhibition problem. The second group of scientists are much more pessimistic. They believe that the *de novo* appearance of oligonucleotides on the abiotic Earth would have been a near miracle. Time will tell which is correct" (Robertson & Joyce 2012 p15). We belong to a third camp who worries that the project of discovering, or inventing ante-RNA life turns on borrowing from Peter to pay Paul. We consider the case for an ancient xenogenome no further; wherever we posit a spontaneous (prebiotic) process, or a biotic process based on lost ribozymes, aficionados of ante-RNA life may insert their favorite xenozymes.[3]

[3] Finally, there is a distributed … known as metabolism first. CNO nucleosynthesis, metabolic networks, lack the data storage principle that separates transmission of knowledge from its expression [metabolism-first][Kaufmann]

One unavoidable question is the order of invention of nucleobases as it is unlikely that all of the standard four nucleobases were available before life commenced …. Crick inosine:uracil code … cytosine late … thymine later and DNA only (section 18)

A life cycle of spontaneous polynucleotide copying, however fitful and error-prone, and folding of these products, however short and unstable, marked the dawn of RNA life. In open-ended exploration of the affordances of their sequence space, communities discovered and exploited ribozymes that facilitated their survival and propagation. In this breakout biotic environment, first likened to a warm little pond, or later a rich organic soup, the comparatively delicate physiology of RNA life played out amongst molecules solvated and folded, at least from time to time, in aqueous solution not too different in from the intracellular milieu of modern life (Darwin 1859; Oparin 1938).

Beyond prebiotic clutter and intractable tars, RNA life faced thermodynamic sinks and kinetic traps. Strand annealing and tertiary folding are slow processes favored by cold, salts, and neutral pH, while unfolding and melting are favored by warmth, dilution, and acidic pH (Tinoco & Bustamante 1999). Fluctuations of one or more of these factors shift the balance from annealing and folding, toward unfolding and melting (Ianeselli et al 2023). Forming thermodynamic sinks, long duplexes with high melting temperatures (T_m) required large environmental fluctuation to melt their secondary structure. Forming kinetic traps, strands became hopelessly entangled when cooled too quickly, or shifted abruptly to higher salts, without reaching their equilibrium folds. All in all, selection favored polynucleotides with melting temperatures only slightly above the ambient high, poised to fold readily without becoming trapped. Other factors, notably metal cations and small organic molecules, including non-coded polypeptides, could chaperone and regulate this folding.

Like the porridge of Goldilocks, physiological conditions within cells have a narrow range of optimal values for temperature, salinity, pH, redox potential, etc. Larger swings arrest growth, if not destroy life, and are at best tolerable, never obligate parts of the life cycle of cellular organisms. The discovery of extremophile prokarya, as well as extremotolerant, or cryptobiotic eukarya including lichen, nematodes, rotifers, and tardigrades, has probed earlier assumptions about the physicochemical limits of life

(Gade et al 2020). Although these species are proof of principle that life can survive, or even thrive, in harsh environments, most if not all of the extant adaptations for extreme environments are derived traits, not vestigial abilities of LUCA, much less earlier life.

Absent the domesticated sources of Gibbs energy and NTP currency of cellular metabolism, the first RNA communities relied directly on environmental fluctuations in their life cycle. Moderate changes of temperature, salts, RNA concentration and pH were needed to alternate between unfolding and melting for template copying, and annealing and folding of working ribozymes. Still greater fluctuations were needed for concentration by adsorption, evaporation or freezing, as well as abiotic activation, or dry-down condensation. All in all, early RNA life paused frequently for the necessary environmental swings, and communities died occasionally from harsh or sustained conditions. Whatever the nature, magnitude and cause of these fluctuations, in the environments conducive to life, they were not so extreme or lasting that they degraded RNA communities, or destroyed their compartmentation, nor so sudden that they trapped RNAs in useless intermediates, nor so rare that polymer life froze, and its evolution ceased.

3. Implementing biological processes in RNA

Natural ribozymes, riboswitches, and their RNPs, probed the earlier concepts of polymer folding and catalysis derived from seminal studies of ribonuclease and other small globular proteins (Anfinsen et al 1961; Anfinsen 1973). Self-folding protein domains are sequences of about 25-200 residues that quickly reach a unique thermodynamic minimum in one cooperative transition under physiological conditions. Small changes in temperature, pH, salts cause sharp transitions between a compact random coil and the native fold. Gibbs energy of protein folding is surprisingly modest, balancing the favorable entropy of polypeptide unfolding against the unfavorable entropy of ordering water by solvent exposed hydrophobic residues, and breaking the tie with the favorable enthalpy of hydrogen bonds, ionic bonds, π-π interactions, and van der Walls pockets in the native fold. Meanwhile, Cyrus Levinthal noticed the formidable search complexity of finding a thermodynamic minimum in polypeptide folding space, sampled through local moves from one rotamer to another (Ramachandran et al 1963; Levinthal 1969). [nucleation-condensation]

A number of qualifications to the ribonuclease story of protein folding soon emerged: (1) protein folding commences during translation; (2) various biogenesis factors or chaperones catalyze difficult transitions, or inhibit undesirable alternatives; (3) the folding landscape of globular proteins is highly funneled under physiological conditions; (4) native folds are deep local minima, but not always global ones; and (5) many natural polypeptides are intrinsically disordered sequences that only fold in association with other proteins, nucleic acids, lipids or specific ligands (Leopold et al 1992; 2024 Nobel prize in chemistry). Finally, (6) amyloid fibrils stabilized by hydrogen bonds between β-strand polypeptide backbones proffer a more or less universal alternative to the native folds of proteins (Dobson 2004). The Gibbs energy of this misfolded state is insensitive to amino acid sequence or composition. Moreover, misfolded proteins can seed or catalyze misfolding of other proteins in prion disease and other proteinopathies.

Before the breakout of polypeptide coding, natural selection likely honed the choice of peptide backbone and the pools of predominately α-amino acids with hydrophobic and hydrophilic sidechains (Orgel). With a small number of backbone and sidechain rotamers, these polymers are flexible over long sequences yet highly constrained over short ones. Over a range of environments, *viz*. aqueous, lipid, flexible or structured RNA, random polypeptides of modest length and different compositions were selected by assembly to form α-helices, β-strands, and reverse turns. After the breakout of polypeptide translation, the genetic code converged on the canonical set of some twenty proteinogenic amino acids. In the map from polypeptide sequences to folded proteins, these residues were exquisitely positioned for folding, as well as substrate recognition and catalysis in the folded protein.

Compared to the wide-mouth, deep funnel energetics and fast kinetics of natural protein folding with, natural ribozymes and riboswitches have narrow-mouth, shallow funnels, and slow folding kinetics (Tinoco & Bustamante 1999; Vicens & Kieft 2022). Thus, the energetic landscapes of RNA folding are flatter with multiple paths from any intermediate to its final fold. Whereas globular proteins undergo concerted folding and denaturation, natural RNAs have sequences of folding intermediates and alternative folding pathways. Indeed, many ribozymes undergo partial unfolding and refolding to another conformation as part of their reaction paths under physiological conditions.

Rivaling the complexity of protein folding, nearly a dozen *regular elements*, viz. topological arrangements of duplexes and unpaired segments, and other 2- or 3-strand motifs with significant translational symmetry,[4] and even more *irregular elements*, viz. unique motifs without duplex or other translational symmetry,[5] have been described in RNA folding. Imperfect hairpins, the shortest sequences that quickly find a stable fold under physiological conditions, were the low-hanging fruit of RNA life. Folding of longer RNA molecules is sequential and hierarchical, *viz*. rapidly folding secondary structures define the early intermediates, while tertiary interactions determine the final fold. These

[4] Any list of regular elements would include coaxial helices, simple hairpins, kissing hairpins, multi-helix junctions, pseudoknots, unpaired leaders or trailers, ribose zippers, A-minors, and G-quadruplexes.

[5] Any list of irregular elements would include kink turns, interdigitating T-loops, loop E motifs, ribose zippers, tetraloops/receptors, T-loops/receptors.

larger, potentially more useful folds are nucleated by hairpins, rearranged through strand displacements or branch migrations, and stabilized by tertiary elements that preserve some secondary elements and rearrange others.

Like concerted folding *simpliciter*, concerted reactions with one principal transition state are rare amongst natural ribozymes and RNPs. Compensating for a relative poverty of functional groups, ribozymes undergo partial unfolding, and a different refolding, in sequential reactions that alternate between slow conformational and fast catalytic steps. In a typical productive sequence of docking - catalysis - redocking - catalysis - undocking, substrates and intermediates are brought to and from a common active site, and their products released. To accomplish this, more-or-less rigid parts rotate at flexible hinges, often with reversible strand displacement, or making and breaking of longer-range interactions. Sometimes the rotation of a single nucleobase, changing its stacking and pairing interactions, distinguishes one conformation from the next.

Sequences of conformations and covalent intermediates reach a zenith with the splicosome and ribosome, macromolecular machines of RNA splicing and protein translation, respectively. Indeed a mechanistic view of the ribosome, with internal movements conveying tRNA intermediates from one site to the next for mRNA decoding and peptidyl transfer, predates both the ribozyme concept and atomic-level descriptions of the ribosome (Watson; Bretscher 1968; Spirin 1968, 2009).

The demarcation between stoichiometric and catalytic reactants is blurred for many ribozymes and RNPs. For instance, self-cleaving and self-splicing ribozymes, which undergo single turnover reactions upon themselves, fail the kinetic definition of catalysts as reactants restored unchanged at the end of the process.[6] In the back-and-forth of ribozyme engineering, and likely evolution itself, a single turnover reaction, where the substrate is a covalent extension of one or both ends of the ribozyme, can be turned into a multiple turnover reaction with encounter of free substrates and departure of free products, and *vice versa*. Substrates and catalysts are further confounded in *substrate-assisted catalysis*, where a conventional stoichiometric

[6] There are analogous self-reactions in proteins, including auto-phosphorylation, auto-proteolysis, and self-splicing inteins.

reactant provides an essential element of the active site, complementing the catalyst *simpliciter*, and in *substrate-induced fit*, where the substrate selects and positions the catalyst as much as the converse.

Nearly all natural ribozymes known today act on RNA substrates to make and break phosphodiester bonds by *transesterification,* or substituting one sugar alcohol by another. In a concerted (S_N2) reaction, the ribose O2', O3' or O5' attacks the phosphate with the inline O3' or O5' alcohol leaving. The rates of spontaneous transesterification (as well as hydrolysis) are extremely low, owing to electrostatic repulsion of the nucleophiles from the shared negative charge of the non-bridging oxygens (Westheimer 1987; Kamerlin et al 2013). Absent some significant difference between substrate and product RNAs in bond strain, secondary structure, or tertiary structure, transesterification is nearly isoergonic, driven primarily by mass action.

Acting on themselves, natural ribozymes catalyze single-turnover reactions with just one transesterification, *viz*. self-cleaving ribozymes, or two consecutive transesterifications, *viz*. self-splicing introns and their splicosomal descendents (Kruger et al 1982; Chillon & Marcia 2021; Garside et al 2021; Wilson & Lilly 2021). In the small, self-cleaving ribozymes, an O2' attacks the vicinal phosphorus, making the 2',3' cyclic phosphate with the O5' alcohol leaving. So long as these cleavage products stay associated, the reverse reaction can relieve the strained cyclic phosphate to restore the original phosphodiester bond. Comparisons of self-cleaving ribozymes show that the loss of entropy upon ligation is greater for the flexible *hammerhead ribozyme* than the rigid *hairpin ribozyme* (Nesbitt et al 1999). Indeed, under high salts and low temperature, the latter ribozyme actually favors ligation over cleavage. Once the cleavage products dissociate, however, the strained 2',3' cyclic phosphate eventually opens by hydrolysis.

Self-splicing introns perform consecutive transesterifications at two distinct phosphodiester bonds called the *splice donor* and *acceptor*, respectively. Group I and group II introns differ in the nucleophile of the first transesterification: In group I introns the O3' of a free guanosine attacks the donor site phosphate, with the O3' alcohol leaving (Cech et al 1981). In group II introns the O2' of an adenosine within the intron attacks the donor site, again with the O3' alcohol leaving (ref). As a result, the upstream

exon ends with the 3’ OH, while the downstream product begins with the guanosine, or a lariat branched at the adenosine, respectively. In the second transesterification, the O3’ of the upstream exon attacks the acceptor site phosphate, with the O3’ alcohol leaving. After both transesterifications, upstream and downstream exons are joined by a phosphodiester bond retaining the splice acceptor site phosphorus. The other product is a linear intron, beginning with the attacking guanosine, or a lariat branched on the attacking adenosine, respectively.

RNase P and the ribosome peptidyl transfer center extend the catalytic repertoire of natural ribozymes beyond the *cis*-transesterifications of self-cleaving ribozymes and self-splicing introns. Both of these ribozymes act in *trans*, that is, on substrates that are not covalent extensions of the catalytic RNA, and mediate multiple turnovers with dissociation of products and association of new substrates. RNase P RNA hydrolyzes the phosphodiester backbone of pre-tRNAs and other RNAs (Guerrier-Takada et al 1983). Hydrolysis differs from transesterification in that the attacking oxyanion comes from water, not ribose. Enthalpically, hydrogen is more electropositive than carbon, so that the water is more polarizable than the alcohols. Entropically, the orientation of water is less constrained than the ribose OH, but its effective concentration is higher. Due to the greater chemical activity of water than the leaving sugar alcohol, mass action favors the forward reaction (hydrolysis) over the reverse (condensation) with *provisos* that solvent freely enters and products freely leave the active site.

Compared to proteins, RNAs have few functional groups for chemical catalysis. Chosen for aromatic stacking and hydrogen bonding, none of the four nucleobases is ionized at neutral pH. Away from neutrality, adenine (pK_a 3.5) and cytosine (pK_a 4.2) can be protonated for general base catalysis, while guanine (pK_a 9.4) and uracil (pK_a 9.3) can be deprotonated for general acid catalysis. Alternatively, magnesium cations can activate oxyanions of water or alcohol, or stabilize the developing negative charge on the leaving group. Like metal-independent RNase enzymes, small self-cleaving ribozymes use general acid-base catalysis for transesterification, yielding 2’,3’ cyclophosphate and 5’ OH products. Like metal-dependent RNase enzymes, RNase P uses metal ion catalysis for hydrolysis, yielding 3’ OH and 5’ phosphate products.

The ribosome peptidyl transfer center catalyzes two distinct reactions, both exergonic. During polypeptide elongation it catalyzes peptide bond formation, an aminolysis of the peptidyl-tRNA trading the acylester for acylamide bond. During polypeptide release, it catalyzes hydrolysis of that same acylester bond. Hydrolysis is favored by mass action of water, while aminolysis is favored by the greater stability of acylamide than acylester bond. Like the catalytic site of RNase P RNA, the chemical activity of water in the ribosome transfer center is far below bulk water. These two reactions of the ribosome peptidyl transfer center, hydrolysis and aminolysis of acyl-esters, are unique among natural ribozymes (Noller et al, 1992; Ban et al 2000; Nissen et al 2000).

Two observations suggest that the catalytic repertoire of ancient ribozymes was likely greater than the surviving examples. First, through a back-and-forth of design and selection, artificial ribozymes have been engineered for novel biochemical and organic reactions (Wilson & Szostak 1999; Cojocaru & Unrau 2021; Deng et al 2022). As a caveat, it is generally unknown what substrates were abundant in the RNA world as prebiotic feedstocks or metabolic intermediates. In particular, it is unclear whether nucleoside triphosphates (NTPs), which are the common energy currency of cellular metabolism, as well as the substrates of polymerase enzymes, became available at some stage of RNA life, or only later, in protein life (section 16). Second, some enzyme cofactors, including nucleotide-derived cofactors, are thought to be vestiges of ribozyme cofactors that predate protein life (Eakin 1963; White 1976; Jadjav & Yarus 2002; Goldman & Kacar 2021; Kirschning 2021).

Whereas biochemists reserve the suffix *-ase* for enzymes and ribozymes that make or break covalent bonds, molecular cell biologists have extended the kinetic concept of catalyst, or its generalization in computer science as *reusable instruction*, to any transition between two definite states (section 19). For an easy example, proteinaceous pores facilitate the diffusion of solutes through cell membranes, but are themselves unchanged in the process. Probing the concept further, a rich trove of natural RNAs, many likely ancient, are found to regulate transcription, protein translation and export, as well as creative functions of heredity. For example, in ribosomal decoding, rRNAs and mRNA catalyze the conformational steps that select one aminoacyl-tRNA from the

working set, leading to peptide bond formation. These catalysts are restored to their local initial states (*modulo* one translocon) at the end of each elongation cycle, and to their global initial states at termination of translation. Unsung as ribozymes, the panoply of RNAs that regulate cellular events are variously dubbed adaptors, carriers, guides, messengers, scaffolds, seeds, switches, or templates. Like catalytic centers that provide a general acid, general base, or metal coordination to lower the activation energy of covalent events, these regulatory RNAs instruct, or increase rates of particular transitions, through binding and positioning of other reactants.

Nearly 60 classes of natural riboswitches have been described that use a handful of mechanisms for ligand-dependent RNA folding (Suddala et al 2023; Olenginski et al 2024). They monitor everything from temperature, inorganic ions such as fluoride or magnesium, to amino acids and nucleotide-related signaling molecules such as cAMP, (p)ppGpp, ZTP, c-di-GMP, c-di-AMP, to charge status of specific tRNAs, or hydrophobic polypeptides exiting the ribosome (Nelson & Breaker 2017; Pavlova et al 2019; Kavita & Breaker 2022). Changing conformation by ligand-dependent refolding, riboswitches regulate options to terminate or continue, as well as choices between forward alternatives, within multi-step molecular processes.

The most kinetically curious, and evolutionarily profound, divide in molecular biology is between regular polymers, such as polypeptide antibiotics or branched-chain oligosaccharides, and hereditary polymers, or evolvable genes. Enzymatic pathways, catalytic realizations of the Kleene's regular operations, *viz*. composition, selection, and iteration, build complex polymers of fixed or indefinite size (Kleene 1952; Moore 1956). In the formal language of processes, each polymer is a run of reactions through the corresponding sequence of intermediates. Any set of enzymes determines a regular set of one or more polymers. These enzymes are themselves evolvable polymers, and thus the regular sets of products can evolve, but these derived polymers have no lineage themselves.

As is well known, the evolutionary status of enzymes, and coded polypeptides generally, is a little more elevated than regular polymers (Crick). Virtually any sequence of proteinogenic amino acids can be made by ribosomal translation of an mRNA. Like nonribosomal polypeptides and other regular polymers, coded polypeptides have no

direct descendants. However, the translation of mRNAs into polypeptides using the genetic codes projects the lineages of these polynucleotides, and their evolvable deoxyribogenes, onto *faux* lineages of their polypeptide products.

Before the emergence of self-folding proteins, one immediate application of polypeptide coding was custom fitting of otherwise disordered polypeptides to fold with, and thus recognize, small ligands, structured polynucleotides, or other polypeptides. In cellular life, molecular complementarity, which Emil Fischer likened to the way a key fits its lock, depends on well-placed sidechains within the protein fold as a whole. For protein domains or RNA aptamers, complementarity is a peculiar one-off relation between the molecular key and its evolvable lock. This is even true where the key is an evolvable protein itself. The importance and difficulty of molecular complementarity in protein recognition is highlighted by the evolution of peptide-guided adaptive immunity, a means of matching protein locks with protein or other molecular keys on the physiological, rather than evolutionary timescale. This remarkably roundabout mechanism of fitting say an antibody to its cognate antigen, entails proteolysis of proteins, and selecting cognate polypeptides on MHC proteins for presentation to cognate T-cell receptors.

As Watson and Crick discovered, the problem of molecular complementarity is greatly simplified when both the lock and its key are polynucleotides (Watson & Crick 1953a,b). Their model of polynucleotide duplexes identified a stereotyped form of complementarity wherein: (1) any polymer has high affinity for a unique complementary polymer, and (2) this affinity is the sum of affinities between nucleotide pairs, so that Gibbs energy comes mainly from the hydrogen-bonded partners, smaller contributions from their immediate neighbors, and very little from more distant sequences.

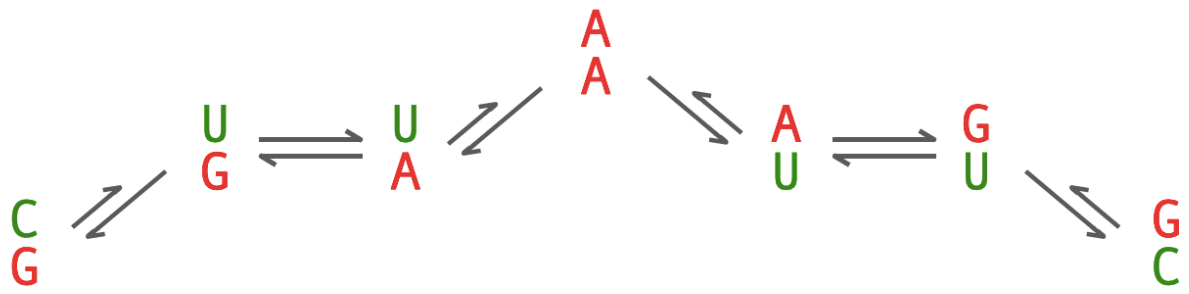


FIGURE 3-1. LOWEST GIBBS ENERGY PATH BETWEEN COVARIANT HELICES

In polynucleotide duplexes, systematic changes, or *covariation*, in both sequences of a helical segment create a family of lock and key combinations. This is used for the rapid prototyping of guide sequences in biological regulation. Certain non-canonical basepairs with isosteric geometry, folding kinetics, and thermodynamic stability near to cognate Watson-Crick pairs, provide paths through the fitness landscape via helical intermediates whose stability, and oftentimes fitness, are near those the original and final duplexes (Figure 3-1). This is useful not only for strictly helical elements, but when the paired sequence can escape a purely helical structure to greater fitness with a specialized bulge, kink, or other sequence-dependent interaction. In RNA lineages, covariation, the imprimatur of paired secondary structure, appears as two sequences in either the same RNA, or in two different RNAs, whose substitutions go hand-in-hand so to preserve their basepairing. Thus, the sequence space is explored without going through intermediates that overly destabilize, or overly stabilize the helix itself. To wit, GC and CG basepairs can be interconverted to AU and UA basepairs, respectively, via their GU and UG wobble pairs without significant disruption of helix stability or distortion of helix geometry. At somewhat greater Gibbs energy, the transversion of AU to UA can proceed through the AA intermediate. Thus, the optimal sequence of a paired region, say for toehold interactions with a third sequence, can be explored through paired intermediates that preserve the stability and geometry of the helix.

Today all of biology can be understood as a hierarchy of search processes on a range of timescales from molecular cellular biology, to development and behavior, to reproduction and evolution itself (sections 12 & 18). The distinction between gene expression and genome replication became clear in later stages of polymer life. Ribogenes and deoxyribogenes template their own reproduction, and thus, have direct descendants, and by iteration, an evolutionary lineage with heritable variation. On shorter timescales of physiology and conservative replication, duplexes formed by Watson-Crick pairing provide a molecular means for storing and copying fixed instructions preserved as sequence information. On longer timescales of creative replication and evolution, these genes are a dynamic data store allowing not just

regular (memoryless) processes, but molecular processes with true memory. With the inclusion of these reprogrammable elements, there was a not just a compiled program to execute, but the means to modify and explore the program itself.

Anticipating our history of polymer life (sections 8-18), we précis the control of gene expression in cellular life, notably regulatory mechanisms based on the ribosome and natural riboswitches (Grundy & Henkin 2006). In cellular metabolism, catabolic pathways are UP-regulated by initial substrates, while anabolic pathways are DOWN-regulated by final products. Activation or inhibition can occur at any level from gene transcription to protein translation to enzyme binding of small effector molecules. In prokarya, transcription and translation are closely coordinated in space and time, *viz*. the DdRP complex transcribing mRNA 5' to 3' from its deoxyribogene is followed by one or more ribosomes translating that same mRNA 5' to 3' into polypeptides. During eukaryogenesis, parasitic group II introns from the mitochondrial endosymbiont drove evolution of the nuclear envelop to insulate the slow process of pre-mRNA splicing in the nucleus, from the fast process of mRNA translation in the cytoplasm (Koonin).

Natural riboswitches regulate metabolic pathways at the levels of transcription or translation in response to small effector molecules. Most riboswitches are integral parts of mRNA leader sequences, and thus, regulate expression of a single gene. From the perspective of the nascent mRNA molecule, the option to continue transcription (anti-termination) can be exercised only once, whereas the option to initiate translation can be exercised by a succession of ribosomes on the same mRNA. Although any *cis*-acting riboswitch regulates only one gene, similar riboswitches found in other genes, allow genome-wide coordination of gene expression in response to the same effector molecule.

Besides *cis*-acting riboswitches found in mRNA leaders, some *trans*-acting regulatory RNAs and RNPs are bistable riboswitches in all but name. These include the SRP Alu domain that pauses ribosome elongation in response to signals from the nascent polypeptide, and a variety of mechanisms based on *trans*-acting guide RNAs (sections 16-18). Small guide RNAs that recognize simple sequences through base-pairing mediate and coordinate genome-wide regulation of deoxyribogenes and their RNA transcripts. The differences between *cis*- or *trans*-action, and between aptamer-

based or sequence-based recognition, affect the portability and evolvability of these regulatory mechanisms.

4. RNA copying | ligase & polymerase ribozymes

At the dawn of RNA life, four types of ribozymes were auspicious discoveries: First, activities that improved the yields of nucleosides and random oligonucleotides from prebiotic feedstocks, or salvaged hitherto dead-end molecules as metabolic intermediates. Second, activities that improved the generality, fidelity or speed of RNA copying, as well as protected or repaired existing polynucleotides. Third, activities that improved genomic or metabolic compartmentation by the selection of self RNAs, rejection of nonself RNAs, concentration of feedstocks, or diffusion of wastes. Finally, no one RNA community could acquire all useful innovations through purely vertical transmission of in-house discoveries. Augmenting abiotic processes of polynucleotide admixture, RNA communities joined a deliberate race to garner useful ribogenes from the pangenome to fit into their regulatory networks. Beyond conservative functions of heredity, viz. faithful vertical transmission, there was strong selection for ribozymes that joined together or broke apart polynucleotides in ways likely to create useful variation for natural selection.

Today we associate polynucleotide copying with formation of the *reverse complement* as the principal intermediate, or the final product. Processive polymerase enzymes of cells and viruses read the template from 3’ to 5’ in one nucleotide steps, and make the reverse complement from 5’ to 3’ by one nucleotide additions. Besides *replicative polymerases* that copy entire chromosomes, *repair polymerases* copy shorter regions. Besides these promiscuous polymerases that replicate or repair diverse templates, there are exclusive polymerases, such as *telomerase*, that copy just one guide sequence. Finally, besides those various template-directed polymerases, *terminal nucleotidyl transferases* extend 3’ ends of polynucleotides with tails of various lengths, nucleotide composition, and even simple sequence, without any polynucleotide template. Nearly all schemes for spontaneous RNA copying also create long duplexes as principal intermediates, suggesting a continuity of products, if not substrates and mechanism, from the dawn of life to the present.

Inspired by Campbell (1991) and others, in section 8 we propose a ribozymatic scheme of RNA copying that duplicates the template in one round of copying without a

reverse complement, or other minus-strand intermediate. Here we introduce some terminology to compare this larger universe of copying schemes. Our first question is how the catalyst moves from one site to another on the template. There is a continuum of schemes between (1) free diffusion of dissociated catalyst in solution, *e.g.*, templated-ligation of large oligonucleotides, (2) constrained diffusion or sliding of the catalyst along the template in either direction, (3) scanning the template in one direction, and (4) processive translocation in definite steps, *e.g.*, primer 3' extension by single nucleotide additions (refs). Processive copyases may have ancillary activities to unwind or otherwise bypass obstacles to translocation and copying.

In processive RNA copying, the polarity of the nascent strand is described in either absolute terms (5' to 3', 3' to 5') or relative to the template (forward, reverse). For example, polymerase enzymes read the template from 3' to 5' and make the copy from 5' to 3', reversing the template polarity. In sections 5-8, we consider schemes that read the template from 5' to 3' as in ribosomal translation of mRNA, and make the copy from either 3' to 5' (reverse), or 5' to 3' (forward). We call the copy a *duplicate* or a *complement* if its nucleobases are the same as the template, or are their Watson-Crick partners, respectively. In this jargon, we say that polymerase enzymes make the reverse complement in the first round of replication, and recreate the forward duplicate in a second round. Whereas most of the copying schemes we discuss are *general* and *faithful*, or more formally, the identity map on all possible template sequences, we point out those few that are only partially defined, or are multivalued maps (aka *degenerate*).

Discovering that cells and viruses use enzymes to replicate and repair their DNA, Arthur Kornberg and colleagues characterized their DNA-directed DNA polymerases, or DdDPs (see Kornberg 1969). Given all four dNTPs, these enzymes can extend a primer on any unique template, reading it from 3' to 5' in one nucleotide steps, while making a complementary copy from 5' to 3' by one nucleotide additions. In an inline S_N2 reaction, the 3' OH of the nascent polynucleotide attacks the dNTP α-phosphate with pyrophosphate leaving. Completing the elongation cycle, the polymerase moves one nucleotide along the template before the next read-add step (refs). Another round of copying recreates the original template.

The paradigm of enzymatic DNA replication was consolidated by discoveries of other template-directed polymerases (Watson et al 2016). Given all four NTPs, DdRPs transcribe duplex DNA into mRNAs or non-coding RNAs (Burma et al 1961; Geiduschek et al 1961; Stevens 1961; Chamberlin & Berg 1962; Furth et al 1962). RdRPs replicate the ribogenome of RNA viruses and viroids, as well as amplify ##RNAs in RNAi (Fraenkel-Conrat 1956; Gierer & Schramm 1956; Haruna et al 1963; Diener 1971; Flores et al 2004; ref). And lastly, RdDPs reverse-transcribe RNAs into DNA in the replication of retroviruses, retrotransposons, and telomeres (Baltimore 1970; Temin & Mizutani 1970; Greider & Blackburn 1989; Collins & Greider 1993). The processivity, speed, and accuracy of polynucleotide copying, including strand separation of duplex template, is driven by Gibbs energy of nucleotidyl transfer followed by pyrophosphate hydrolysis. During transfer, one phosphoester bond is formed at the expense of one phosphoanhydride bond. During hydrolysis, a second phosphoanhydride bond is lost to water. Beyond polymerase enzymes *simpliciter*, a handful of steps to complete the replication cycle engaged molecular biologists for years to come. Any short list would include template selection, initiation including primer synthesis and annealing, termination including end modifications, unwinding duplex templates and separating template and copy strands, supercoiling and linking, and bidirectional replication of leading and lagging strands.

With the central dogma of DNA replication, RNA transcription, and protein translation in place, molecular biologists asked how these biopolymers and their polymerases evolved. Manfred Eigen noticed two constraints on replication under purifying selection, one on polynucleotide length, and another on sequence fidelity (Eigen 1971; Eigen & Schuster 1979). If the genome were not to degenerate in absolute length, the rate of replication must exceed the rate of decomposition. If it were not to degenerate in sequence information, each generation must make at least one error-free copy. A simple argument showed that the effective sequence length sustainable under purifying selection is about k^{-1} where the *error rate* k is measured per nucleotide per replication. In light of these constraints, John Maynard-Smith noticed the *Catch-22* in the evolution of replication, *viz*. for a polymerase gene to replicate accurately enough to evolve, a polymerase enzyme must first evolve to replicate it (Maynard-Smith 1983).

Studying *in vitro* replication of RNA viruses, Solomon Spiegelman and colleagues showed that viral polymerase enzymes (1) favor their native template over other RNA or DNA templates, (2) faithfully replicate rare sequence variants, and (3) make copies of copies *ad nauseum* (Mills et al 1967; Kacian et al 1972). With Qβ polymerase, the only viral protein required for *in vitro* replication, supplied by the investigators, viral RNA was merely a template for copying, not an mRNA constrained to encode a useful protein. Absent natural selection for full-length functional proteins needed for infective particles, viral RNA tolerated any mutations that did not disrupt polymerase recognition, priming or replication. Selecting for faster replication in serial transfer experiments, shorter sequences with enhanced recognition, dubbed “little monsters”, soon took over the RNA population.

Extrapolating the idea of relaxed selection to living cells, Richard Dawkins popularized the idea of selfish genes, *viz*. free-loaders that compete for replication within the genome, but make no apparent contribution to fitness of the community as a whole (Dawkins 1976). In both prokarya and eukarya, repeated sequences found dispersed throughout the genome were explained as mobile parasites that multiply within a genome, and invade related hosts, or more rarely, distant ones (*cf*. Doolittle & Sapienza 1980; Orgel & Crick 1980). From this new perspective, the natural state of the genome was a Hobbesian war of all genes against all. Success of selfish sequences in the intracommunity competition comes at a cost to the community as a whole in its competition with other communities.

Species vary widely in their caution and curiosity about other species, or the balance of genome conservation (VGT) and creativity (HGT). However, no species can succeed without continual tinkering and occasional borrowing. Prokarya have lean-and-mean core genomes with relatively little memory to spare for in-house R & D. The immediate advantages of exploiting known affordances, as well as emergent advantages from exploring unknown ones, drive species to partake ergodically of the pangenome. Still each gamble to obtain a useful operon carries unknown risks of welcoming a Trojan horse, or opening a genomic pandora’s box.

In the emergence of eukarya, the evolutionary search strategy shifted from a small core genome and large pangenome, to *sexual reproduction*, a form of controlled mixis

between large core genomes of the same biological species. In the sexual species, really a new middle level of selection, a population of genomes explores concurrently, and useful innovations are readily brought together in one genome, and eventually fixed in the population. In this staid Mendelian life cycle of meiosis and fertilization, alleles of each locus engage in tourneys mediated by chromosome pairing, exchange and segregation. Unlike free mixis, alleles of one locus are in direct (Mendelian) competition with each other, but poised for cooperation with alleles of other loci.

Whereas genomes of prokarya are gamed by *mobile selfish elements*, those of eukarya are gamed as well by *sessile selfish cliques.* These self-dealing cliques are intra-chromosomal, inter-chromosomal, or even inter-genomic conspiracies among non-contiguous loci, *viz*. B-chromosomes, meiotic drive, nuclear-mitochondrial incompatibility, segregation distortion, *etc* (refs). Known as the *biological species concept*, the evolutionary signature of sexual reproduction is abrupt speciation, allowing winner-take-all competition between sister species. These macroevolutionary face-offs between closely related species enforce the Mendelian compact against gangs of genome cheats who are otherwise favored by purely microevolutionary competition within the sexual population.[7] Despite the universality and regularity of speciation, the genetic barriers between sister species appear arbitrary, and as diverse as biology itself, with no one mechanism of reproductive isolation, whether pre- or post-zygotic, predominant. No doubt, more predictable barriers would present an open invitation to form selfish cliques in the incipient species, or what economists call insider trading.

---

[7] There are really two macroevolutionary roles for reproductive isolation between sister species: On the one hand, these barriers ensure that when cliques or parasites up-end the genome and reduce its fitness, the uninfected sister species can continue the lineage. On the other hand, two sister species can specialize apart, exploring genome variations that are mainly advantagous at the group, not individual level, that is, have strongly positive-frequency dependent selection. In what is called *ecological speciation*, sister species diverge in preferred habitat, food source, pollinators, etc., or life history traits such as flowering time, etc. Finally, besides genes that increase fitness at the species level, genes that increase fitness at the individual level can contribute to macroevolution. This obviously applies to alleles that predate the speciation event, but are not yet fixed, that continue their tourneys within each species, as well as contribute to the macroevolutionary success of one sister species. Less obviously, it applies to alleles acquired by one sister branch, but not the other, through HGT soon or long after the speciation event.

Two contrasting approaches emerged in the search for tractable models of the evolution of polymer life: In the more eclectic approach, introduced by Eigen, they considered a community of polymers of one or more types with coupled cycles of reproduction. In the more focused approach, associated with the RNA world, they considered this one type of polymer, and even one ribozyme, capable of catalyzing its own replication (Pace & Marsh 1985; Sharp 1985; Gilbert 1986; Orgel 1986). Like the polymerase enzymes of cells and viruses, this ribozyme was a general catalyst of genome replication, including itself as an unfolded template, or ribogene. In the late 1980s there were three main arguments for an extinct copyase ribozyme: First, there was no clear alternative to a ligase or polymerase for RNA life moving beyond the likely limits of spontaneous copying, in speed, processivity, generality, and especially fidelity. Second, there were clear hints that RNA had greater catalytic versatility than the two then known natural ribozymes, *viz*. self-splicing introns and RNase P. Finally, there were suggestions that self-splicing introns themselves might provide the basis for template-directed copying (Cech 1986; Doudna & Szostak 1989).

To reincarnate RNA life, organic chemists sought to recreate its elusive prebiotic feedstock of random oligonucleotides and spontaneous process of polynucleotide copying (Joyce & Orgel 1993; Orgel 2003, 2004). Meanwhile, molecular biologists sought to exhume, design or select a ribozymatic process of polynucleotide copying (Szostak 2012). Just where to start, and what to look for in terms of substrates and products, was deeply uncertain: thus, (1) seeking a palimpsest of the lost copyase, some focused on the origins and functions of natural ribozymes, riboswitches and RNPs; (2) surmising this copyase ribozyme has no extant descendants, others abandoned homology and parsimony, settling for some semblance of biochemical continuity of the stoichiometric reactants with cellular metabolism; (3) finally, some turned from the familiar biochemical reactions of cellular metabolism under physiological conditions, entirely, toward activated forms of organic matter and non-equilibrium enivronments known from geochemistry and astrophysics. Whether they modeled modern metabolic reactions or conjectural prebiotic reactions, the latter two camps entertained artificial ribozymes as proof of concept.

Though neither process is yet known, we shall assume that lost ribozymes of the RNA world produced longer and more accurate products than spontaneous RNA copying. In any RNA world path from spontaneous to ribozymatic copying, diffusive ligation was a likely intermediate stage, or even a final alternative to processive addition. Like spontaneous copying, the lost template-directed ligase likely made little distinction between the complementary strands. Whatever the intermediate stages, over time copyase ribozymes overcame constraints on their templates, as well as degeneracy, or other errors in their products. As ribozymatic copying became more general and more faithful, these catalysts doubtless became more selective for replication of self ribogenes. A clear distinction between full-length templates and nascent copies emerged with processive RNA copying, allowing differential replication of plus- and minus-strands.

In the quest for a copyase ribozyme, Thomas Cech and colleagues turned to the group I self-splicing intron (Cech 1986; Zaug & Cech 1986a, 1986b; Been & Cech 1988). They modeled a template-directed elongation cycle as transesterifcation of the terminal nucleotide of a feedstock oligonucleotide to the 3' OH of the nascent polynucleotide with the oligonucleotide leader leaving, now shortened by one nucleotide. Some proximate source of Gibbs energy was still missing to turn these back-and-forth transesterifications into a processive elongation cycle with unidirectional docking-transfer-undocking-translocation. Meanwhile, Jack Szostak and colleagues engineered the group I self-splicing intron for template-directed ligation of short oligonucleotides on a complementary strand (Doudna & Szostak 1989). Remarkably, their ribozyme, cleverly designed with pencil and paper, was further improved by selection *in vitro* (Green & Szostak 1992). Rather than start with a self-splicing intron, or another natural ribozyme, these investigators now selected an RNA ligase ribozyme *ab initio* that uses oligonucleotide 5' triphosphates as substrates and forms 3'-to-5' phosphodiester linkages (Bartel & Szostak 1993; Ekland et al 1995). Further engineering and selection *in vitro* yielded a polymerase ribozyme using nucleoside triphosphates whose overall reaction, if not mechanistic steps, resembled natural RNA polymerase enzymes (Johnston et al 2001). [Hobert email PMID]

Like the investigations of spontaneous copying, early investigations of copyase ribozymes identified a panoply of obstacles to the speed, processivity, fidelity and generality of copying. A functional copyase would need to unwind secondary structure in the template, to separate the copy from the template, to untether the template from the ribozyme, and perhaps, to initiate and terminate copying at specific sites. Some saw these simply as open research problems, while others emphasized the difficulty or even impossibility of their solution and called for radical alternatives (Crick & Orgel 1973; Joyce & Orgel 1993; Robertson & Joyce 2012; Hud et al 2013; Fialho et al 2020).

Thirty years of research on the artificial class I ligase have shown us something of what an RNA polymerase ribozyme can do (Wachowius & Holliger 2021). Descendants of the original ligase ribozyme today can recognize a general primer-template helix in *trans*, adding hundreds of nucleotides in 3'- 5' linkages with fair speed and accuracy. While the problem of a processive elongation cycle, within the larger cycle of initiation and termination, remains difficult, there are few firm conclusions about what a polymerase ribozyme can never do. Beyond incremental improvements, there are unexplored avenues, and doubtless unimagined ones, including novel non-equilibrium environments, such as freezing and thawing in icy brines, novel mineral and organic cofactors, as well as novel substrates and reactions, such as trinucleotide triphosphate additions (Le Vay & Mutschler 2019). All in all, the powers of RNA catalysis are considerable, and the uncertainties of framing the problem even greater. We cannot expect a few decades of laboratory experiments to recapitulate the explorations of 200,000,000 years of RNA life. *All of this invites a healthy optimism about the ultimate liklihood of ribozymatic RNA copying, and an equally healthy agnosticism about its eventual form.*

In the past several decades of exhuming or reinventing ribozymatic RNA copying, the argument for this missing-link of RNA life from the biochemical virtuosity of polyribonucleotides has gotten considerably stronger. These include discoveries of natural RNA-guided mechanism of recombination, genome defense, and gene regulation, as well as successful design and selection of artificial ribozymes that catalyze RNA-guided reactions from RNA copying to peptide synthesis. Meanwhile, studies of protein translation reveal that the ribosome is a natural ribozyme for RNA-

directed polymer elongation. In sections 5-7 we précis earlier conjectures that the ribosome and tRNAs descend from a lost process of RNA copying. In sections 8-12 we present our scheme of processive RNA copying by this lost duplisome and its dRNAs. In sections 13-18 we trace the saga of ribosome life, and the legacy of the duplisome.

5. RNA copying | permuted reverse duplicate

The RNA world poses two great mysteries for comparative genomics. One is a *mystery of absence*, *viz*. something needed yet not clearly found, and the other a *mystery of presence*, *viz*. something found yet not clearly needed. Notwithstanding self-splicing introns and nuclear splicosomes for RNA recombination, nor various RNA-guided enzymes for RNA (DNA) replication, cleavage, modification and regulation, there is no clear vestige of the RNA-directed ligase or polymerase ribozyme used for RNA copying. To us and others, it seems untenable that this central ribozymatic process of RNA life vanished without a trace. Meanwhile, the rRNAs and tRNAs of protein translation common to all cellular life have no clear progenitors in the RNA world. It seems untenable that this central ribozymatic process of protein life sprang from nowhere.

To get around the need today for coded proteins in protein translation, molecular biologists mooted bold schemes of polypeptide translation predating the ribosome based on direct interactions of peptidyl- and aminoacyl-tRNAs brought together by codon-anticodon pairing with mRNA. Peptidyl transfer is kinetically facile, requiring only a fair approximation of nucleophile and electrophile, and peptide bond formation is energetically favorable, providing a proximate source of Gibbs energy to drive processive translation. Modifying Carl Woese's *reciprocating ratchet* model of translocation, Crick and colleagues proposed that the primitive genetic code had 8 codons of the form RRY read by 8 tRNA isoacceptors with anticodon loops of the form 3' UGYYRUU 5' (Woese 1970; Crick et al 1976; Eigen & Schuster 1979). But Woese noticed there was no mechanochemical coupling of peptide bond formation to translocation, without which his mechanism was not a unidirectional ratchet, but merely reversible strand displacement (Woese 1980; Weiss & Cherry 1993).

If ribosomes were truly indispensable for polypeptide translation, a bold alternative was that early ribosomes, like their tRNAs, were made entirely of RNA (Woese 1967; Crick 1968; Orgel 1968). "If indeed rRNA and tRNA were essential parts of the primitive machinery, one naturally asks how much protein, if any, was then needed. It is tempting to wonder if the primitive ribosome could have been made entirely of RNA. Some parts

of the structure, for example the presumed polymerase, may now be protein, having been replaced because a protein could do the job with greater precision. Other parts may not have been necessary then, since primitive protein synthesis may have been rather inefficient and inaccurate. Without a more detailed knowledge of the structure of present day ribosomes it is difficult to make an informed guess" (Crick 1968 p371).

Mostly ending the quest for ribosome-free mechanisms of polypeptide translation, there is now ample evidence that the ribosome decoding and peptidyl transfer centers are themselves ribozymes formed from the small and large subunit rRNAs, respectively (Noller et al 1992; Moore & Steitz 2002). Moreover, there is a fair case that bacterial T-box riboswitches are vestiges of extinct tRNA charging ribozymes (Ishida et al 2020; Lu et al 2024). But all of this evidence for translation mediated by primitive rRNAs, tRNAs and other ribozymes *without assistance of coded proteins* merely whets the appetite for an explanation of just what the progenitors of these RNAs were doing *before the breakout of polypeptide translation*.

In "An RNA replisome as the ancestor of the ribosome" John Howland Campbell (1938-2021) at UCLA David Geffen School of Medicine dissolved both mysteries of RNA life, the disappearance of polynucleotide replication ribozymes without descendants, and the appearance of polypeptide translation ribozymes without ancestors (Campbell 1991). There was no ancient polymerase ribozyme to exhume, he suggested, but an entirely different mechanism of processive RNA copying that read the template from 5' to 3' in *triplet codons*, while adding *triplet duplicons* to a nascent copy from 3' to 5'. Unlike enzymatic DNA replication, or RNA transcription, where one and the same (d)NTP that reads the template is immediately added to the nascent product, in this scheme of ribozymatic RNA copying, as in protein translation, adaptor RNAs separated the template from its product, and hence, reading from addition, in both space and time.

Campbell's proposal was the first of a half dozen schemes for RNA copying based on conjectural ancestors of the ribosome and its transfer RNAs (Weiss & Cherry 1993; Gordon 1995; Poole et al 1998; Yakhnin 2007; Noller 2010/2012). Here we adopt a uniform terminology to describe those published schemes, and to compare them to our own. To avoid confusion with the ribosome and tRNAs of protein translation, we

refer to their proposed progenitors used in RNA replication as the *duplisome* and *donor RNAs (dRNAs)*, respectively. In all of these copying schemes, the template is read from 5’ to 3’ in one *codon* steps by matching to the dRNA *anticodon*, while the nascent product is lengthened by one *duplicon* additions.

The size and location of the duplicon vary with each scheme of RNA copying, affecting all aspects of the elongation cycle from dRNA loading to decoding and nucleotidyl transfer to translocation. In a family of schemes that make the reverse complement, the same triplet acts first as anticodon, and then as duplicon (Weiss & Cherry 1993; Gordon 1995; Poole et al 1998). Like the replication cycle of polymerase enzymes, one round of copying creates a reverse complement of the template, while a second round recreates the original polynucleotide. In another family of schemes that make a duplicate without an intermediate complement, both rounds of Watson-Crick pairing occur *within the polynucleotide elongation cycle itself*, one in loading the dRNA with a cognate duplicon, and the other in decoding the template RNA. In two such schemes the duplicon and its complementary anticodon are proper parts of the dRNA (Campbell 1991; present paper), while in the third scheme, the duplicon is a free trinucleotide (Noller 2010/2012). *Importantly, whereas Campbell’s duplicon comprises the last three nucleotides of the loaded dRNA, our duplicon comprises the first two nucleotides of the dRNA.* Finally, the *translocon* comprises the coaxial helices of template and dRNAs translocated within the duplisome; its distance moved is either three nucleotides as in protein translation, or just two nucleotides (present paper).

Era	Polynucleotides	Polypeptides
early RNA life	spontaneous copying	random
late RNA life	duplisome	random
polypeptide life	duplisome →	ribosome
protein life	RdRP enzyme	ribosome

TABLE 5-1. RIBOSOME LIFE

Like the discovery that birds are living descendants of dinosaurs hidden in plain sight, Campbell proposed that ribosomes are surviving descendants of a lost RNA replisome hidden in plain sight. Table 5-1 lists the major stages of ribosome life: First, in late RNA life, spontaneous RNA copying *without support of evolved ribozymes* was replaced by processive copying mediated by the duplisome and its dRNAs. After the breakout of polypeptide translation, RNA duplication continued alongside translation. In the parlance of synthetic biology, the late duplisome and early ribosome were orthogonal polymerases that copied or translated template RNAs, respectively. Finally, in late protein life, the duplisome and its dRNAs were retired in favor of an RNA-directed RNA polymerase (RdRP) enzyme. By then, the ribosome and its tRNAs were wildly successful exaptations ensconced in their roles of protein translation, and increasingly reliant on coded proteins, *viz*. ribosome (tRNA) biogenesis factors and RNA modification enzymes, ribosomal proteins, aminoacyl tRNA synthetases, translation factors, and signal recognition particle.

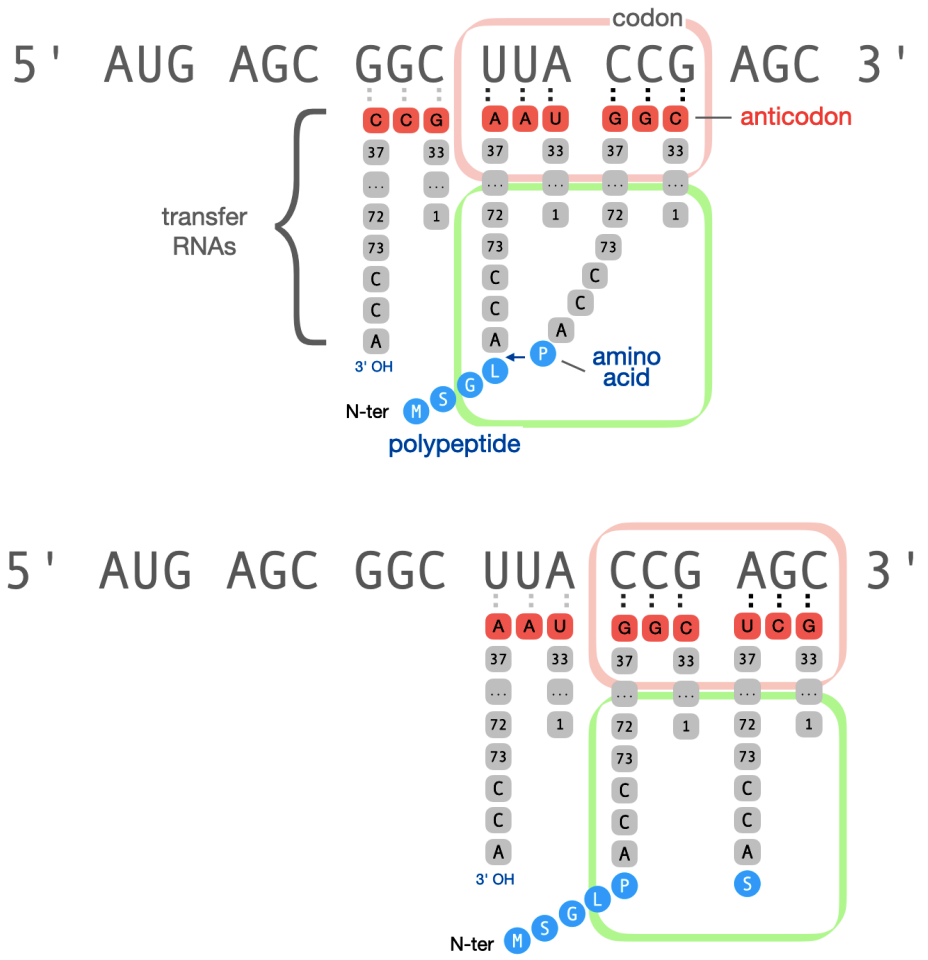


FIGURE 5-1. POLYPEPTIDE ELONGATION CYCLE

Figure 5-1 depicts one cycle of polypeptide elongation in the ribosome. In the upper panel, the α-amine of the aminoacyl-tRNA in the A-site attacks the acylester of the peptidyl-tRNA in the P-site, transferring its nascent polypeptide chain. In the lower panel, the ribosome has moved along the mRNA to the next codon and a new aminoacyl-tRNA has entered the decoding center and been accommodated. Ghosts of deacyl-tRNAs that left the E-site are shown for reference. Figure 5-2 depicts one cycle of polynucleotide elongation in Campbell's scheme of RNA copying. The template is read from 5' to 3' in codon triplets by the duplisome. The duplicon comprises the last three nucleotides of the dRNA, which Campbell suggested might be ancestral to the universal $_{74}$CCA$_{76}$ at the 3' end of mature tRNAs. For nucleotidyl transfer, he implied that O3' of the duplicon-dRNA attacks the polynucleotidyl-dRNA, with freed dRNA

leaving. Thus, the product is made from 3’ to 5’ by addition of duplicon triplets. Ghosts of freed dRNAs that left the E-site are shown for reference.

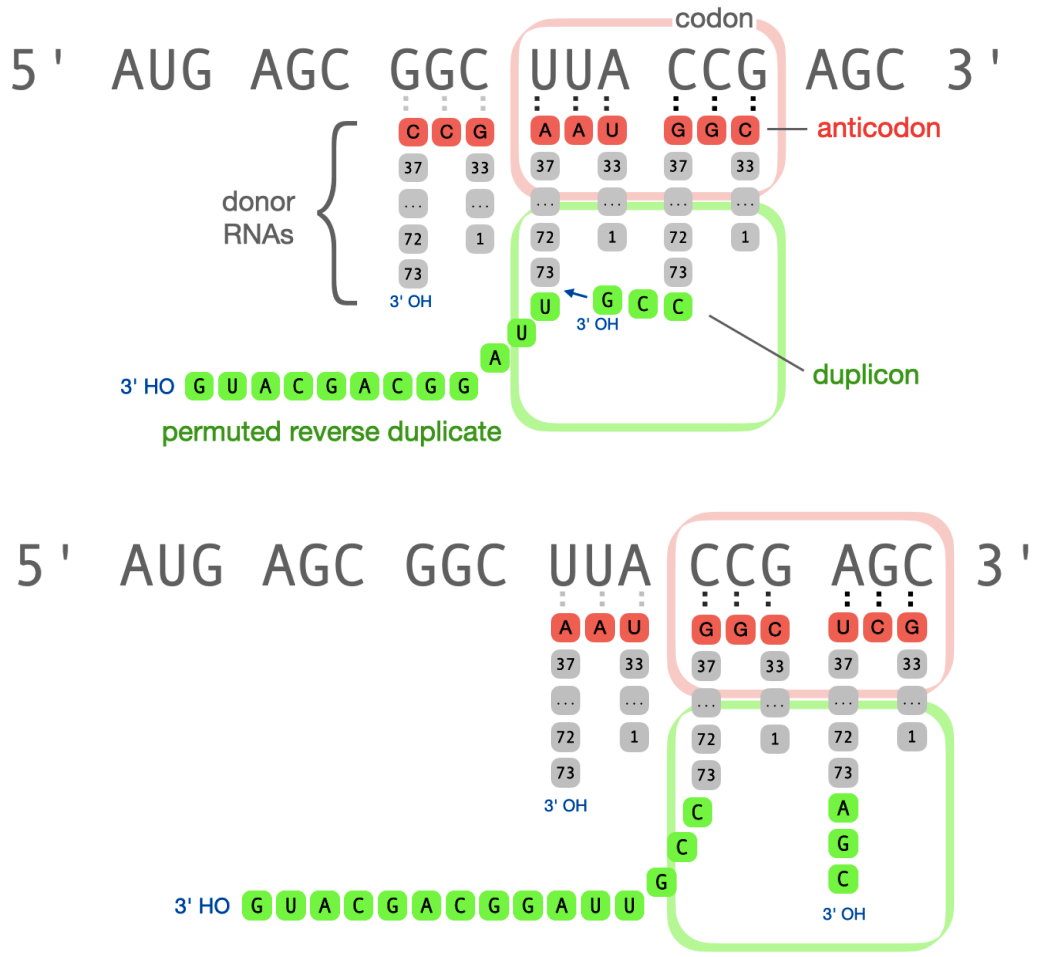


FIGURE 5-2. POLYNUCLEOTIDE ELONGATION CYCLE (CAMPBELL 1991)

Inspection of Campbell’s scheme of RNA duplication shows three serious problems. The first problem is the baroque structure of the copy itself (Figure 5-3). This *permuted reverse duplicate* is neither a forward duplicate identical to the template, nor the familiar reverse complement of polymerase enzymes, but a reverse duplicate that transposes first and third nucleotides of each codon. One virtue, Campbell noticed, is that templates cannot anneal to such copies, and hence, the long duplexes formed in spontaneous RNA copying neither arise during duplication, nor afterward. Like the familiar reverse complement, however, the permuted reverse duplicate still needs a second round of copying to recreate its original template.

```
5' AUG AGC GGC UUA CCG AGC 3'  template (+)
3' GUA CGA CGG AUU GCC CGA 5'  permuted reverse duplicate (-)
5' AUG AGC GGC UUA CCG AGC 3'  copy of copy (+)
```

FIGURE 5-3. PERMUTED REVERSE DUPLICATE

The second problem is that the cycle still needs some means to repeatedly and accurately reload dRNAs with duplicons consumed in polynucleotide elongation. Campbell himself only hinted that the duplicon is somehow self-loaded onto the dRNA from its anticodon. Any scheme of anticodon-directed dRNA loading presents two distinct subproblems, *viz*. placing the anticodon into spatial approximation with the duplicon, and the chemistry of loading. Allowing that the proximity problem can be solved, to avoid a regression of Matryoshka dolls, any plausible chemistry of anticodon-directed dRNA loading must be strictly simpler than the overall copying scheme. That is, we cannot invoke some unknown, general mechanism of template-directed ligation or primer extension *as this is the whole point of the duplisome and its dRNAs as a polynucleotide copyase*. The final problem is that templates require as many as 64 different dRNA isoacceptors for faithful copying, one for each codon triplet. This is significantly more than the number of tRNA isoacceptors in modern genetic codes.

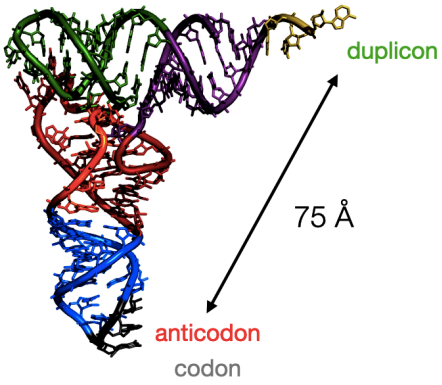


FIGURE 5-4. SPATIAL SEPARATION BETWEEN ANTICODON AND DUPLICON

Like peptidyl transfer in the ribosome, nucleotidyl transfer in the duplisome entails spatial separation of the anticodon in the decoding center from the duplicon in the

polymer transfer center (Figure 5-4). Like tRNA charging outside the ribosome, dRNA loading outside the duplisome poses a molecular 'recognition at a distance' problem, *viz*. how to match a duplicon to the anticodon at another end of the molecule. For tRNA charging, some 20 aminoacyl tRNA synthetase (aaRS) enzymes recognize one or more tRNA isoacceptors, charging the 3' end of their acceptor arms with the cognate amino acid. To recognize and charge their tRNA substrates, these enzymes must span the distance from the 3' end of the acceptor arm to determinants of tRNA identity scattered at various sites along the acceptor stem, the elbow and variable arm, to the anticodon arm and anticodon itself, as much as 75 angstroms away (ref).

Irrespective of just how the tRNA charging code and its entourage of aaRS enzymes evolved, any dRNA loading ribozyme doubtless used approximation of the anticodon for duplicon selection. One way to overcome the distance between the anticodon and duplicon triplets is to form homodimers of dRNA isoacceptors. In "Possible ancestral functions of the genetic and RNA operational precodes and the origin of the genetic system" Juan Martínez-Giménez and Rafael Tabarés-Seisdedos extrapolated their earlier proposal for primitive charging of tRNAs to loading of dRNA (Martínez-Giménez and RafaelTabarés-Seisdedos 2002, 2021). In their scheme, the recognition and loading at a distance problems are solved by transient dimers of identical dRNAs that bring the anticodon of one into approximation with the duplicon of the other (Figure 5-5). This mating of dRNAs multiplies Campbell's original problem of maintaining working sets of 64 dRNA isoacceptors, to forming 64 homodimers without interference from 2016 = (64 x 63) / 2 possible heterodimers.

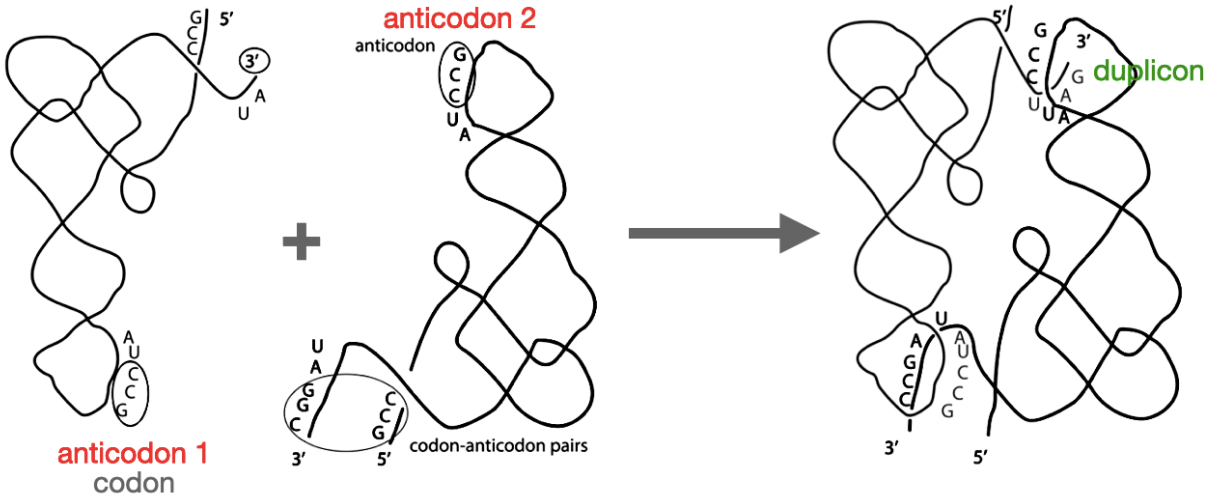


FIGURE 5-5. ANTICODON-DIRECTED dRNA LOADING FROM TRANSIENT HOMODIMERS (MARTÍNEZ-GIMÉNEZ & TABARÉS-SEISDEDOS 2021)

6. RNA copying | reverse complement

In “Speculations on the origin of ribosomal translocation” Robert Weiss and Joshua Cherry, at the University of Utah Department of Human Genetics, mooted a very different scheme of RNA copying inspired by what was then known of the ribosome (Weiss & Cherry 1993). Although their stated object was the origin of translocation and reading frame maintenance, they in fact proposed a general scheme of processive copying fed by a pool of random oligoribonucleotides. Like Campbell’s, their duplisome reads the template in codon triplets from 5’ to 3’ and adds duplicon triplets to the growing 5’ end of the nascent polynucleotide. Whereas Campbell explicitly modeled dRNAs with a central anticodon arm and 3’ duplicon on the tRNA, Weiss and Cherry allowed donor oligoribonucleotides of indefinite sequence and length. Their only requirement was some central triplet that functioned first as the anticodon to read the codon, and then as the duplicon added to the nascent polynucleotide (Figure 6-1). Absent a dRNA loading step based on duplicon-anticodon pairing, codon-anticodon recognition is the only base-pairing step in their elongation cycle, and hence, the Weiss-Cherry copying scheme creates the *reverse complement* of the template.

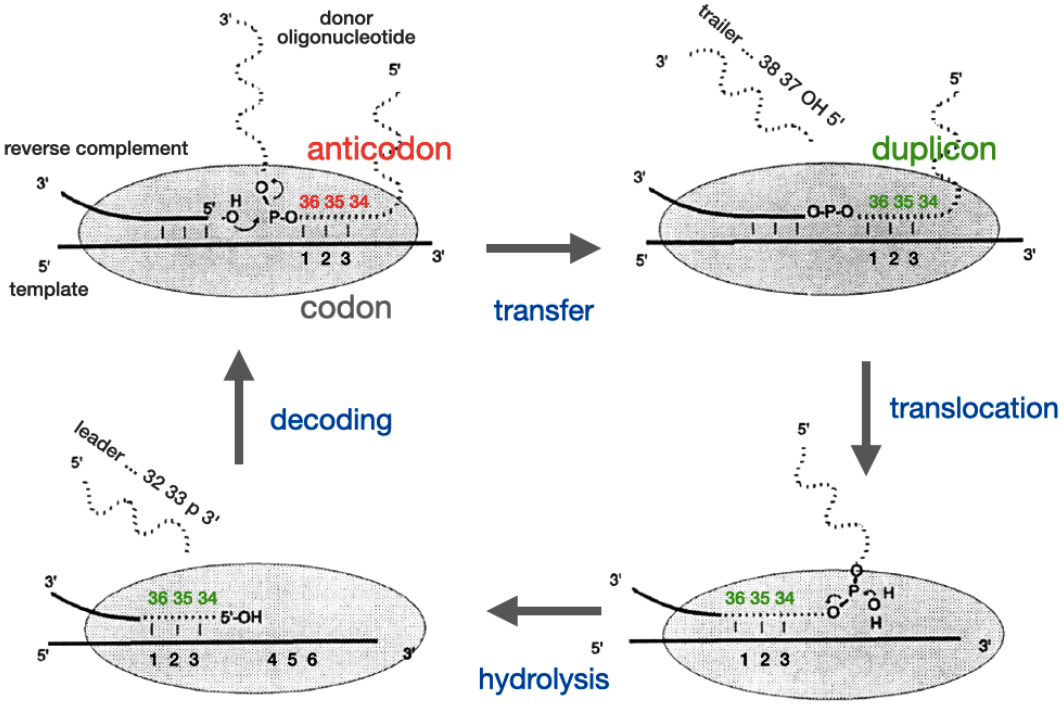


FIGURE 6-1. POLYNUCLEOTIDE ELONGATION CYCLE (WEISS & CHERRY 1993)

Weiss & Cherry (1993) proposed a four-step elongation cycle with an alternation of conformational and covalent steps: DECODING - TRANSFER - TRANSLOCATION - HYDROLYSIS (Figure 6-1). For discussion, we number the anticodon of the donor oligonucleotide $_{34}$NNN$_{36}$ although the leader and trailer sequences are indeterminate (Figure 6-2). In the transfer step of the elongation cycle, O5' of the nascent polynucleotide attacks the donor phosphodiester bond $_{36}$NpN$_{37}$ with trailer O5' N37 leaving. In the hydrolysis step, water attacks $_{33}$NpN$_{34}$ with the duplicon O5' N34 leaving. Thus, at the end of the cycle, the nascent polynucleotide is lengthened by one duplicon HO N34 N35 N36 p, while the trailer HO N37 N38 // and then the leader // N32 N33 p of the donor oligonucleotide are each released.

```
Weiss-Cherry 5' // N32 N33 p  + HO N34 N35 N36 p  +  HO N37 N38 // 3'
Gordon       5' // N32 U33 OH +  p N34 N35 N36 OH + G p N37 N38 // 3'
```

FIGURE 6-2. LEADER + DUPLICON + TRAILER OF DONOR OLIGONUCLEOTIDE

There were three apparent virtues of the Weiss-Cherry scheme of processive RNA copying: First, it uses random oligonucleotides to read the codon and donate the duplicon. Thus, there is no requirement for reloading these donors beyond the feedstock processes that make all four mononucleotides, and the condensation process that lengthens oligonucleotides by random monomer addition or oligomer ligation. Second, their scheme makes the familiar reverse complement of replicative polymerase enzymes (albeit from 3' to 5' by triplet additions). The idea that duplex RNA was a copying intermediate suggested evolutionary continuity, if not actual parsimony, from spontaneous through ribozymatic to enzymatic RNA copying.

Third, the Weiss-Cherry scheme has a proximate source of Gibbs energy to drive processive RNA elongation in its hydrolysis step. The authors suggest that the duplisome tapped this phosphodiester bond hydrolysis to ratchet translocation, unfolding hairpins and similar obstacles in the template. Without this (or another) proximate source, they doubted that the thermodynamics of the initial substrates and final products alone could drive a robust general process of RNA copying. Thus, reliance on mass action from a vast excess of random oligonucleotides would be

vulnerable to imbalances or deficiencies in these donor pools. Similarly, reliance on proprietous secondary and tertiary structure of substrates and products required a delicate balancing on the duplisome of template unfolding and refolding, as well as separation and folding of the nascent product. In their final remarks, the authors speculated on how polypeptide elongation arose from polynucleotide elongation by replacing donor oligonucleotides with aminoacylated transfer oligonucleotides, and nucleotidyl transfer with peptidyl transfer.

Weiss & Cherry (1993) inspired another scheme of RNA copying that reads the template from 5’ to 3’ in codon triplets, and makes the reverse complement from 3’ to 5’ by addition of duplicon née anticodon triplets to the nascent polynucleotide. In “Were RNA replication and translation directly coupled in the RNA (+protein?) world, Karl Gordon, at CSIRO Canberra Division of Entomology, modified their scheme in three ways: First, he modeled the donor oligonucleotides as primitive tRNAs of definite length and folding. Second, he modeled nucleotidyl cleavage and transfer reactions on the group I self-splicing intron, the best understood ribozyme at the time. And third, he proposed that RNA copying was driven by coupling nucleotidyl transfer to a concurrent peptidyl transfer.

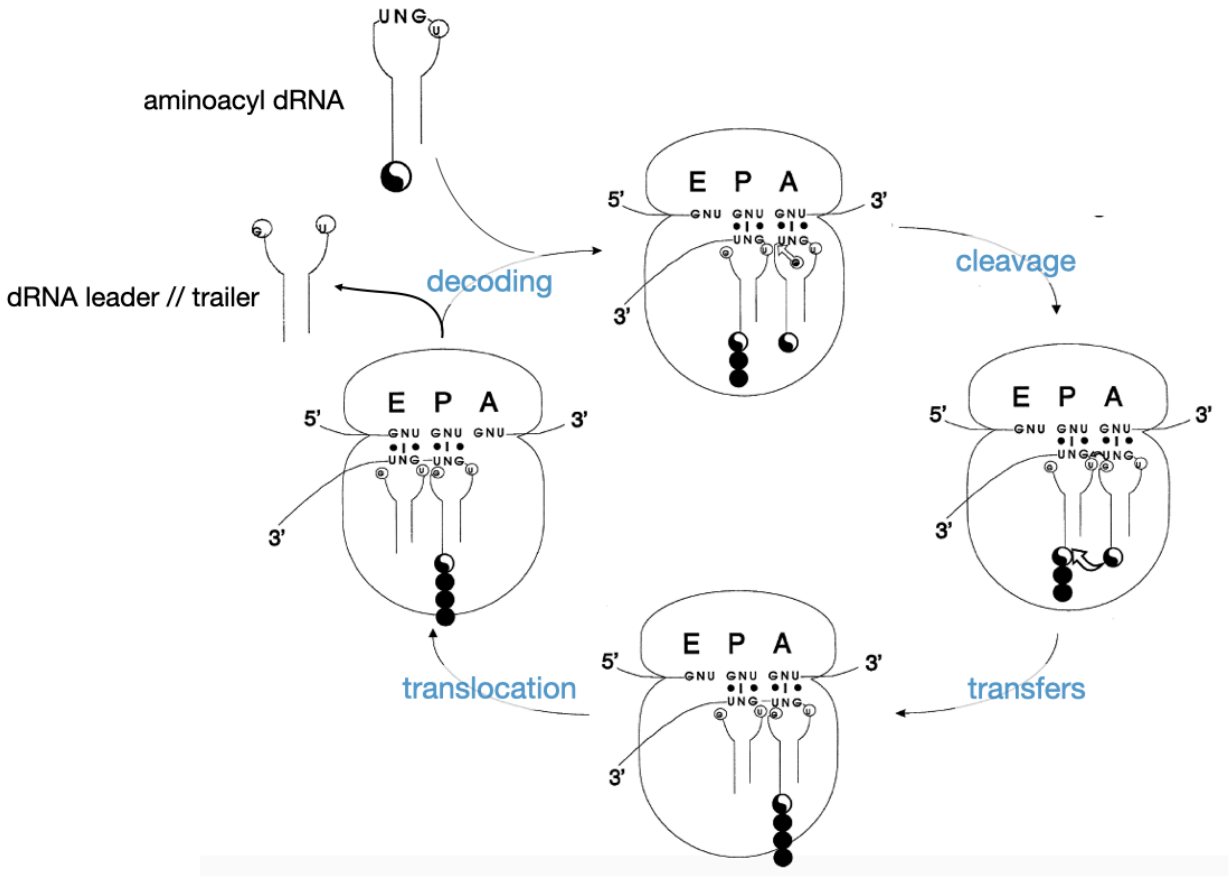


FIGURE 6-3. POLYNUCLEOTIDE ELONGATION CYCLE (GORDON 1995)

Gordon (1995) proposed an elongation cycle with the covalent steps sandwiched between two conformational steps: decoding - cleavage - transfers - translocation (Figures 6-3, 6-4). Decoding, cleavage and nucleotidyl transfer occurred in one catalytic center, while peptidyl transfer occurred in another. He conjectured that the first center, ancestor of the ribosome decoding center, arose from a group I self-splicing intron in the small unit rRNA. For parsimony with the splice site requirements of group I self-splicing intron, he proposed that the primitive decoding center read four codons of the form GNU using four dRNA isoacceptors with anticodon loops of the form $_{38}$NNUNGUN$_{32}$. Gordon modeled cleavage and nucleotidyl transfer on the first and second transesterification reactions of the self-splicing intron. In particular, he proposed that a free guanosine attacks the donor site in the cleavage step (red nucleoside in Figure 6-2).

Whereas the Weiss-Cherry polynucleotide elongation cycle is exergonic, driven by its hydrolysis step, the analogous cleavage step in the Gordon cycle is nearly isoergonic because of the lower chemical activity of guanosine than water. Therefore, to drive processive elongation, Gordon proposed that nucleotidyl transfer was coupled to concurrent peptidyl transfer in the large subunit. In the primitive relation of amino acid charging and dRNA decoding, one ribozyme charged all four dRNA isoacceptors at their 3' ends from a common pool of amino acids. Absent any particular charging code between dRNA isoacceptors and amino acids, this was not true translation, just random polypeptides formed as by-products of processive RNA copying.

Gordon noticed that two features of his RNA copying scheme might solve problems of recognition, initiation, and termination in ribogene replication. First, the 3' and 5' ends of the finished copies comprise the trailer and the leader of the initial and final dRNAs, respectively. Gordon speculated that all templates might carry a complete cloverleaf dRNA at both ends that were effectively regenerated during copying. Second, whereas a paucity of donor oligonucleotides with the cognate anticodon triplet might fortuitously terminate replication in the Weiss-Cherry scheme, Gordon noticed that a dedicated STOP-codon, for which there was no cognate dRNA, could allow codon-directed polynucleotide termination. The possibility of STOP-codons is closely tied to the mechanism of dRNA loading, and how the working set of loaded dRNAs is maintained, including any omissions.

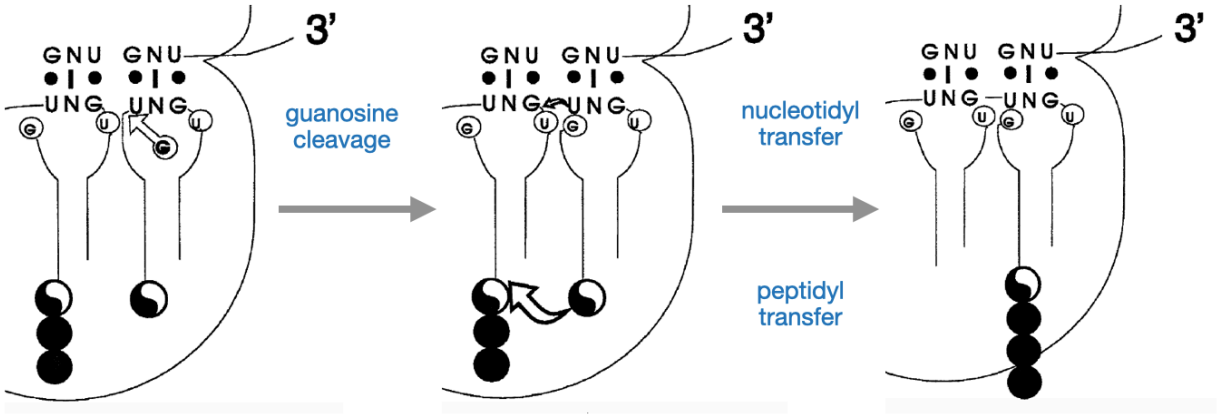


FIGURE 6-4. COVALENT STEPS IN POLYNUCLEOTIDE ELONGATION CYCLE (GORDON 1995)

Gordon constrained both ends of any possible evolutionary path of his duplisome decoding-cleavage-nucleotidyl transfer center, fixing one ancient ribozyme (group I self-splicing intron) as its ancestor and another one (ribosome decoding center) as its descendant. For apparent parsimony with the former, he sacrificed generality of RNA copying and took on the daunting problem of dRNA religation. This restriction of decoding to templates of the form // GNU GNU GNU // or perhaps // RNY RNY RNY // vitiates his scheme as a general process of RNA copying. No matter how fast and accurate, it is unlikely that processive copying that thins out the space of evolvable genes in this way, could supplement, much less replace spontaneous copying. Whereas Weiss & Cherry (1993) relied on feedstock processes for their donor oligonucleotides, Gordon (1995) required an unknown ligase ribozyme as well to reload the set of four dRNA isoacceptors with anticodons $_{36}$GNU$_{34}$. Besides this anticodon loading ligase, an additional charging ribozyme was required for aminoacylation of the dRNA 3' ends. Whereas his polynucleotide ligase ribozyme is purely *deus ex machina*, the aminoacyl dRNA synthetase ribozyme is parsimonious with aminoacyl tRNA synthetase ribozymes required in any model for the breakout of polypeptide translation from the RNA world (section 14).

In "The path from the RNA world", David Penny and colleagues, at the Massey University Institute of Molecular Biosciences, promised that "By focusing on the function of the protoribosome we develop a plausible model for the evolution of a protein-synthesizing ribosome from a high-fidelity RNA polymerase that incorporated triplets of oligonucleotides" (Poole et al 1998, abstract). Their main figure, captioned "An ancient RNA replicase as the precursor of the ribosome," is a mishmash of the Weiss-Cherry and Gordon schemes (Figure 6-5). Like Gordon, donor oligonucleotides are modeled as tRNA-like cloverleafs with an amino acid esterified at its 3' end. They allude to "ribozyme-catalyzed cleavage and ligation functions similar to modern splicosomes," but do not specify the covalent intermediates and by-products of these duplisome reactions. Whereas Gordon proposed to tap peptide bond formation to drive polynucleotide elongation, Penny and colleagues propose to tap the greater free

energy of amino acid release from the dRNA by simple hydrolysis. Outside the duplisome, unspecified ribozymes of dRNA anticodon loading and amino acid charging complete their polynucleotide elongation cycle.

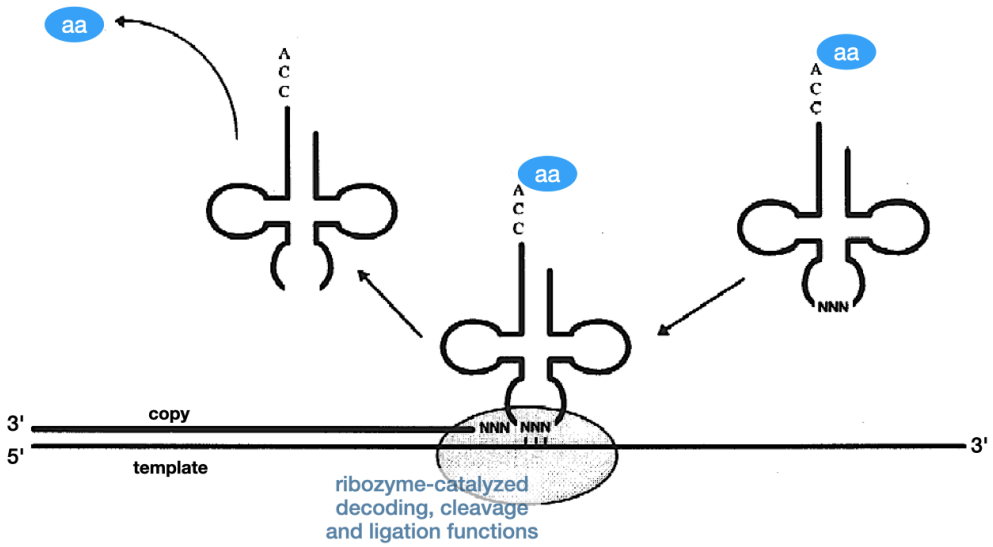


FIGURE 6-5. POLYNUCLEOTIDE ELONGATION CYCLE (POOLE, JEFFARES & PENNY 1998)

7. RNA copying | degenerate forward duplicate

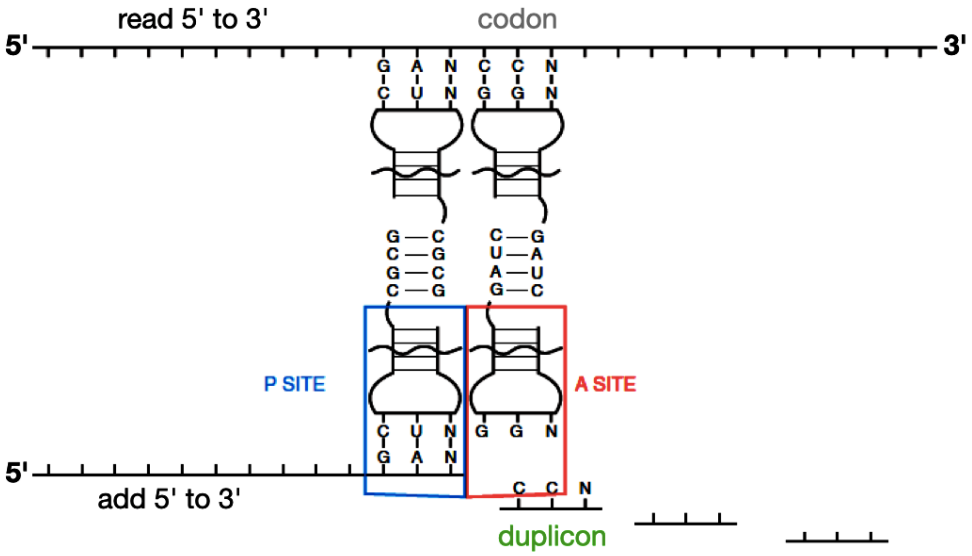


FIGURE 7-1. POLYNUCLEOTIDE ELONGATION CYCLE (NOLLER 2010/2012)

In a remarkable scientific decade, ribosomes in various stages of protein translation were characterized at atomic resolution by X-ray crystallography, and then cryo-electron microscopy (Ban et al 2000; refs; Ramakrishnan 2011). In light of this more detailed picture of decoding, peptidyl transfer, and translocation, Harry Noller at the University of California Santa Cruz Sinsheimer Laboratories pondered functions of the ribosome and tRNAs that predated protein translation. In a talk on the “Evolution of protein synthesis from an RNA world”, he conjectured that the ribosome decoding center evolved for polynucleotide duplication, and was later co-opted for polypeptide translation (Noller 2010/2012). In his copying scheme, the RNA template is read from 5’ to 3’ in codon triplets and its forward duplicate is made from 5’ to 3’ by addition of free duplicon triplets through an unspecified reaction (Figure 7-1). The forward duplicate avoids troublesome copying intermediates, *viz*. there is no useless minus strand to copy again, much less a long duplex to unwind (Table 7-1).

	long duplex	minus strand
forward duplicate	NO	NO
reverse duplicate	NO	YES
reverse complement	YES	YES

TABLE 7-1. RNA COPYING INTERMEDIATES

In lieu of dRNA loading *outside of the duplisome* (Campbell 1991), Noller invoked dimeric decoders inside the duplisome, one for decoding the template, and another for duplicon addition (Figure 7-1). In his scheme, dRNAs formed stable homodimers, so that the anticodon of one subunit reads the codon, while the identical anticodon of the other subunit loads the duplicon to add. Noller noticed that with superwobble of the first anticodon position, just 16 isoacceptors could read all 64 codon triplets, and similarly, there are exactly 16 self-complementary quadruplets, *viz*. palindromes of four nucleotides. Without suggesting any particular relation between anticodons and palindromes, he proposed that dRNA homodimers were held together by such quadruplets at their 3' end (Figure 7-2). While stable marriages of 16 homodimers by self-complementary tails are more plausible, both energetically and kinetically, than transient matings of 64 homodimers (*pace* Martínez-Giménez & Tabarés-Seisdedos 2002, 2021), strong selection was needed to maintain any one-to-one relation between anticodons and palindromes. But worries about the homodimers in his scheme just bury the lede: *the inference that every third nucleobase in the product is unspecified, or degenerate, vitiates its usefulness for RNA copying.*

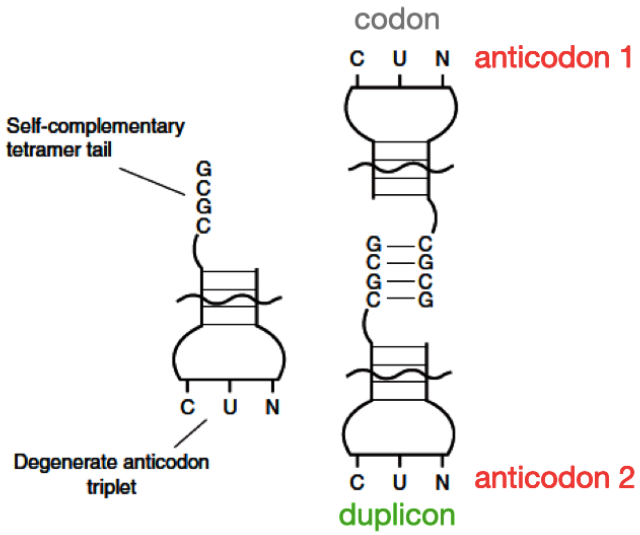


FIGURE 7-2. 16 STABLE dRNA HOMODIMERS (NOLLER 2010/2012)

8. RNA copying | faithful forward duplicate

In Figure 8-1 we graph all of the citations among the original publications on the *duplisome hypothesis*, *viz*. the conjectural origin of ribosome and tRNAs as a means of processive RNA copying. Each artful paper raised interesting talking points about the breakout of polypeptide translation from an RNA world. Campbell was all but lost in the literature, perhaps because his permuted reverse duplicate was unprecedented, if not frankly bizarre. Weiss-Cherry and Gordon attracted some interest, perhaps because their product is the familiar reverse complement of replicative polymerase enzymes (albeit made from 3' to 5' by triplet additions). As more was learned of protein translation, however, their schemes seemed less and less likely as actual ancestors of the ribosome and tRNAs. Similarly, the degenerate forward duplicate in Noller was occasionally noticed, but not seriously critiqued (*pace* Shirokikh 2023). Finally, two monographs gathered together the original references, but could hardly do justice to their eclectic ideas (Bernhardt 2012; Morgens 2013).

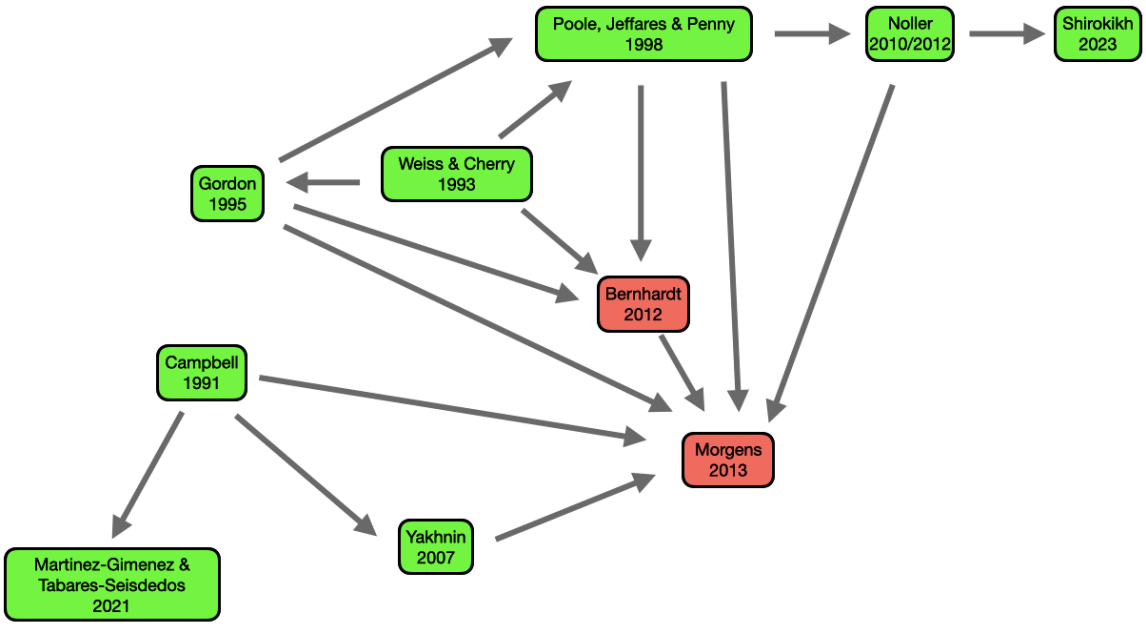
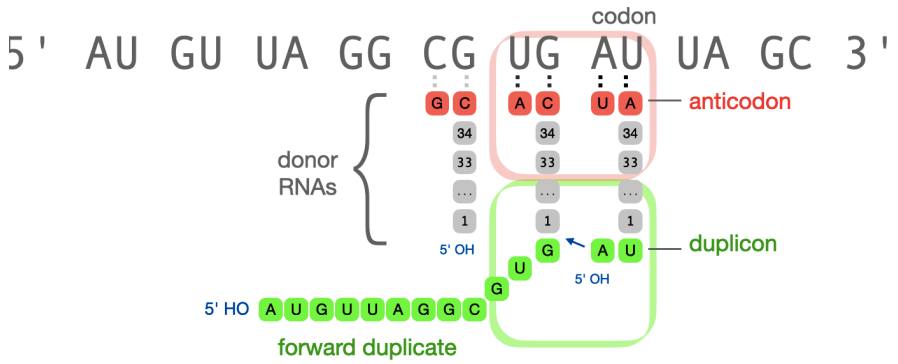


FIGURE 8-1. THE DUPLISOME HYPOTHESIS - HISTORY OF CITATIONS

Coming upon Noller (2010/2012), who cited Poole et al (1998), who in turn cited Gordon (1995) and Weiss & Cherry (1993), and finally exhuming Campbell (1991) with Google Scholar, we admired their bold object to explain the origin of protein translation from a lost process of RNA copying. These eclectic schemes each failed, not for obscure reasons, but patent defects. Whereas Campbell (1991) derived the ribosome decoding and peptidyl transfer centers from duplisome decoding and nucleotidyl transfer centers, respectively, the other schemes tasked the duplisome small subunit with both decoding and nucleotidyl transfer, leaving its large subunit with peptidyl transfer (Gordon 1995), amino acid release (Poole et al 1998), or no task at all (Weiss & Cherry 1993; Noller 2012). Whereas Weiss & Cherry (1993), Gordon (1995), and Poole et al (1998) made the *reverse complement*, Campbell (1991) and Noller (2012) made the *permuted reverse duplicate* or *degenerate forward duplicate*, respectively. Thus, none of these copying schemes made the simplest, and arguably most desirable product, a *faithful forward duplicate*.

Here we present a simple scheme of RNA copying that retains the principal virtues of earlier schemes without their patent defects. *The key difference of our proposal from Campbell (1991) is that the first two nucleotides of loaded dRNAs, not their final three nucleotides, comprise the duplicon.* Two obvious virtues of this revised dRNA structure are that the product of RNA copying is a *faithful forward duplicate* made 5’ to 3’, and just 16 dRNAs are needed to copy any template, not Campbell’s 64 isoacceptors nor Noller’s 16 homodimers (Figure 8-2).



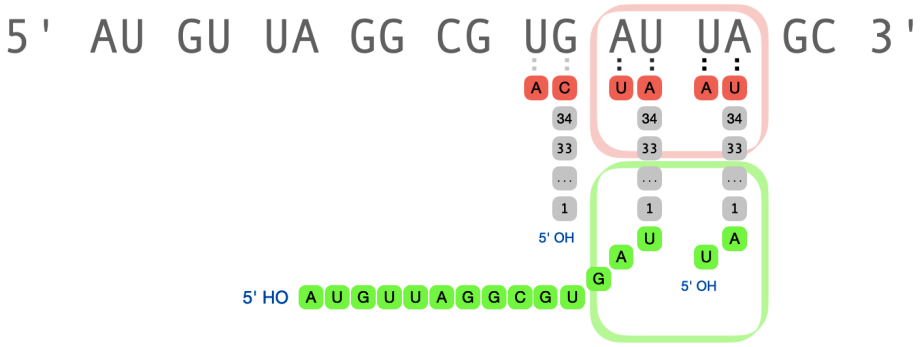


FIGURE 8-2. POLYNUCLEOTIDE ELONGATION CYCLE (PROVISIONAL)

Our provisional elongation cycle depicted in Figure 8-2 is just the start of investigation, not its conclusion. *The seemingly minor change in the size and location of the duplicon constrains all aspects of polynucleotide copying, and raises unexpected questions about the evolution of ribosome life.* In the next three sections we discuss the structure of dRNAs and their ribozymatic loading from random oligonucleotides (section 9), the elongation and termination reactions of the duplisome nucleotidyl transfer center (section 9), the energetics and kinetics of the elongation cycle (section 10), and the mechanism of decoding (section 11). In the final six sections we discuss the origin of the duplisome (section 12), the origin of tRNAs (section 13) and their aminoacylation (section 14), the breakout of polypeptide translation (section 15), the saga of protein life (sections 16 & 17), the DNA world (section 18), and précis the logic of discovery that underlies the evolution of knowledge, and invention of search levels, through polymer life and beyond (section 19).

9. From dRNA loading to nucleotidyl transfer

To turn our provisional scheme of RNA copying from section 8 into a definite elongation cycle, we must specify how the nascent polynucleotide is transferred within the duplisome, and how the freed dRNA is reloaded afterward. After pondering the likely chemistry of polynucleotide transfer and dRNA loading separately, we have converged on one parsimonious solution to both problems. Here we explain how dRNAs are loaded from the pool of random oligomers in a condensation-hydrolysis sequence, or *ratchet*, catalyzed by ribozyme P, progenitor of the catalytic RNA of RNase P. Conversely, polynucleotides are transferred in a hydrolysis-condensation sequence, or *powerstroke*, catalyzed by the duplisome nucleotidyl transfer center, progenitor of the ribosome peptidyl transfer center.

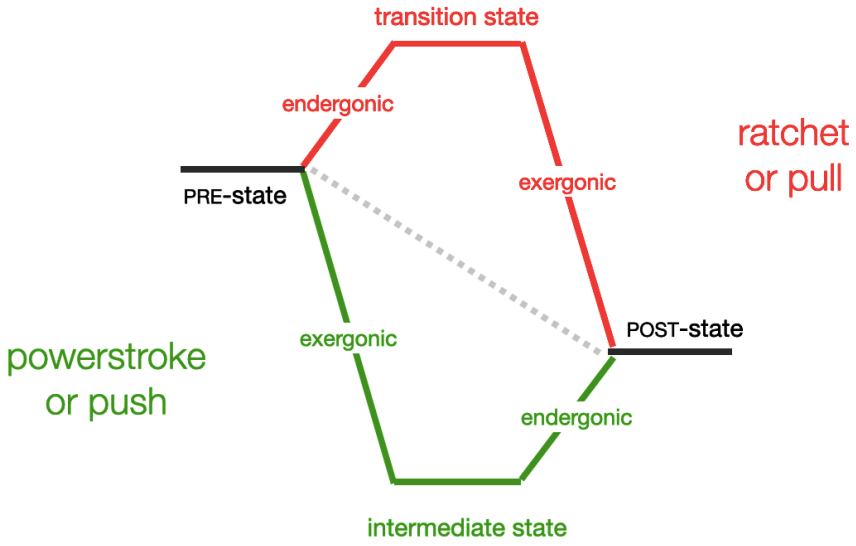


FIGURE 9-1. REACTION PATHS

Before we discuss the biochemistry of dRNA loading and nucleotidyl transfer, we précis basic kinetics and energetics of reaction paths. Figure 9-1 shows an energetically allowable transition from PRE-state to POST-state of lower Gibbs energy. Energetics determine the equilibrium ratios of states, but not how quickly equilibrium is reached. One useful reference is the Maxwell-Boltzmann theory of mixing two ideal gases determined by their kinetic energy of translation, a purely entropic phenomenon

of diffusion that can be accelerated by active mixing. In general there is no diffusion-like or monotone kinetic path from the PRE-state to the POST-state, only more circuitous pathways. The two simplest, non-monotone paths entail (1) an increase in Gibbs energy from the PRE-state, to some unstable intermediate state (aka TRANSITION-state), and then a compensatory fall to the POST-state, or (2) a decrease in Gibbs energy to some stable intermediate state and then a compensatory rise to the POST-state. Pathways that drive an endergonic step with an exergonic step just before or just after are called *powerstroke* and *ratchet* mechanisms, respectively. For purpose of discussion, we refer to these two-step reactions as *pushing* and *pulling*, respectively, and refer to their endergonic and exergonic steps as *uphill* and *downhill*, respectively.

Our key ideas in polynucleotide transfer and dRNA loading are: (1) tight control of water at the active site allows the downhill hydrolysis to respectively push or pull the uphill condensation, and (2) the sites of hydrolysis and condensation are offset by two nucleotides (aka one duplicon) between two dRNAs in transfer, or along one dRNA in loading. This cycle of RNA elongation solves the *water problem* of polymer condensation in bulk solvent, and frames more nuanced questions about the energetics of duplisome life (section 10). Finally, considering basic problems of combination and repair in the RNA world, we trace the duplisome nucleotidyl transfer center back to a primordial ligase ribozyme, while we trace ribozyme P, and perhaps the decoding center itself, back to a complementary repair ribozyme (section 12).

Rather than solving the anticodon-duplicon matching at a distance problem with homodimers (*pace* Noller 2010/2012; Martínez-Giménez & Tabarés-Seisdedos 2021), we propose each dRNA had two conformations called *closed* and *open*. Anticodon and duplicon come together in the closed conformation for dRNA loading, then come apart in the open conformation for decoding and polynucleotide transfer, respectively, at spatially separated sites. To be definite, we adopt as a provisional model of the ancestral dRNA, nucleotides 1-36 of a modern tRNA, with a 15 base-pair stem in the closed conformation, capped by an 8 nucleotide D-loop (14-21), for a total of 38 nucleotides, including the duplicon (Figure 9-2). Using this numbering, the duplicon N2 N1 pairs with the anticodon N35 N36, N1 pairs with N34, and so forth. Our choice of dRNA structure, motivated by the need to pair duplicon and anticodon for loading,

constrains the entire scheme of RNA copying (sections 8-11). From this, we reconstruct a rich history of polymer life, from the origins of RNA duplication, to the breakout of polypeptide translation, and onward, to the retirement of the duplisome and beyond (sections 12-18). Explaining the structure and functions of dRNAs as we go, we draw attention in several places to minor variations on this particular model.

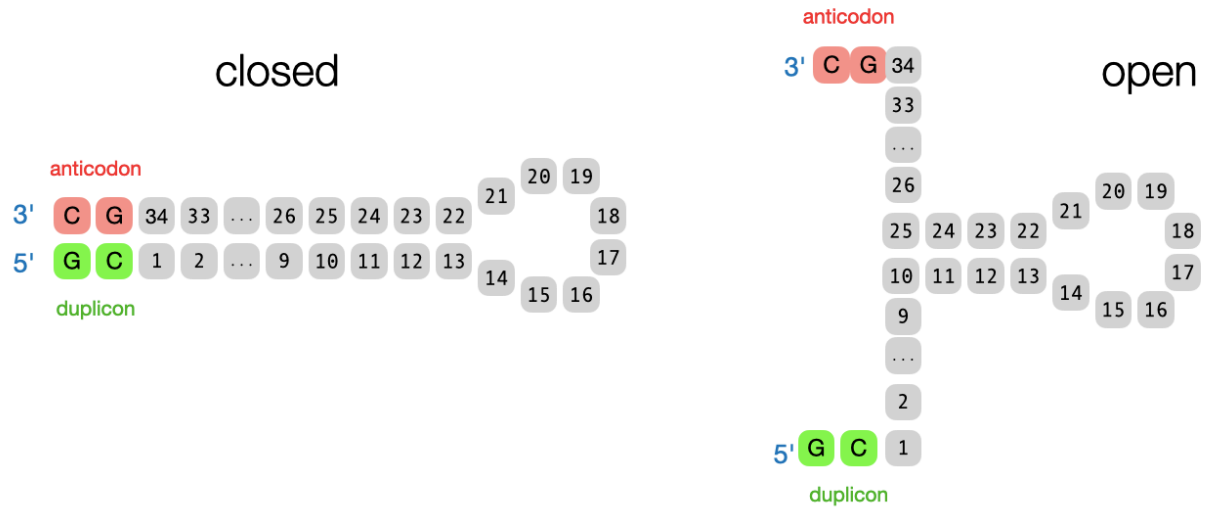


FIGURE 9-2. dRNA CONFORMATIONS

Found in all kingdoms of cellular life, the RNA component of RNase P removes the 5’ leader of pre-tRNAs and eclectic other substrates (Guerrier-Takada et al 1983; Gößringer et al 2021; Phan et al 2021). In bacteria, RNase P cleaves pre-tRNAs, pre-tmRNA, pre-SRP RNA, pre-rRNA and certain mRNAs. In eukaryotes, gene duplication of the catalytic RNA gave rise to two very similar complexes called RNase P and RNase MRP that share most of their accessory proteins (Welting et al 2006; Coughlin et al 2008). The former cleaves pre-tRNAs and certain pre-snoRNAs, while the latter cleaves pre-rRNA and mitochondrial primers, as well as certain lncRNAs and mRNAs.

In both natural and artificial substrates, an unpaired 5’ leader is cleaved from a paired stem, e.g., a simple helix and loop for pre-SRP RNA, or coaxial helices and loop for pre-tRNAs, leaving hydroxyl and phosphate at the 3’ and 5’ ends, respectively (McClain et al 1987; Kirsebom & Trobro 2009). Studies on pre-tRNAs and their mimics have identified key features in substrate recognition, positioning and hydrolysis. Two

features at some remove from the scissile bond assist binding of pre-tRNAs to bacterial RNase P. First, in what has been dubbed a *molecular ruler*, the pre-tRNA elbow binds the interdigitating T-loop motif of the specificity domain, placing the topological junction of leader/stem at about the right position in the catalytic domain for trimming (Chan et al 2013; Lehmann et al 2013; Zhang & Ferré-D'Amaré 2016a,b). Second, sequence-independent interactions of the distal leader with the accessory protein increase the affinity of RNase P for pre-tRNAs relative to mature tRNAs, allowing product release, and preventing product inhibition.

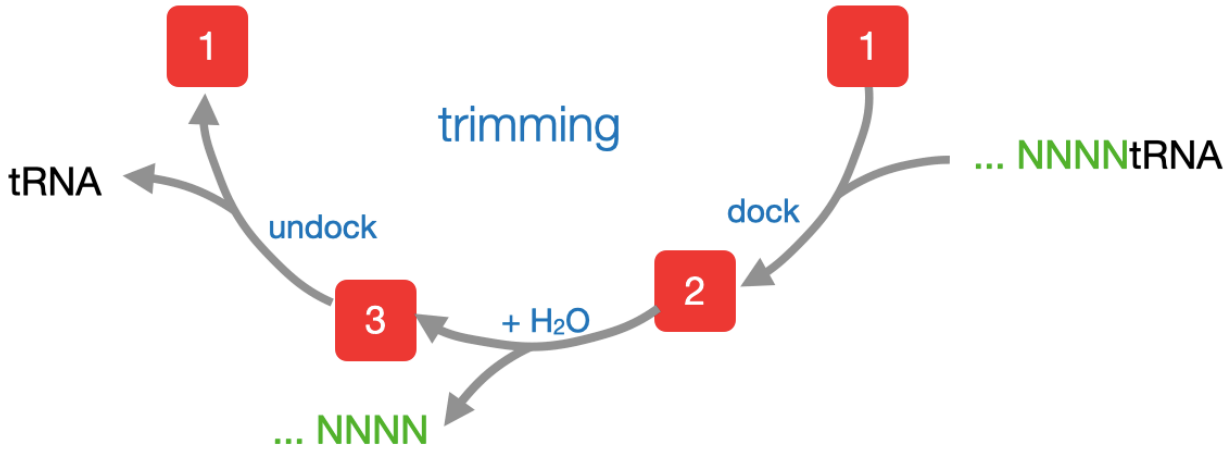


FIGURE 9-3. pre-tRNA TRIMMING BY RNase P

Once interactions with the pre-tRNA elbow and distal leader form the substrate encounter complex with RNase P, additional interactions dock the scissile bond at the active site in the catalytic conformation (Lan et al. 2018; Zhu et al 2022). The preferred cleavage site is not specified by the nucleotide sequence, but by the topological junction between the single-strand leader and double-strand stem. The crucial steps are unwinding any fortuitous pairing between leader and trailer preceding the mature stem, and rotating nucleotide U51 of helix P4.[8] Unwinding entails stacking A248 of junction J5/15 on base-pair N1:N72, thereby unstacking N73 from the acceptor stem and exposing the Hoogsteen edge of A248 to pair with N1. In bacteria, where the

[8] Numbers of the RNase P RNA sequence are from *Escherichia coli*.

terminal 3’ C74 C75 A76 of the mature tRNA are encoded in the primary transcript, G292 G293 U294 in the internal loop of P15[9] pair with these nucleotides, holding the proximal trailer flayed away from the 5’ leader of the pre-tRNA (Kirsebom & Svard 1994). With the substrate unwound, U51 rotates into the active site, coordinating one Mg^{++} that generates the hydroxide nucleophile, while a second Mg^{++} assists the leaving group departure. After S_N2 hydrolysis, the leader quickly dissociates, while a slower return from catalytic to encounter conformation releases the mature tRNA (Tallsjo & Kirsebom 1993). Under physiological conditions, the holoenzyme completes about … turnovers per second.

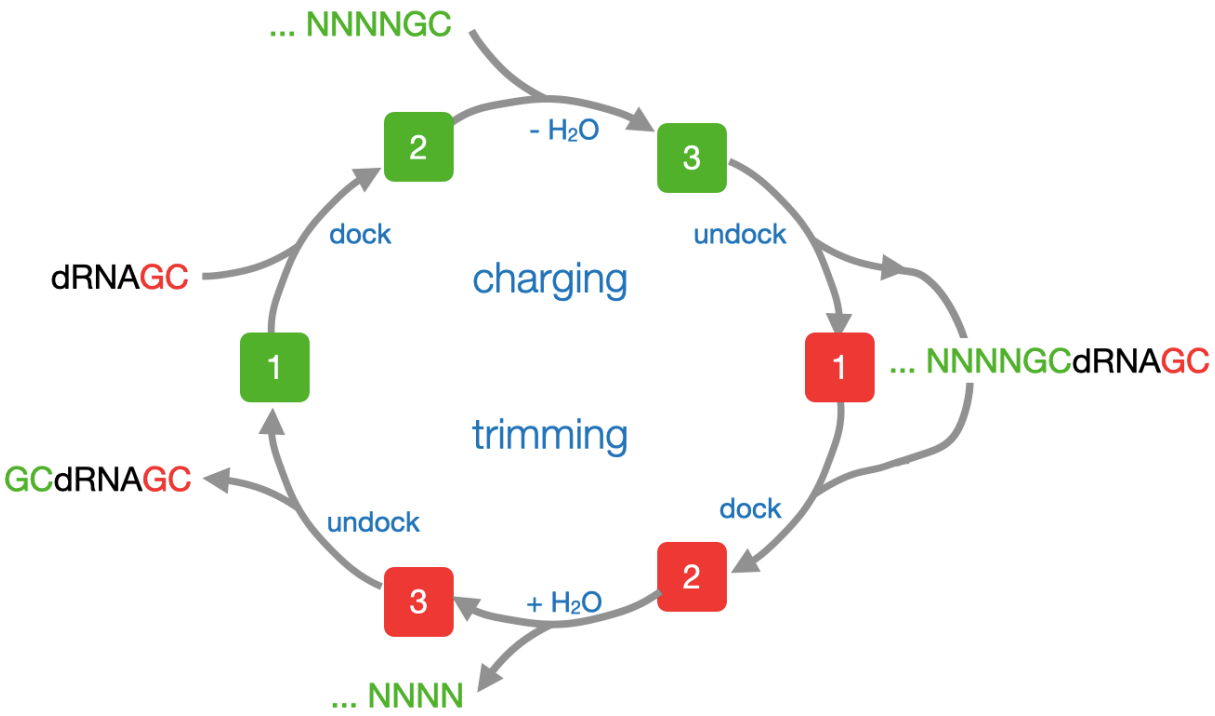


FIGURE 9-4. dRNA LOADING BY RIBOZYME P

We propose that dRNAs were reloaded from the pool of random oligomers in a coupled condensation-hydrolysis sequence catalyzed by an ancestor of P RNA called

[9] Internal loop L15 is absent in the P RNA of eukarya and many archaea?, consistent with the post-transcriptional addition of the terminal 3’ CCA after pre-tRNA processing by RNase P.

ribozyme P for sake of discussion. We propose this ribozyme loaded the free dRNA from a cognate oligomer via the sequence of docking-condensation-redocking-hydrolysis-undocking. For convenience, we refer to the covalent steps of condensation and hydrolysis as *charging* and *trimming*, respectively, and the complete sequence as *loading* (Figures 9-3 & 9-4). Complexed with ribozyme P, the dRNA anticodon samples the pool of oligomers for a match. At the end of the loading sequence, the input oligomer is shortened by two nucleotides at its 3' end, while the free dRNA is now loaded with these same two nucleotides at its 5' end. There is no net change in the number of RNA molecules or phosphoester bonds, but free energy of base-pairing between duplicon and anticodon (- 4-6 kcal /mol) favors charging cognate oligomers over non-cognate ones, and helps drive the overall reaction toward loading.

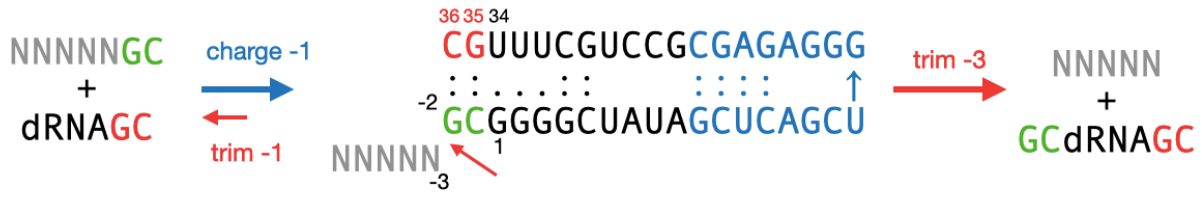


FIGURE 9-5. dRNA LOADING FROM COGNATE OLIGOMER

Our model of dRNA loading conjectures that (1) the tightly coupled, condensation-hydrolysis sequence was a quasi-reversible process, and that (2) the active site shifted between the two catalytic steps. The irreversible maturation of pre-tRNAs by RNase P, *viz*. docking-hydrolysis-undocking, is a vestige of dRNA trimming; there is no similar vestige of dRNA charging. If our energetic ratchet is novel for ribozymes, pulling the first step (condensation) uphill via a second step (hydrolysis) downhill, self-splicing introns are examples *par excellence* of staggering reaction sites by redocking the intermediate (refs).

In dRNA loading, the scissile bond rewinds from the site of condensation N1↑N1 to the site of hydrolysis N3↓N2. Thus, A248 stacks on N1:N34 during charging, and re-stacks on N2:N36 during trimming. Other interactions used for pre-tRNA positioning on

RNase P were different or absent in dRNA loading. The length of the closed dRNA stem, say 13 or 15 basepairs for charging and trimming, respectively, was comparable to the 12 or 13 basepairs in the coaxial helices of acceptor- and T-arms in pre-tRNA trimming today. But only a primitive D-loop, not the modern D/T elbow was available to position the scissile bond by a ruler mechanism. Whereas pre-tRNAs have a sizable 3' trailer, beginning N73 C74 C75 A76 in bacteria, dRNAs had only the unpaired anticodon N35 N36 during charging, and no trailer at all during trimming. Finally, ribozyme P had no coded accessory protein in the RNA world.

There is indirect evidence from comparative genomics, and direct evidence from mutant pre-tRNAs, that the preferred site of RNase P cleavage can be wound in either direction by matches that lengthen the paired stem, or mis-matches that shorten it. In some cases both a decrease in cleavage at the canonical site and an increase in off-site cleavage have been observed. A natural -1 shift in the cleavage site due to pairing with the unusual discriminator C73 creates the 8 base-pair acceptor stem of bacterial $tRNA^{His}$, compared to the 7 base-pair stem of most tRNAs (Orellana et al 1986). In rat nuclear pre-$tRNA^{Lys}$, base substitutions that leave N1:N72 paired are well-tolerated, while substitutions that unpair these nucleotides reduce cleavage at the canonical site without an apparent increase in +1 cleavage (Paisley & Van Tuyle 1994). In yeast nuclear pre-tRNAs a mismatch at N1:N73 appears important to prevent leader-trailer pairing from winding the preferred cleavage site (Lee et al 1997).

Kinetic and structural studies on cleavage site selection in pre-tRNAs are hard to interpret because remote features, not just the topology and sequence of the cleavage site, affect substrate docking and catalysis. Substrates lacking the specific elbow and trailer interactions of pre-tRNAs may better illustrate topological docking of the scissile bond. In *Escherichia coli* pre-SRP RNA, the 5'-leader is cleaved at the same site U24↓G25 whether from the intermediate hairpin H1 formed after transcription of the first 36 nucleotides, or from the mature hairpin formed after transcription of all 138 = 24 + 114 nucleotides (Fukuda et al 2020). In living cells, the leader is likely removed co-transcriptionally once the H1 hairpin forms.

Minimal substrates retaining few of the features of pre-tRNAs …. Thus, bacterial P RNA can efficiently cleave the (length?) 5' leader from a model substrate with a 12

basepair stem and 7 nucleotide loop, and 5 nucleotide 3' trailer NCCAN (McClain et al 1987). A partial shift to -1 cleavage was seen in an artificial hairpin-loop substrate with ?leader? 4 bp stem, GAAA tetraloop, and 3' trailer CCAC (Brännvall et al 2007).

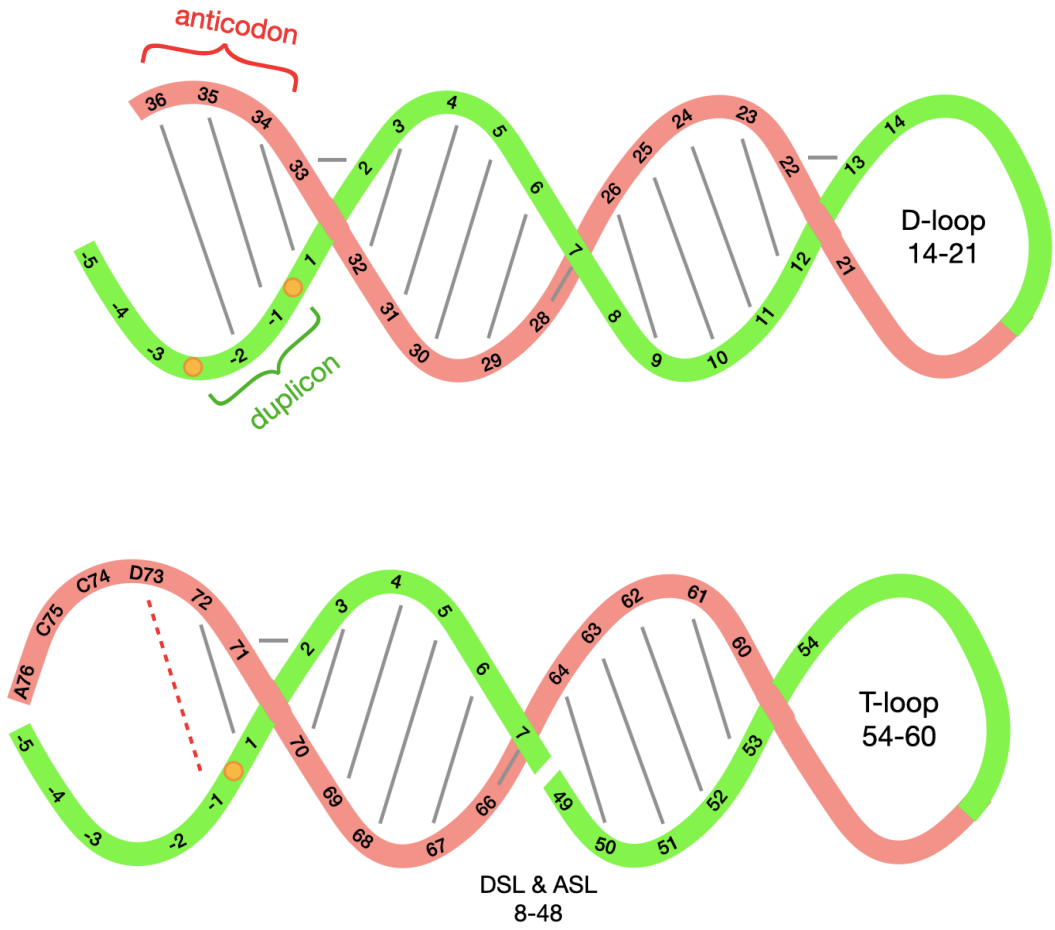


FIGURE 9-6. dRNA LOADING (TOP) VERSUS pre-tRNA TRIMMING (BOTTOM)

In Figure 9-6 we compare dRNA loading by ribozyme P and pre-tRNA maturation by RNase P. Scissile phosphates are shown as brown circles. The closed dRNA is modeled as an RNA duplex topped by D-loop (Dickerson et al 1982; Saenger 1984). It has been charged at N1↑N1 with a cognate oligomer, and awaits trimming at N3↓N2. dRNAs end with N36, the third position of the anticodon. In the pre-tRNA, the acceptor and T arms coaxial helices are topped by the T-loop, where a gap marks the D and

anticodon arms. The pre-tRNA awaits trimming at N1↓N1. The critical mismatch or bulge between N1 and N73 is shown by a dashed red line.

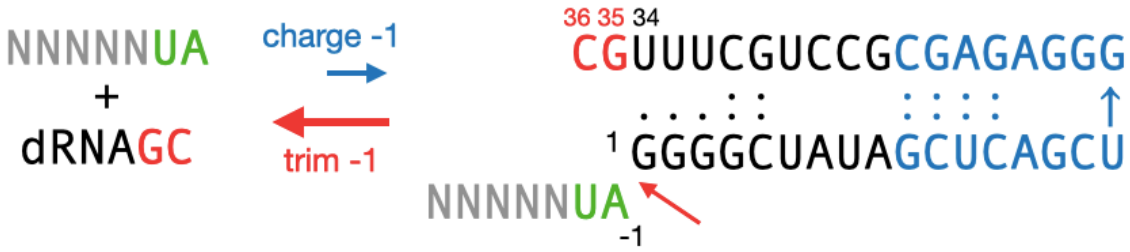


FIGURE 9-7. dRNA CHARGING WITH NON-COGNATE OLIGOMER

Until the excess leader is trimmed away, an overloaded dRNA likely cannot enter the duplisome decoding center, much less accommodate the nucleotidyl transfer center. But misloaded dRNAs whose duplicon mismatches the anticodon at one, or rarely both positions, were a likely source of substitution errors, and might even cause misreading of codons (section 11). For non-cognate oligomers, the ratios of forward and reverse reactions are shifted in both charging and trimming (Figure 9-7). Thus, *mispaired* dRNAs are more likely to dissociate than mischarge, and *mischarged* dRNAs are more likely to trim at N1↓N1 without winding to N3↓N2. Interestingly, *misloaded* dRNAs might be *retrimmed* via a second hydrolytic cleavage at N1↓N1 (Figure 9-8). As the loss of one phosphoester bond separates the trimming at N3↓2 from the retrimming at N1↓N1, this rejection pathway for *misloaded* dRNAs is a simple form of kinetic proof-reading (Hopfield 1974; Ninio 1975).

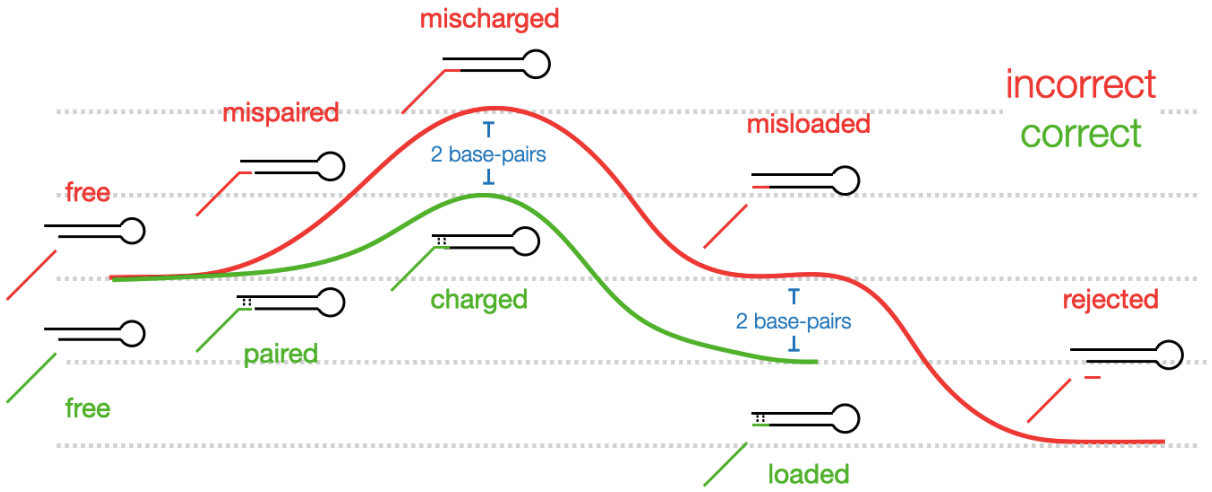


FIGURE 9-8. REJECTION OF MISLOADED dRNA

Our scheme of dRNA loading catalyzed by ribozyme P requires a phosphate at the 5’ end of free dRNAs, and leaves one at the 5’ end of loaded dRNAs. Obvious mechanisms of polynucleotide transfer use O5’ of the loaded dRNA for nucleophilic attack on the phosphodiester bond of the polynucleotidyl-dRNA directly, or else, the polynucleotidyl 2’,3’ cyclic phosphate intermediate formed by strand scission as in self-cleaving ribozymes. But to use the O5’ oxyanion as nucleophile, we must remove the 5’ phosphate from the loaded dRNA *before it enters the A-site*, and restore the 5’ phosphate to the free dRNA *after it leaves the E-site*. To wit, loading and transfer reactions could be coupled together via a *dRNA 5’-phosphotransferase* that transfers this phosphate from loaded to free dRNAs (Figure 9-9). Or, rather than this isoergonic shuttle, a *duplicon-dRNA 5’ phosphatase* and *free-dRNA 5’ kinase* could work in tandem to drive the elongation cycle from some high-energy phosphate donor.

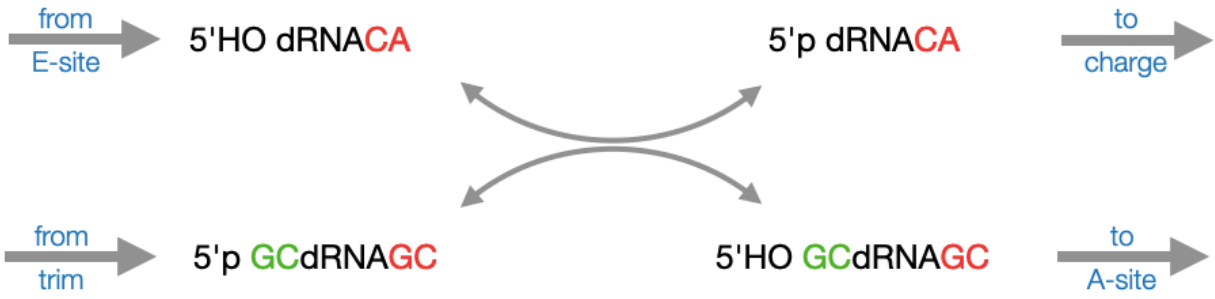


FIGURE 9-9. dRNA 5' PHOSPHOTRANSFERASE

Figure 9-10 depicts the substrates and products of nucleotidyl transfer by a concerted transesterification where O5' oxyanions of the duplicon-dRNA and the freed dRNA are the inline nucleophile and leaving group, respectively. The reaction passes through a trigonal bipyramidal (sp^3d) transition state, or a more stable phosphorane intermediate. The electron movements are complemented by a series of proton movements, or shuttle, mediated by the substrates, transfer center, and perhaps water. After transfer, the P-site dRNA begins 5' HO N1, while the A-site dRNA carries the nascent polynucleotide chain elongated by one duplicon. Importantly, whereas peptidyl transfer in the ribosome is exergonic because of the greater stability of amide than acylester, nucleotidyl transfer is nearly isoergonic, with no net change in the number or kinds of covalent bonds.

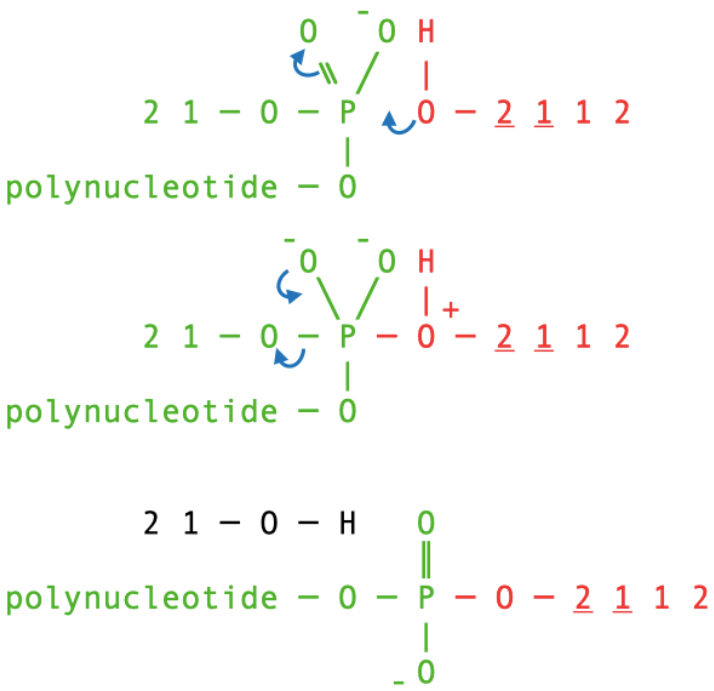


FIGURE 9-10. NUCLEOTIDYL TRANSFER BY CONCERTED TRANSESTERIFICATION

Reversible strand scission in self-cleaving ribozymes suggest a two-step pathway for nucleotidyl transfer mediated through a polynucleotidyl 2',3' cyclic phosphate intermediate (Figure 9-11). Converting the abortive cleavage-religation cycle of self-

cleaving ribozymes into a productive sequence, polynucleotidyl-dRNA in the P-site undergoes strand scission by the vicinal O2’ to free the 5’ HO dRNA and retain the nascent polynucleotide with 2’,3’ cyclic phosphate. In the second step, O5’ of the duplicon-dRNA in the A-site attacks the strained cyclic phosphodiester bond, recreating a polynucleotidyl-dRNA, now elongated by the new duplicon and its cognate dRNA.

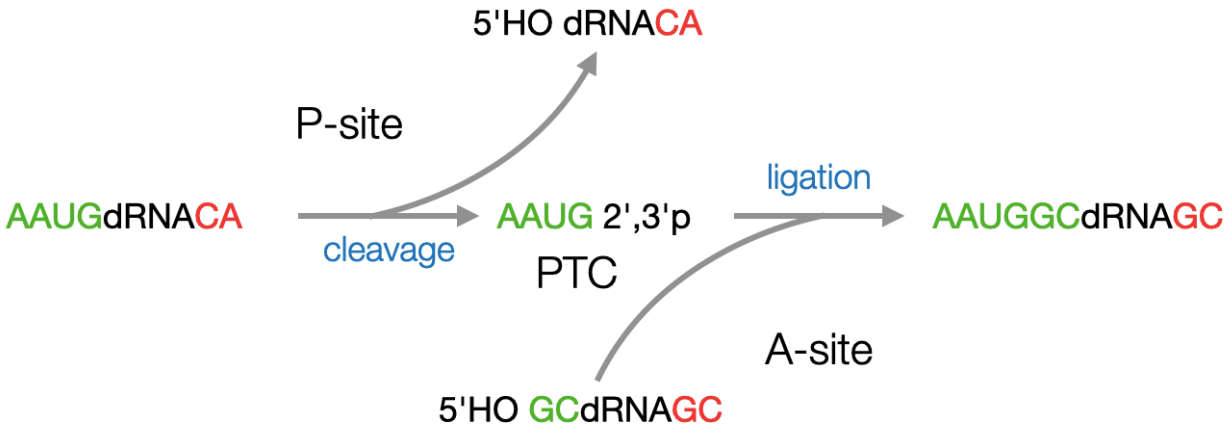


FIGURE 9-11. NUCLEOTIDYL TRANSFER BY SEQUENTIAL TRANSESTERIFICATION

Both elongation cycles for RNA duplication outlined above use ribozyme P for dRNA loading and the 5’ oxyanion of duplicon-dRNA as nucleophile for nucleotidyl transfer. A serious, perhaps fatal drawback of these schemes to couple dRNA loading and polynucleotide transfer is that they require one (*phosphotransferase*) or even two (*phosphatase*, *kinase*) extinct ribozymes for manipulating dRNA 5’ ends, *in addition to ribozyme P known through extant RNase P*. What if instead, we simply retain the 5’ phosphate of the newly loaded dRNA as it enters the duplisome A-site? This alternative nucleotidyl transfer reaction, or *transphosphorylation*, is formulated in Figure 9-12. Like transesterification, this reaction is isoergonic, but unlike the former, it is sterically impossible as a concerted mechanism. Below we propose a sequential mechanism of transphosphorylation, again without precedent in natural or artificial ribozymes.

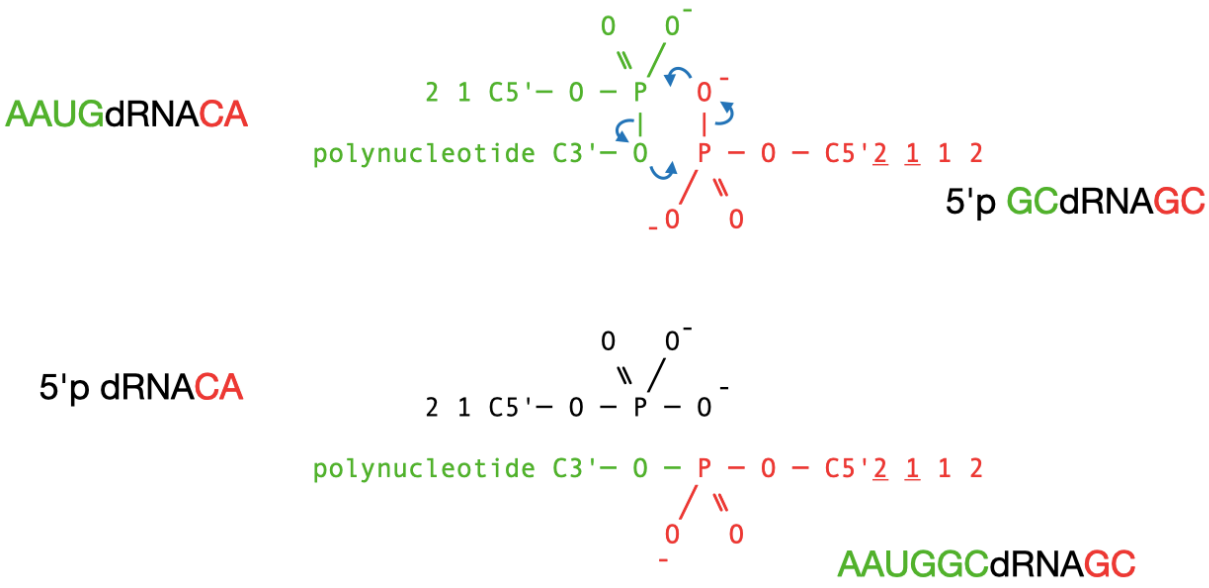


FIGURE 9-12. CONCERTED TRANSPHOSPHORYLATION

Inspired by the conjectural condensation-hydrolysis sequence in dRNA loading catalyzed by ribozyme P, we now consider a hitherto obscure possibility, *viz*. the 3’ OH of the nascent polynucleotide attacks the 5’ phosphate of the duplicon-dRNA as the *second step* of a hydrolysis-condensation sequence, confined from bulk solvent within the duplisome nucleotidyl transfer center. So long as free water is kept from the reaction center, there seems no obstacle to this tightly coupled hydrolysis-condensation sequence with a complementary proton shuttle (Figure 9-13).

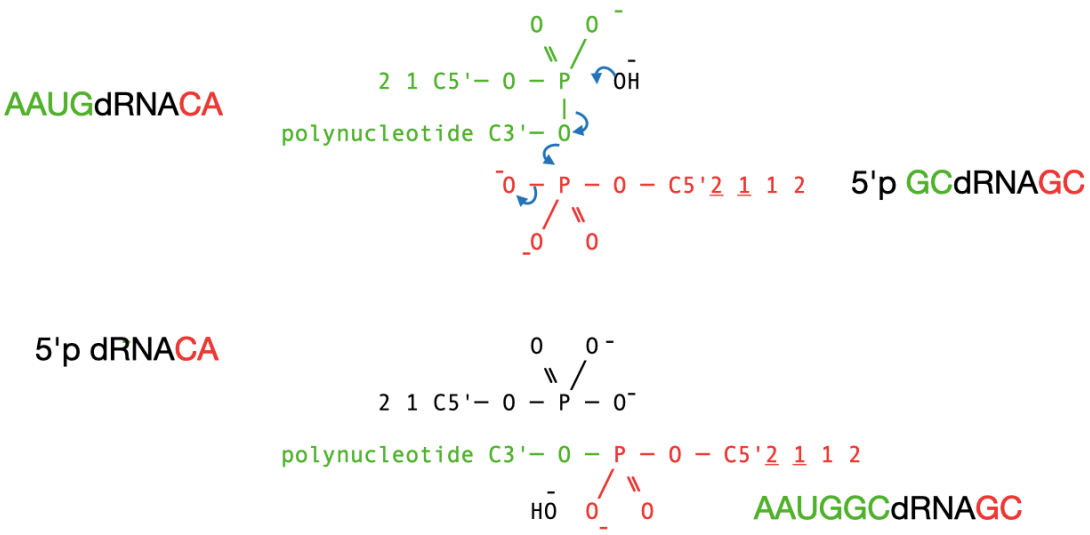


FIGURE 9-13. POLYNUCLEOTIDE TRANSFER VIA SEQUENTIAL HYDROLYSIS & CONDENSATION

In Figure 9-14, we depict the catalytic events in (left) dRNA loading by ribozyme P and (right) polynucleotide transfer by the duplisome center. Small blue arrows connect each nucleophile (tail) to its electrophile (head). P-site and A-site reactants in the duplisome are highlighted in orange and blue, respectively. In dRNA loading condensation precedes hydrolysis, while in polynucleotide transfer hydrolysis precedes condensation. If hydrolysis is not followed by condensation, say when the A-site is empty, elongation terminates in release of the nascent polynucleotide through the P-site exit tunnel.

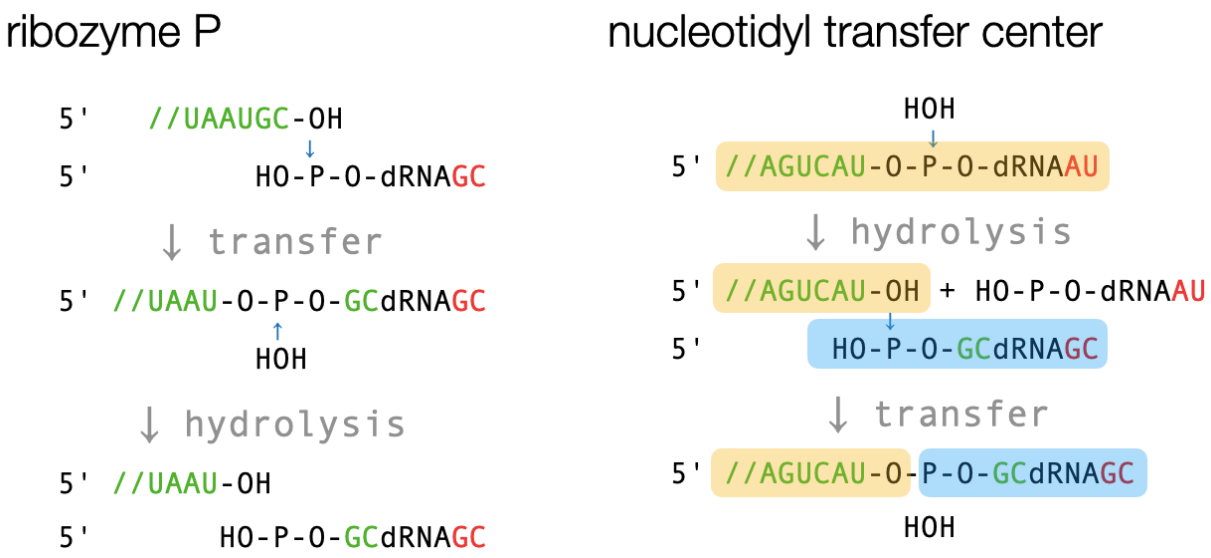


FIGURE 9-14. CHEMISTRY OF POLYNUCLEOTIDE ELONGATION CYCLE

The peptidyl transfer center at the heart of the ribosome large subunit is a roughly symmetrical pocket formed from P- and A-regions along the central loop of domain V, about 180 nucleotides overall, that possibly arose by a tandem duplication.[10] Thought to be the oldest part of the large subunit, this center likely originated as a free-standing ribozyme in the RNA world (Bokov & Steinberg 2009; Petrov et al 2014). One obvious suggestion is that this original ribozyme dubbed the *protoribosome* had a role in random polypeptide synthesis, catalyzing essentially the same acylester-to-amide transfer on RNA carriers as the ribosome (Agmon et al 2009; Tamura 2015). Whereas ancestors of the protoribosome likely acted on other RNA world substrates, its role as a peptidyl transferase is proposed to have emerged *before the core transfer center began accretion of RNA extensions and functions along its way to becoming the large subunit rRNA*.

Two observations have been taken as evidence for a peptidyl transferase ribozyme as progenitor of the ribosome large subunit. First, artificial ribozymes can catalyze peptide bond formation by positioning model substrates relative to one another (?

[10] The peptidyl transfer center comprises P-region H74 H89 and A-region H90 H93 of the large subunit rRNA (Krupkin et al 2011). The central loop of this center is closed by helix H73, while the region between H73 and H74 forms the start of the nascent polypeptide exit tunnel. The folded peptidyl transfer center has an inner core of some 120 nucleotides, and an outer shell of 60 nucleotides that includes the P- and A-loops.

Weber & Orgel 1980; Lohse & Szostak 1996; Zhang & Cech 1997; Tamura & Schimmel 2003). In one model reaction, the α-amine of phenylalanine, tethered to the ribozyme 5' end can attack the carbonyl carbon of aminoacyl-esters on the 3' OH of adenosine monophosphate, albeit with 20,000-fold slower turnover than the ribosome elongation cycle (Zhang & Cech 1997). Compared to water or alcohols, the non-protonated primary amine is a strong nucleophile with a pK_a about 8. The tetrahedral transition state has a relatively low activation energy, and the products are energetically favored. [aa-AMP models]

Second, like peptidyl transferase ribozymes selected *ab initio*, core fragments of the ribosome peptidyl transfer center can catalyze peptide bond formation between aminoacyl-minihelices, or even simpler tRNA mimics (Bose et al 2022; Kawabata et al 2022). Heterodimers of the P- and A-region cores, or homodimers of the P-region core alone, are active. Having striped away later additions to the ribosome and its tRNAs, these core fragments of the ribosome transfer center and tRNA acceptor arms, are proposed to recapitulate an important milestone in the evolution of polymer life. "Although doubts and caveats remain, Yonath's and Tamura's work seems to recapitulate a milestone on the road from primordial organic molecules to the ribosome used by the last common ancestor or all living things" (Dance 2023).

We propose that the progenitor of the ribosome peptidyl transfer center was a duplisome nucleotidyl transfer center that catalyzed polynucleotide elongation by sequential hydrolysis-condensation reactions. *If so, extrapolation from the ribosome peptidyl transfer center to a ribozyme (pace protoribosome) that made random polypeptides is mistaken.* In section 12 we trace the nucleotidyl transfer center back to a primordial ligase ribozyme, a workhorse of non-templated combination and repair in early RNA life. In sections 13 to 15 we trace the duplisome nucleotidyl transfer center and its dRNAs forward to the ribosome peptidyl transfer centers and its tRNAs through two new codon-directed adaptor-mediated reactions, polynucleotide termination and polypeptide elongation (Table 9-1). Excepting peptide bond formation, all of the reactions in Table 9-1 consume the equivalent of one water molecule. Desolvation, or excluding bulk water from the active site, averts side reactions that compete with polymer elongation. In polynucleotide ligation … (section 12). In nucleotidyl transfer,

one activated water is needed for hydrolysis, and in the absence of condensation, results in polynucleotide termination and release (section 13). In polynucleotide termination ….(section 14).

REACTION		P-SITE SUBSTRATE	A-SITE SUBSTRATE	H_20	BOND LOST	BOND MADE	ΔG kcal / mol
polynucleotide	ligation	5' polynucleotide	3' polynucleotide	Y	phospho ester	phospho ester	~ 0
"	elongation	polynucleotidyl dRNA	duplicon dRNA	Y	phospho ester	phospho ester	~ 0
"	release	polynucleotidyl dRNA	terminator tRNA	Y	phospho ester	none	- ?
polypeptide	elongation	polypeptidyl tRNA	aminoacyl tRNA	N	acyl ester	acyl amide	-6
"	release	polypeptidyl tRNA	terminator tRNA	Y	acyl ester	none	- ?

TABLE 9-1. EVOLUTION OF POLYMER TRANSFER CENTER

Our proposed origin and evolution of the polymer transfer center helps to explain the entropic mechanism, substrate promiscuity, and evolutionary plasticity of the ribosome peptidyl transfer center. In a simple analysis of catalysis, the uncatalyzed reaction proceeds from substrate (S) through transition state (T) to product (P) (Figure 9-15). *For book-keeping we include the catalyst (E) in this reaction pathway as an inert component, or non-reactant.* In purely enthalpic catalysis, substate and product interact weakly with the catalyst, associating and disassociating from the catalyst with negligible change in Gibbs energy. The important interaction is stabilizing the transition state, and thereby, lowering the activation energy of reaction. In purely entropic catalysis, the substate and product interact strongly with the catalyst, docking and undocking with significant change in Gibbs energy. The transition itself entails negligible activation energy so the temperature dependence of reaction reflects only entropy costs. As a rule, entropic catalysis is important for bimolecular reactions where the catalyst plays host to two substates, while enthalpic catalysis is important for unimolecular reactants with a single substrate, not counting water. The thermodynamic

distinction between enthalpic and entropic catalysis is seen empirically in an Arrhenius plot of log k versus T^{-1} for the temperature-dependence of the reaction (ref).

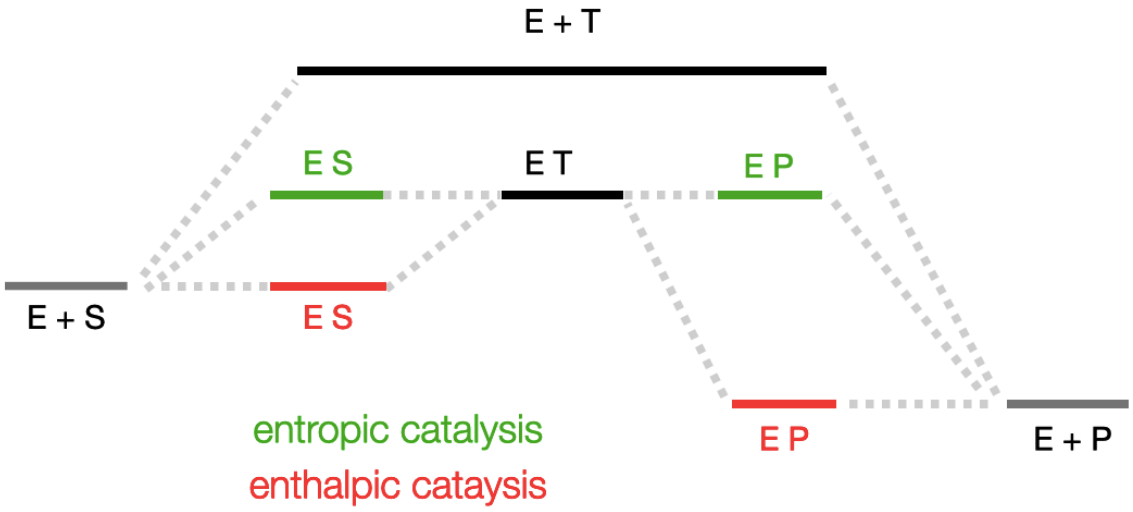


FIGURE 9-15. ENTROPIC VERSUS ENTHALPIC CATALYSIS

Most enzymes are predominantly enthlapic catalysts that lower the Gibbs energy of the TRANSITION-state through some combination of distributed or localized non-covalent bonds to protons *(*general acid/base), electrophiles/nucleophiles (H-bond donor/acceptor), monopoles (ionic bond), or dipoles (electrostatic bond). The ribosome peptidyl transfer center is a predominantly entropic catalyst that raises the Gibbs energy of the PRE- and POST-transition states by constraining the geometry of bound reactants (Sievers et al 2004; Schroeder & Wolfenden 2007). By positioning the electrophile and nucleophile, as well as lubricating exit of the nascent polypeptide, the transfer center and exit tunnel accelerate peptide bond formation some 10^5-10^7 fold.

Allowing a variety of amino acid sidechains, the ribosome interacts with universal features of the substrates and products, notably the $_{74}$CCA$_{76}$ handle at tRNA 3' ends, and the backbone of nascent polypeptides. Thus, the electrophile (carbonyl carbon) is positioned by pairing peptidyl-tRNA C74 C75 with G2251 G2252 in the P-loop atop H80, and A-minor interaction of A76 with A2450-C2501 (Samaha et al 1995; Kim & Green 1999; Nissan et al 2000). Similarly, the nucleophile (α-amine nitrogen) is

positioned by pairing aminoacyl-tRNA C75 with G2553 in the A-loop atop H92, as well as A-minor interaction of A76 with U2506-G2583.

The main propulsive (pushing) force on nascent polymers during elongation results from the transfer and translocation reactions that replace the P-site nucleotidyl-dRNA or peptidyl-tRNA by equivalent ones, now lengthened by two nucleotides, or one amino acid, respectively. Pulling on the nascent polypeptide, a network of conserved hydrogen bonds between the backbone of the three proximal residues and rRNA nucleotides G2061 A2062 U2506 ensures that it maintains an extended β-strand conformation needed for peptide bond formation and tunnel passage (Syroegin et al 2023). Additional tensile (pulling) forces result from the configurational freedom of segments emerging from the cramped tunnel, as well as the Gibbs energy of protein folding, and co-translational secretion (Leininger et al 2019). Whether pushing or pulling, these forces are transmitted along the polypeptide backbone with little attenuation by the tunnel wall.

In peptide bond formation, aminolysis has replaced hydrolysis, but one activated water molecule is still needed for polypeptide termination and release (section 15). Entropic trapping and desolvation are the predominant, but not sole contribution of the ribosome to peptide bond formation. Various methods including atomic structures, kinetic studies, rRNA mutations, and substrate mimics have identified likely enthalpic interactions that lower the free energy of the tetrahedral transition state by activation of the nucleophile, or stabilization of the leaving group, as well as proton relays that complement the electron movements (Zaher et al 2011; Polikanov et el 2014).

Early molecular biologists first demonstrated the independence of peptide bond formation from decoding *tout court* using nonsense suppressor mutations that changed tRNA anticodons, as well as using Raney-nickel to convert cysteine to alanine on the aminoacyl-tRNA (Chapeville et al 1962; von Ehrenstein et al 1963). Since then, for purposes of probing the substrate promiscuity of the ribosome, as well as applications to expand the genetic code, more refined methods have been developed for misacylation of tRNAs, e.g., ribozymatic charging *in vitro* using the *flexizyme*, and enzymatic charging *in vivo* using orthogonal tRNA/aaRS pairs.

The ribosome peptidyl transfer center and its tRNAs are finely tuned to form peptide bonds between α-L-amino acids. In its natural substrate promiscuity, it catalyzes peptide bond formation between any of some 20 x 20 = 400 pairs of proteinogenic amino acids at comparable rates. One difficulty of this balancing act is revealed by its slower rate when one or both residues are the imido acid α-L-proline (Doerfel et al 2013). Peptide bond formation with an α-D-aminoacyl-tRNA in the A-site is about 10^3 slower than with the corresponding α-L-aminoacyl-tRNA (Englander et al 2015).[11] The reason is that accommodation of the sidechain in the A-site cleft between C2452 and U2506 rotates Cα to position its H, not its amine for inline attack on the peptidyl-ester (Melnikov et al 2019). Besides its unfavorable attack geometry, the nucleophile is too far away from tA76 in the P-site for activation by hydrogen bonding to O2' or leaving group stabilization by protonation of O3'.

Synthetic biology has begun to explore the plasticity of the ribosomal peptidyl transfer center by design or selection of rRNA variants that polymerize α-D-amino acids, β-amino acids, α-hydroxyl acids, and more exotic monomers (Dedkova & Hecht 2019; Kofman et al 2021). Some monomers, such as β-, γ- and δ-amino acids or dipeptides, lengthen the polymer backbone. The native transfer center can accommodate β-amino acids, but longer monomers require modification of the rRNA or tRNAs to better position the reactants (Maini et al 2015). Whereas these products are polypeptides, α-hydroxyl acids produce polyesters, and other monomers produce polymers with more exotic bonds (Fahnestock & Rich 1971).

[11] Beyond their normally low concentrations in cells, α-D-amino acids are excluded at various steps in protein translation including (1) tRNA charging by aaRS enzymes, (2) aa-tRNA proof-reading deacylases, (3) aa-tRNA binding to EF-Tu, (4) reactivity in the peptidyl transfer center, and (5) ribosome arrest of the incorporated peptide.

10. What drove RNA elongation?

For the dawn of RNA life, we stipulated a feedstock of random oligonucleotides and compartmentation of spontaneous copying, but were otherwise agnostic about abiotic materials, energy sources, and reaction mechanisms (sections 1-2). In sections 8 and 9, we proposed that the duplisome and its dRNAs, tapping the existing feedstock of random oligomers, supplanted the original process of spontaneous RNA copying. Both parts of this processive elongation cycle, polynucleotide transfer within the duplisome and dRNA reloading without, were further analyzed as two covalent steps with ribozyme-bound intermediates. To wit, uphill condensation was coupled to downhill hydrolysis occurring immediately before (polynucleotide transfer) or after (dRNA loading). Like dRNA loading and polynucleotide transfer taken separately, the overall cycle of polymer elongation driven by oligomer shortening was thermodynamically isoergonic, and likely kinetically reversible (Ross & Deamer 2016).

If dRNA loading and polynucleotide transfer *tout court* do not explain the processivity of RNA duplication, we must look elsewhere for a source of Gibbs energy to drive the elongation cycle forward. First, we consider (and rule out) the possibility that duplication was pulled forward by favorable free energy of product folding. Next we consider how mass action of the [loaded dRNA] / [free dRNA] ratios might drive elongation forward. Although the oligonucleotide feedstock could shift these ratios to favor polymer elongation, this mechanism is vulnerable to pausing or even reversing with the vagaries of oligonucleotide supplies. We consider (and rule out) other ways to concentrate or activate these substrates, and remove or inactivate their by-products.

Having said all that we can about the initial inputs and final outputs of RNA duplication, we ask whether there was some proximate source of Gibbs energy, comparable to peptide bond formation, or GTP hydrolysis for polypeptide elongation in the ribosome, that operates within the RNA elongation cycle itself? Eschewing additional covalent intermediates, here we suggest the duplisome and its dRNAs formed a simple heat engine, harnessing the daily heating and cooling of the Earth's surface to drive RNA elongation. Our key idea is that useful work was extracted from a thermal cycle of opening dRNA hairpins during the hot day and closing them during the

cold night. Remarkably, the slow, but steady addition of one duplicon per day afforded sufficient generations for natural selection of RNA life. But with duplication times of 1 to 10 years, the requirements for covalent stability of RNAs placed strong constraints on habitable environments, and demand for mechanisms of damage prevention and repair.

In Figure 10-1 we show the reactants for RNA duplication from a pool of loaded dRNAs. The duplisome is a common catalyst restored to its initial state at the termination of copying, having cycled through various intermediate states each elongation step. Similarly, the template is unfolded for copying, but ultimately refolded to its original structure. Formally, it is an autocatalyst in that the major product, or duplicate is identical to the template, excepting any errors in copying, and possibly the ends.[12] Loaded and free dRNAs are stoichiometric reactants, *viz*. substrates and by-products, respectively. The net reaction does not change the total number of RNA molecules (or equivalently, the number of phosphoester bonds) but redistributes their lengths and sequence. Hence, the Gibbs energy available to drive duplication comes from (1) the reactions that reload dRNAs, and (2) differences of secondary or tertiary structure between these substrates (duplicon-dRNAs), their by-products (freed dRNAs), and the major product (duplicate RNA). Gains of base pairing and stacking in the folded duplicate are countered by losses of pairing and stacking between duplicon and anticodon in loaded dRNAs. Notwithstanding extremely stable RNAs with favorable stacking, base triples, ion binding *etc*., it seems unlikely that vagaries of product folding could drive a robust general mechanism of RNA copying (Weiss & Cherry 1993).

[12] In section 12 we consider problems of end replication and repair, including how the nucleotidyl transfer center, or ribozyme P might release the final dRNA from the 3' end of the duplicate.

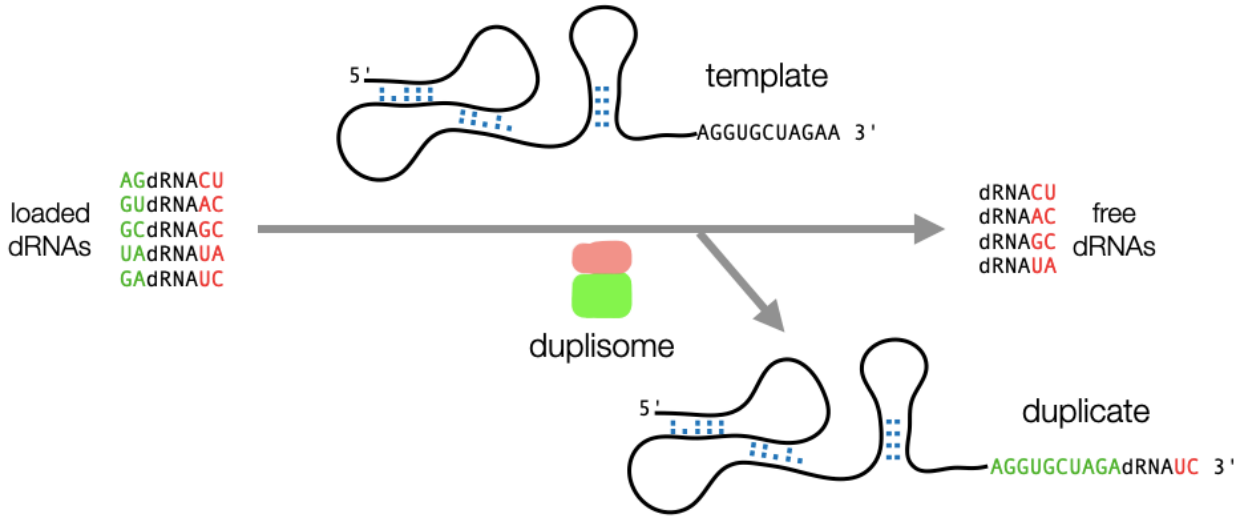


FIGURE 10-1. MASS ACTION OF LOADED AND FREE dRNAs

Could mass action of loaded dRNAs drive elongation? The reaction in Figure 10-1 is thermodynamically isoergonic, and perhaps kinetically reversible, with an equilibrium constant shown in Equation 10-1. The indices σ_i count how many times each of the 16 different duplicons occurs in the copy. The reaction is driven toward duplicon addition when loaded dRNAs are in excess, and toward subtraction when free dRNAs are in excess. Taxing our intuition, we imagine codon-directed subtraction from senescent polynucleotides as the reverse of codon-directed addition to nascent polynucleotides. In this spooky exercise of microscopic reversibility, we picture de-exit of closed dRNA into the E-site, opening this free dRNA for reverse translocation along with the polynucleotidyl-dRNA, reverse nucleotidyl transfer, and finally, closing the duplicon-dRNA with duplicon displacement of the codon from its anticodon, followed by de-entry from the A-site.

$$K_{eq} = \frac{k_+}{k_-} = \frac{[template + duplicate]}{[template]} \cdot \prod_{i=1}^{16} \left(\frac{[free]}{[loaded]} \right)^{\sigma_i}$$

EQUATION 10-1. MASS ACTION OF LOADED / FREE dRNAs

In Figure 10-2 we show the reactants in RNA duplication from a pool of random oligonucleotides. Once again, the template is an autocatalyst, while the duplisome, ribozyme P and free dRNAs are common catalysts. Of the stoichiometric reactants, oligomers consumed in dRNA loading are returned, shortened by two nucleotides at their 3' end, as by-products. Despite the rearrangements of primary structure, there is no net change in the number of RNA molecules or phosphoester bonds. Unspecified reactions that (1) consume shorter oligonucleotides to produce longer ones, as well as (2) replenish the mononucleotide pools, complete the elongation cycle. These feedstock reactions elevate the [loaded dRNA] / [free dRNA] ratios to drive elongation by producing longer oligomers (say 6 nucleotides or more) that are preferred substrates for dRNA loading, and consuming shorter oligomers that cause product inhibition in dRNA loading.

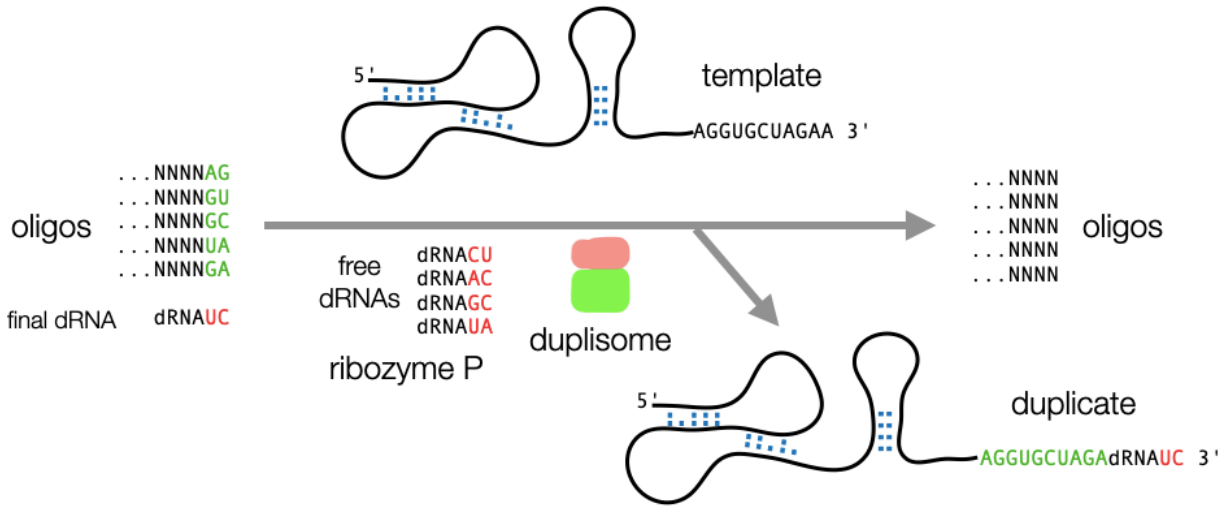


FIGURE 10-2. INPUT AND OUTPUT OLIGOMERS

Tracing back to LUCA, isoergonic and endergonic reactions are metabolically coupled, more or less directly, to the hydrolysis of ATP or GTP. Beyond the favorable Gibbs energy of the peptidyl transfer reaction (Table 9-1), driven indirectly by tRNA charging enzymes with *ATP transferase* activity, polypeptide elongation is coupled directly to translation factors with *GTP hydrolase* activity. Was there any comparable source of chemical free energy for RNA elongation, and a means of tapping it? In section 9, we noticed that the nucleotidyl transfer center might use the 5' OH in the A-

site, removing the 5’ phosphate of loaded dRNAs before duplisome entry, and restoring this phosphate to free dRNAs after duplisome exit. If so, a pair of ribozymes, *viz*. duplicon-dRNA 5’ phosphatase and dRNA 5’ kinase, might tap a high-energy phosphate donor to effectively shift the [substrate] / [by-product] ratios to drive elongation. There is great interest and uncertainty about prebiotic chemical activation and the first ribozymes to tap those high-energy donors. It is unclear which, if any transactions of the cellular NTP currency preceded, rather than followed, the evolution of ribozymatic RNA copying, or indeed protein translation itself (section 16). Beyond duplicons *tout court*, we pursue no other activated intermediates, preferring our parsimonious proposal that the polynucleotide 3’ OH made by hydrolysis in the P-site condensed immediately with the 5’ phosphate of the duplicon-dRNA in the A-site (section 9).

Eschewing fortuitous product folding, mass action of [loaded dRNA] / [free dRNA] ratios, or activated intermediates, was there any proximate means to drive RNA elongation? The non-equilibrium environments and Gibbs energy sources of the hadean Earth, both steady and episodic, are challenging to comprehend. One exception is sunlight which delivered a large flux of energy to the Earth’s surface, modulated in daily and yearly cycles by planetary motion. Acknowledging this streetlight (or better, sunlight) bias in our search, here we consider whether radiant energy from the hadean Sun drove RNA duplication. Rather than coupling the elongation cycle to sunlight directly, we suggest that the opening and closing of dRNAs within the duplisome was entrained to daily heating and cooling of the Earth’s surface, driving regular addition of one duplicon per day. One simple model is that duplisome life domesticated the principal energy source of spontaneous copying, *viz*. dry-down condensation in shallow pools, for processive copying in the physiological mileau.[13] To wit, *the thermal duplisome worked where the heat was not so extreme to melt all duplexes, nor to evaporate all water to dryness, that is, under conditions that allowed concurrent functions of RNAs as ribogenes and ribozymes.*

[13] A thermal physiological process might also have arisen from spontaneous RNA copying driven by regular freezing and thawing of shallow icy brines (ref).

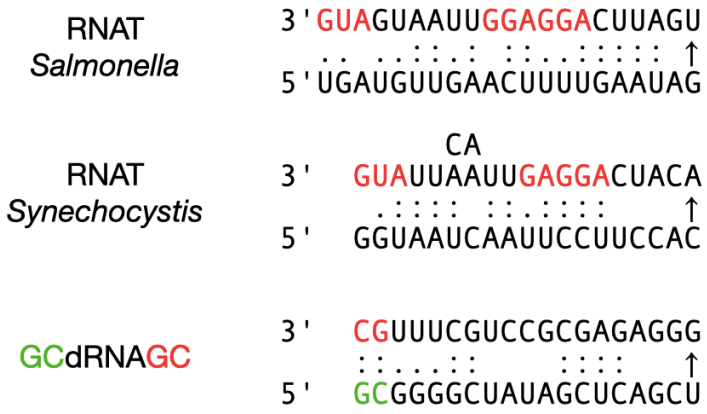


FIGURE 10-3. RNA THERMOMETERS & LOADED dRNA

RNA hairpins are used in bacteria as thermal riboswitches to control mRNA translation in response to temperature (Kortmann & Narberhaus 2012). In a change of perspective, we regard RNA hairpins as thermal ribomotors that perform mechano-chemical work. In typical RNA thermometers found in the 5' UTR of an mRNA, a critical hairpin masks the Shine-Dalgarno (SD) box, preventing protein translation in the cold. Upon warming, melting of this hairpin unmasks the SD box and initial AUG for translation. In Figure 10-3 we compare the critical hairpins of two RNA thermometers to GCdRNAGC modeled on bacterial $tRNA^{Ala}$. The SD box and AUG start codon in these bacterial RNA thermometers are shown in red.

$$W = \Delta T / T_m \cdot \Delta H_{open}$$

EQUATION 10-2. WORK AVAILABLE FROM HOT OPENING AND COLD CLOSING OF HAIRPIN RNA

The principle of our heat engine is that RNA hairpins are harder to pry open when cold, than when warm. The Gibbs energy available from this cycle can be calculated from the melting temperature (T_m) and opening enthalpy (ΔH_{open}) of the hairpin, along with the temperature difference (ΔT) between the hot and cold baths (Privalov 2012). If the melting temperature is say 350 K, and these baths are 35 K apart, up to 10% of the

enthalpy of melting can be extracted as useful work, assuming the enthalpy and entropy of opening are insensitive to temperature (Equation 10-2). By opening the duplicon-dRNA in the A-site during the day and then closing the freed dRNA in the P-site during the night, the duplisome extracts as much as 10 kcal / mol useful work, comparable to hydrolysis of 1 GTP to GDP in cellular metabolism (Lehninger). In Figure 10-4 we depict a Gedanken cycle of opening GCdRNAGC in the warm bath and then closing dRNAGC in the cold bath. Using the RNAfold Vienna program to estimate the free energies of opening and closing these hairpins at 310 K and 273 K, respectively, we obtained a net free energy of nearly -9 kcal / mol / cycle (Gruber et al 2008; http://rna.tbi.univie.ac.at/).

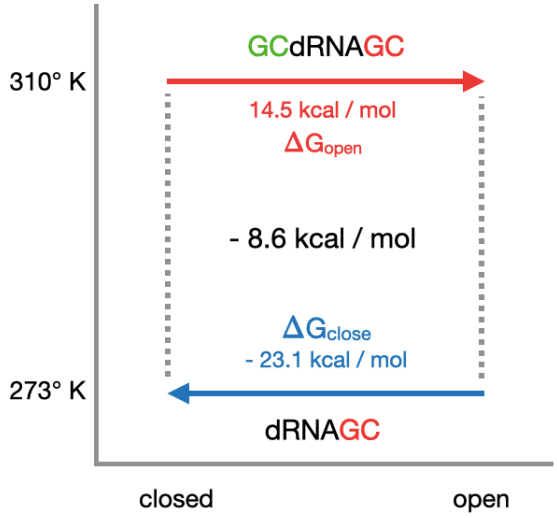


FIGURE 10-4. HEAT CYCLE OF dRNA OPENING AND CLOSING

In section 9 we proposed that dRNAs close for duplicon loading, and open for nucleotidyl transfer, respectively. Here we suggest that the loaded dRNA opens, and the freed dRNA closes *within the duplisome itself*, to drive decoding and translocation, respectively. Before discussing this elongation cycle, we remark on the secondary and tertiary structure of dRNAs. If closed dRNAs were more-or-less regular hairpin extensions of the D-arm of modern tRNAs (*cf*. Figure 9-3), what might open dRNAs look like? In Figure 10-5 we model an open dRNA on the tertiary structure of yeast $tRNA^{Phe}$ (Kim et al 1974; Robertus et al 1974). None of the interactions with the 3' half (nucleotides 37-76) would have been present, but any stacking or hydrogen bonding

between nucleobases contained strictly within the 5' half of the modern tRNA (nucleotides 1-36) may have been present in the dRNA ancestor. These include conserved base triples U8-A14-A21 and A9-U12-A23 that pin the **duplicon-leader** back against the **D-arm**. Conceivably, this augmented D-helix held the leader open, **hinged** back between A9 and G10, to accommodate the nucleotidyl transfer center. Conversely, the **trailer** from G26 to the anticodon A35 A36, has no interactions with the D-arm, suggesting it swung freely to pair with the template in the decoding center.

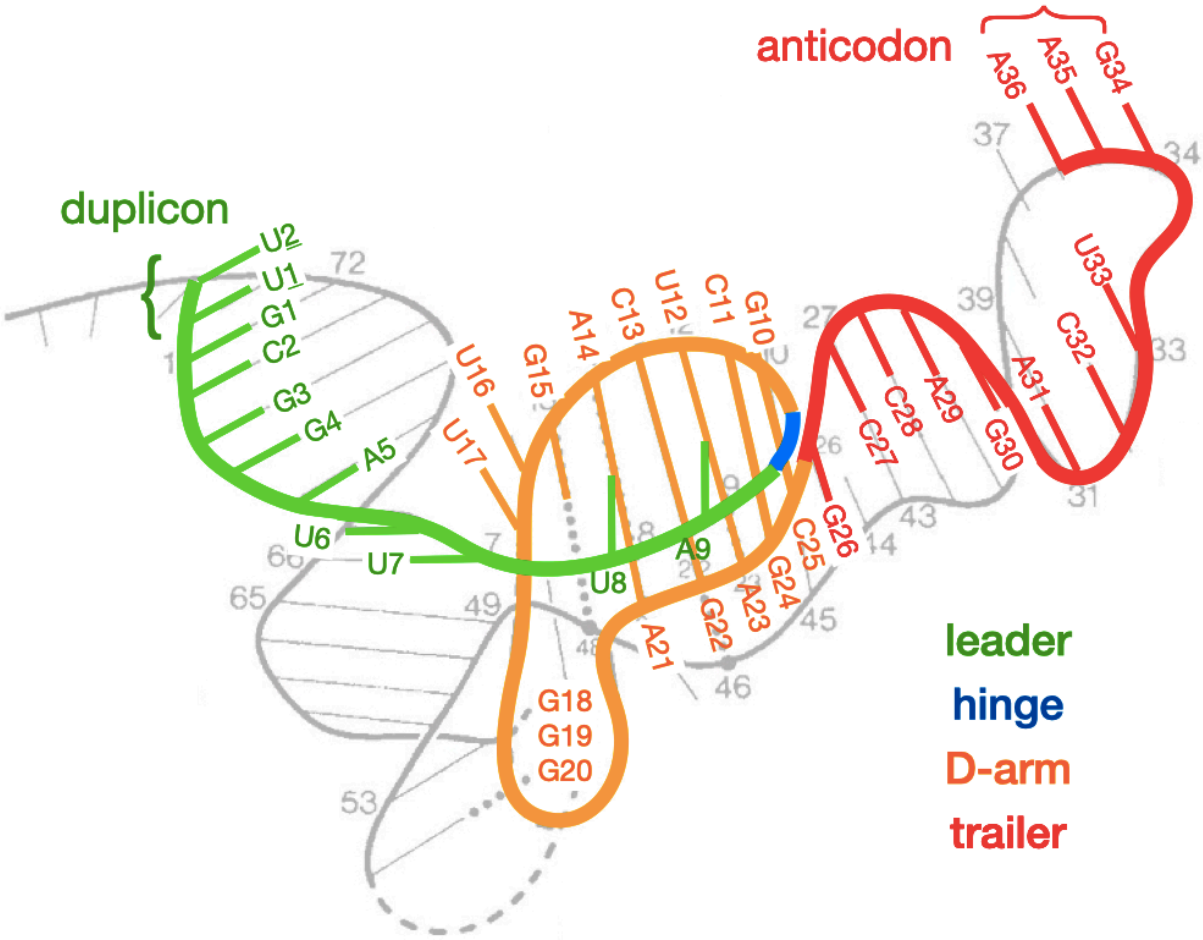


FIGURE 10-5. OPEN dRNA THREADED ON tRNA (AFTER ROBERTUS ET AL 1974)

In the duplisome decoding center, the dRNA is pried open via codon-anticodon pairing that displaces the duplicon and holds on until the entire leader has folded back onto the D arm as the day warms (section 11). Large deformations and rigid rotations of aminoacyl-tRNA during accommodation may be vestiges of this opening of duplicon-dRNA (ref). Indeed releasing factors, likely protein mimics of lost polypeptide

terminator tRNAs, show analogous movements of accommodation (ref). After nucleotidyl transfer, the freed dRNA closes to exit from the duplisome. We suggest this duplisome-bound opening of the duplicon-dRNA during the hot day, and closing of the freed dRNA during the cold night provided the Gibbs energy to drive an otherwise, nearly isoergonic elongation cycle for RNA duplication. Although dRNAs combine motoring with loading and transfer in an elegant way, thermal drive is not limited to hairpin structures. Thus, regions of the rRNAs themselves might also undergo reversible changes in secondary or tertiary structure to drive elongation. Perhaps relevant here, the Alu domain of SRP RNA occupies the ribosome factor binding site to pause elongation (Ahl et al 2015).

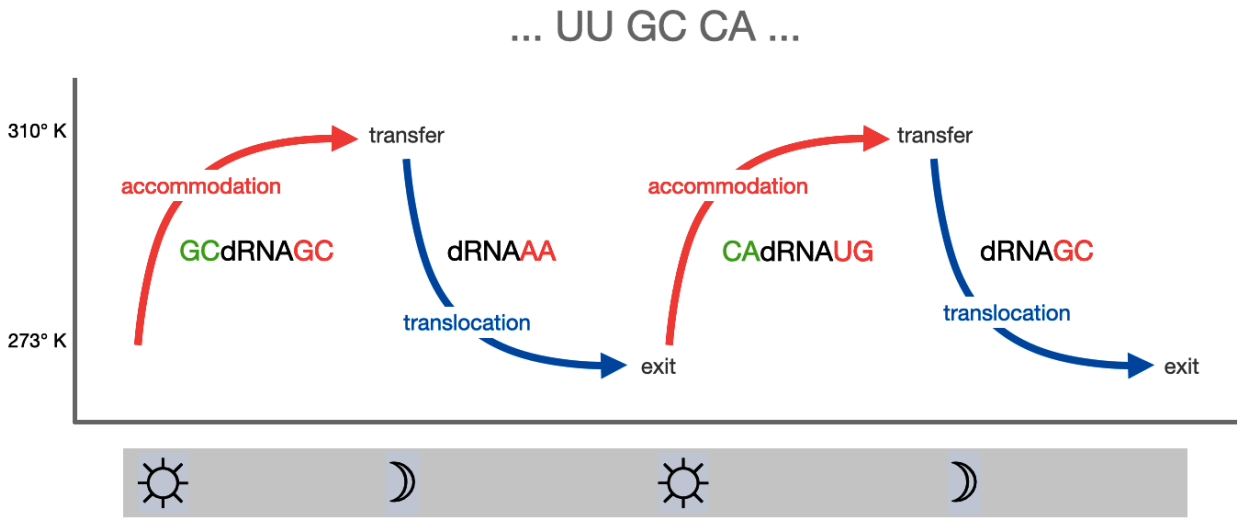


FIGURE 10-6. THERMAL ELONGATION CYCLE

In Figure 10-6 above we illustrate the thermal elongation cycle where heating drives decoding and cooling drives translocation with dRNAs acting as thermal ribomotors. On the first morning, closed GCdRNAGC enters the A-site where its anticodon gains a toehold with the cognate codon GC, prying the dRNA open to accommodate. Near midday the nascent polynucleotidyl-dRNA and duplicon-dRNA undergo an isoergonic exchange in the nucleotidyl transfer center. On the first evening the free dRNAAA closes, driving translocation and exit. On the second morning the closed CAdRNAUG enters the A-site for another decoding. Thus, the original GCdRNAGC exits as a freed dRNAGC on the second evening. In this way, one new duplicon is added each day, yet

any particular dRNA takes two full days from entry to exit. A full reaction cycle of two successive days is shown in Figure 10-7. Finally, we notice that the sequential reactions of dRNA loading by ribozyme P may have been entrained to the diurnal temperature cycle, charging at night when the entropic cost of substrate docking was low, and trimming in day when the entropic benefit of product release was high.

...UUdRNAAA + GCdRNAGC ⇔ dRNAAA + ...UUGCdRNAGC

...UUGCdRNAGC + CAdRNAUG ⇔ dRNAGC + ...UUGCCAdRNAUG

FIGURE 10-7. COVALENT REACTIONS OF ELONGATION CYCLE

Could duplisome life have evolved adding (at most) one duplicon per polynucleotide per day? There are really two questions concerning the conservative and the creative role of heredity, respectively: (1) Could this rate of synthesis have outpaced decomposition? (2) And would there be enough rounds of variation and selection for Darwinian evolution? Stipulating the first answer is yes, the second is clearly yes as well. If we assume one duplicon was added per day over a 6 month growing season, we could add some 360 nucleotides per year. Under these conditions 3600 nucleotides would be duplicated in 10 years, sufficient for modern rRNAs. The generation time could, of course, be shorter if ancestral rRNAs were duplicated in smaller fragments.[14] The population-doubling time could be shorter if polyduplisomes shared one template just as polyribosomes share one mRNA. Finally, the Earth's rotation has slowed since the hadean when there were likely twice as many days per year (Gordon & Mikhailovsky 2021).

A doubling time, or even a generation time, of just 1 year is conceivable for duplisome life. But even if a generation required 10 years, and duplisome life lasted only 200 million years, there would have been 20 million generations, a figure comparable to say the mammalian radiation of recent evolution. Major stages and transitions in duplisome life, from duplisome-mediated combination and codon-

[14] Besides the universal partition of SSU and LSU rRNAs, there are three precedents for further partition: In eukaryotes, cytoplasmic LSU rRNA is fragmented into 5.8S and 28S rRNAs. In protosomes, 28S rRNA is fragmented into nearly equal halves called 28Sα and 28Sβ (Natsidis et al 2019). Finally, in *Euglena gracillis*, 28S rRNA is fragmented into multiple pieces (Matzov et al 2020). ?insect 5.8S cleavage?

directed polynucleotide termination to the breakout of polypeptide translation and retirement of RNA duplication, are discussed in sections 12-17.

But what of the chemical stability of the RNA backbone and its nucleobases? Protected from environmental insults, *viz*. UV radiation, exogenous nucleophiles, biotic nucleases, RNA in neutral solution is still prone to spontaneous scission of the backbone, as well as hydrolytic deamination (cytosine) or depurination (adenosine, guanosine). At 20-25 °C, the half-life of one phosphodiester linkage in water at pH 6-7 is estimated between 5-1000 years (Eigner et al 1961; Li & Breaker 1999; Wolfenden 2011; Ross & Deamer 2016; Lonnberg 2023). The stability of any phoshodiester linkage is significantly greater in paired regions. Similarly, the rate of spontaneous deamination of cytosine decreases over 100-fold in paired regions. In some models of spontaneous RNA copying, much of the ribogenome may have been paired much of the time, and an estimated 50-70% of nucleotides in diverse RNPs, ribozymes, and riboswitches are base paired (virtual circular genome). So long as one and same polyribonucleotide acts as both ribogene and ribozyme, however, exposed unpaired regions were a fact of RNA life.

It is unclear whether a copying time of 180 days for an RNA of 360 nucleotides is fast enough to stay apace its half-life from spontaneous decomposition, estimated anywhere from 5-1000 days. What is clear, however, is that common forms of covalent chemical damage placed strong constraints on diurnal elongation, or any comparably slow process of RNA replication: First, habitable environments were constrained by the need to minimize RNA decomposition, favoring conjectures say that ribozymes functioned in dark and icy brines. Second, there was strong selection for shorter, or less damage-prone ribozyme sequences and folds, including pairing as much as possible. Finally, there was strong selection for ways of protecting critical sequences, as well as repair mechanisms. We return to some of these problems in section 12.

Whatever the original solutions of RNA life for folding and regulation of ribozymes, as well as replication, protection and repair of ribogenes, in later stages of polymer life, enzymes augmented and displaced ribozymes as ephemeral gene activities (sections 16 & 17), and deoxyribogenes replaced ribogenes as durable genes (section 18). Thus, the average phosphodiester bond in duplex DNA has an estimated half-life of 31 million

years at 25 °C in neutral water (Wolfenden 2011), and remarkable mechanisms have evolved for detection and error-free repair of common DNA damage. Attesting the stability of deoxyribogenomes under conditions of dormancy, as well as cellular mechanisms of resistance, plant seeds about 2000 years old, as well as bacterial spores about 250,000 years old have been germinated (Setlow & Christie 2023).

11. The RNA code

The universal genetic code maps codons to amino acids, and by iteration, maps mRNAs to polypeptides. This code is the product of two simpler relations involving tRNAs. Complexed with its cognate aminoacyl-tRNA synthetase (aaRS), each tRNA is charged at its 3' end with the correct amino acid, and then, complexed with the ribosome decoding center, matched at its anticodon to the mRNA codon. The composite function is many-to-one, *i.e*. *synonymous codons* assign the same amino acid, and only partially defined, *i.e*. *nonsense codons* assign no amino acid. Iterating from one codon to an entire mRNA, deacyl- and peptidyl-tRNAs are shifted to the E- and P-sites, respectively, tugging the next codon into the A-site.

The RNA code maps codons to duplicons, and by iteration, maps templates to duplicates. Like the amino acid code, this older code is the product of two relations involving dRNAs. Complexed with ribozyme P, each dRNA is charged at its 5' end with the correct duplicon, and then, complexed with the duplisome decoding center, matched at its anticodon to the template codon. The composite function is one-to-one, and formally, the identify. Iterating from one codon to an entire template RNA, freed- and polynucleotidyl-dRNAs are shifted to the E- and P-sites, respectively, tugging the next codon into the A-site.

The polymer mappings from template RNA to duplicate RNA, or from mRNA to polypeptide, are remarkably accurate, but not perfect. As in protein translation, the fidelity of RNA duplication depends on the accuracy of various steps including (1) dRNA loading by ribozyme P, (2) duplicon-dRNA selection by the duplisome decoding center, (3) polynucleotide transfer in the duplisome nucleotidyl transfer center, and (4) translocation of dRNAs and template. Errors in each step produced characteristic defects in the duplicate, i.e., nucleobase substitutions from errors in dRNA loading and decoding, premature polynucleotide release from errors in nucleotidyl transfer, and indels from frameshifts during translocation and decoding. Several mechanisms have evolved to suppress these errors, and improve the fidelity of polymer mapping. Some of these mechanisms arose in polynucleotide elongation and others pertain specifically

to polynucleotide termination (sections 13 & 14) and polypeptide translation (sections 15 & 16).

One likely way of improving the accuracy of RNA copying is to lengthen the codon-anticodon helix from two to three basepairs (Grosjean & Westhof 2016). Naively, this comes at the cost of increasing the number of dRNAs from 16 to 64 (*pace* Campbell 1991), or allowing degeneracy of the product (*pace* Noller 2010/2012). Here we suggest a way to attain nearly the accuracy of triplet decoding with only 16 dRNA isoacceptors dubbed *superwobble*. *We are agnostic whether dRNA superwobble decoding replaced doublet decoding outlined in section 8, or was in fact the original decoding mechanism.* Before we refine our concepts of step size in polynucleotide elongation, we précis five intertwined concepts of step size in polypeptide elongation (Table 11-1). To wit, reading matches *one codon* with *one anticodon*, transfer adds *one residue* to the nascent polypeptide, and translocation moves tRNAs *one site* in the ribosome, and concurrently, moves the mRNA *one translocon*.

	read codon & anticodon	add duplicon	move dRNA	move template
duplication superwobble	2 ✻ nt	2 nt	1 site	2 nt
translation superwobble	2 ✻ nt	1 aa	1 site	3 nt
translation WC triplet	3 nt	1 aa	1 site	3 nt

TABLE 11-1. FIVE CONCEPTS OF STEP SIZE IN PROCESSIVE POLYMER ELONGATION

The codon and anticodon, *viz*. stretches of mRNA and tRNA matched at the decoding center, define the size of one another, and proved to be triplets (Nirenberg et al 1963; Holley et al 1965; Jones & Nirenberg 1966). However, the naive idea of *complementary triplets* proved inadequate as the rules of nucleobase matching as well as the underlying mechanism vary by nucleotide position (Ogle et al 2001; Demeshkina

et al 2012). Thus, the first two codon positions are restricted to Watson-Crick basepairs, but the third position is more promiscuous. In his *tRNA wobble hypothesis*, Crick noticed that an anticodon U34 could read either A or G in the third codon position, reducing the number of tRNAs needed for the protein code (Crick 1966). In this, and other cases, the map from codon to anticodon is not invertible because the identity of the third codon position is partly lost in decoding, and two or more tRNA isoacceptors can read the same codon. In mitochondria and plastids, as well as reduced bacterial genomes such as *Mycoplasma*, a single tRNA with unmodified U34 can read a four-codon box, e.g., $tRNA^{Ala}$ (UGC) reads all four alanine codons GCN (Bonitz et al 1980; Heckman et al 1980; Andachi et al 1989; Rogalski et al 2008; Alkatib et al 2012). Known as *tRNA superwobble*, the adaptor and decoder merely confirm the presence of the third nucleotide, but ignore its identity.

The translocon proved to be a triplet as well, so that the ribosome translates a "comma-free" succession of codons with no gaps or overlaps. In deciphering the protein code, molecular biologists had considered other formal possibilities including triplet codons with doublet translocons, or doublet codons with triplet translocons (Crick 1968). The former in effect reads odd-numbered nucleotides twice, first in the third codon position, and then again in the first codon position, while the latter ignores the identity of every third nucleotide, *viz*. what is now known as superwobble decoding.

The two formal requirements for faithful RNA copying are that the duplicon matches the template, and that the translocon is the same length as the duplicon, not one nucleotide more, nor one less. Thus, there is no reason that the decoding center cannot match codon-anticodon triplets, so long as the duplicon and translocon are both doublets or both triplets. Here we modify the scheme of doublet decoding with doublet addition from section 8, to incorporate dRNA superwobble, *viz*. monitoring the presence of the third position nucleotide, but ignoring its identity. There are two likely advantages to triplet decoding with doublet addition. First, as discussed below, the third basepair improves decoding accuracy by increasing the stability of the codon-anticodon helix (Grosjean & Westhof 2016). Second, as discussed in section 12, we suggest that the wobble position monitors template termination.

Early theories of ribosomal frameshift errors explained reading frame iteration as an active mechanism of translocation, followed by passive entry of aminoacyl-tRNA into an empty A-site. In fact, the division of labor between translocation and tRNA selection is more fluid, so that the incoming tRNA helps to determine, not simply respect, the current frame. Among the most common errors are shifts of the reading frame upstream on the mRNA by 1 nucleotide called -1 frameshifts (Figure 11-1). Frameshift can occur by *one tRNA slippage* in the POST-translocation ribosome with an empty A-site. Here the peptidyl-tRNA slips -1 nucleotide, facilitated by a slippery mRNA sequence X XXY, as well as depletion of the cognate aminoacyl-tRNA. Frameshift can also occur after peptidyl transfer by *two tRNA slippage* in the PRE-translocation ribosome. Here the deacyl-tRNA and peptidyl-tRNA together slip by -1 nucleotide, facilitated by slippery mRNA sequence X XXY YYZ, and programmed by downstream hairpins or pseudoknots. Frameshift can perhaps also occur during decoding itself by what we call *two tRNA scrunch* where nucleobase N34 of the peptidyl-tRNA flips out of the anticodon stack, allowing nucleobase N36 of the incoming aminoacyl-tRNA to read the -1 nucleotide (Licznar et al 2002; Atkins & Bjork 2009).

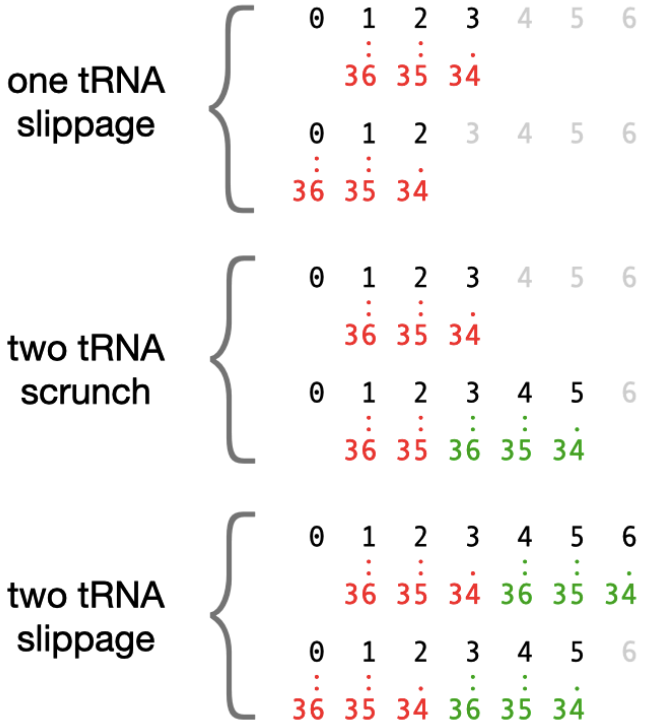


FIGURE 11-1. RIBOSOMAL MECHANISMS OF -1 FRAMESHIFT

One tRNA slippage, two tRNA scrunch, and two tRNA slippage, may be distinguished by their intermediate states and regulation, but the nascent polypeptide is identical for all three pathways. Analogous mechanisms of one and two tRNA slippage have been proposed for +1 frameshifts as well as two tRNA spread where the peptidyl-tRNA in the P-site not only defends its wobble position, but spreads out to the next nucleotide downstream, forcing the incoming aminoacyl-tRNA to read the +1 codon. Two tRNA spread may explain how a tRNA with an enlarged anticodon loop pairs normally in the A-site yet causes +1 shifts in the P-site.

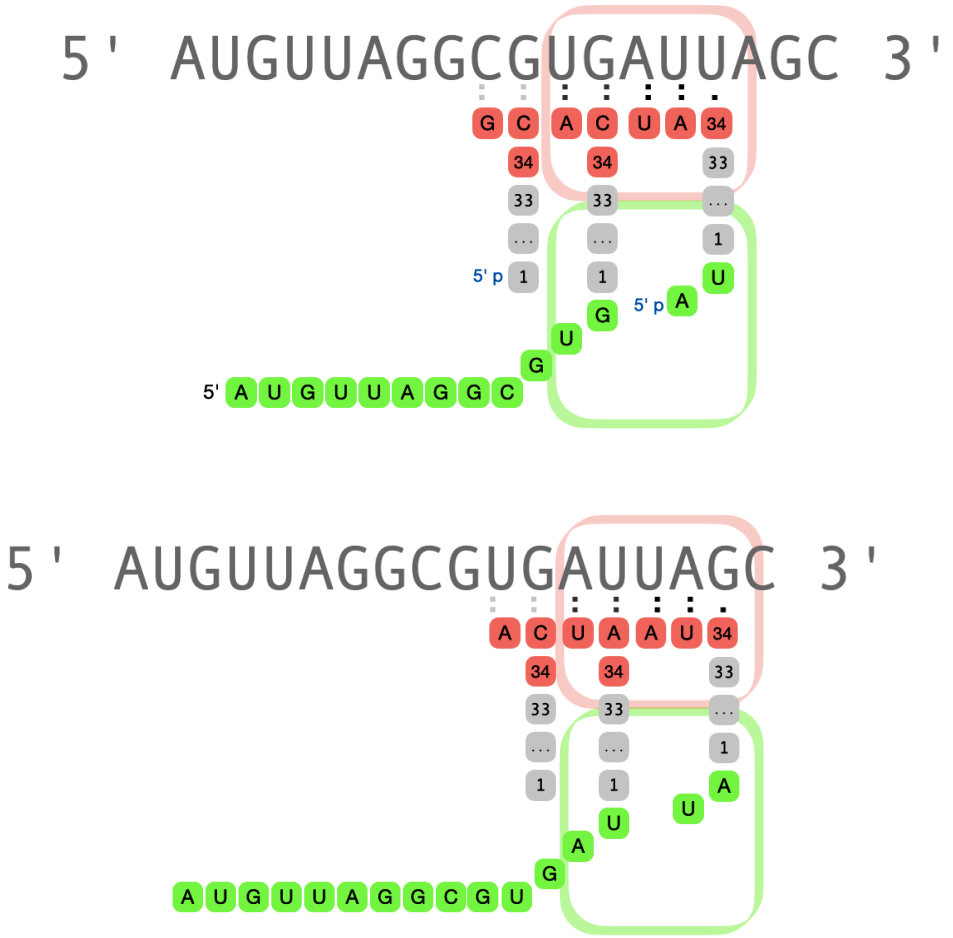


FIGURE 11-2. dRNA SUPERWOBBLE DECODING WITH DOUBLET ADDITION [EDIT FIGURE]

Here we suggest that the duplisome and dRNAs combined superwobble decoding with doublet addition in the normal cycle of RNA elongation (Figure 11-2). In this best of both worlds, the triplet codon-anticodon helix provided stability for accurate

decoding, while the doublet duplicon and translocon provided faithful copying with just 16 dRNA isoacceptors. *To be clear, two dRNA scrunch was the normal mechanism of duplisome decoding whereas two tRNA scrunch, if it occurs at all in the ribosome, is a form of -1 frameshift error.* Finally, although dRNA U34 superwobble became the universal mechanism of triplet decoding in RNA duplication, analogous cases of tRNA U34 superwobble are derived, not primitive characters in protein translation. Indeed tRNA superwobble has arisen independently a number of times in bacteria and eukaryal organelles.

The fidelity of dRNA decoding in polynucleotide elongation was determined, we suggest, by kinetic competition between the codon and duplicon for their shared anticodon (Simmel 2023). To wit, the duplisome sampled prospective duplicon-dRNAs that entered the A-site in their closed conformation. Any non-cognate or near cognate dRNA left quickly, allowing the duplisome to sample another prospective donor. Sampling the set of 16 isoacceptors without replacement, it would need 8 1/2 trials on average to encounter the cognate dRNA. More realistically, sampling these isoacceptors with replacement, it would need 16 trials on average. Biases in the frequencies of codons and isoacceptors could increase or decrease the mean sample size in decoding. Once the duplisome encounters the cognate donor at the A-site, the codon can displace its duplicon from the anticodon, retaining the dRNA through this toehold until it fully opens as the day warms and the duplicon leader arm is accommodated in the nucleotidyl transfer center. We illustrate this toehold competition and dRNA opening in Figure 11-3. In the parlance of polynucleotide strand exchange, the *codon-invader* **GCG** displaces the *duplicon-incumbant* **GCA** from the *anticodon-substrate* **UGC**.

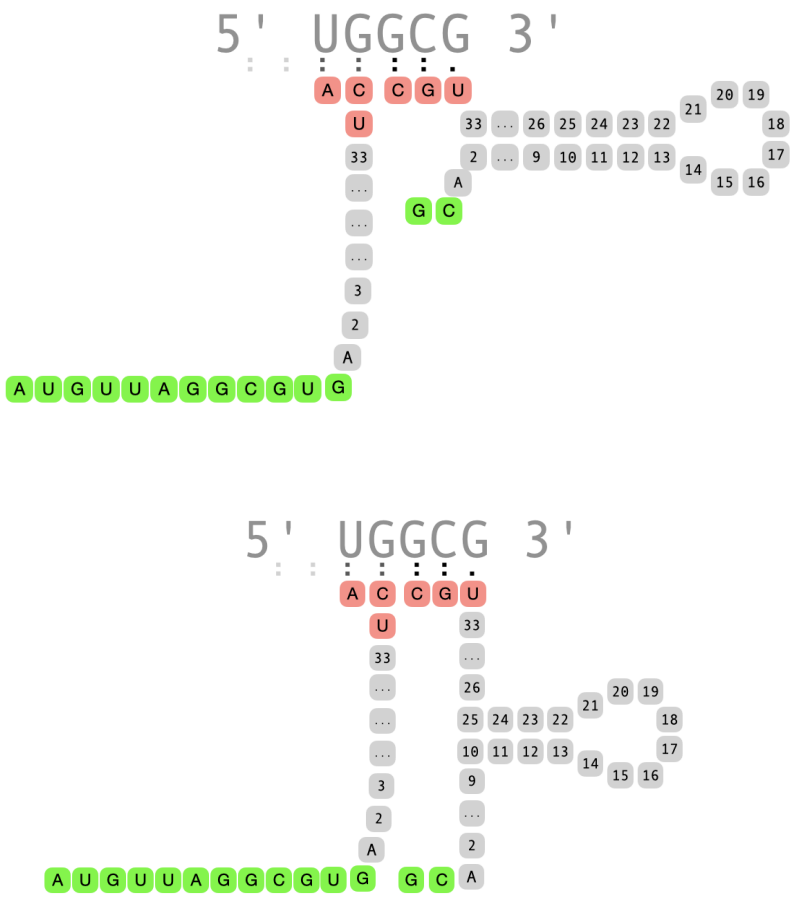


FIGURE 11-3. DECODING BY DUPLICON DISPLACEMENT

There are two serious rivals to cognate dRNAs in the competition for decoding: near-cognate dRNAs that match only the first or the second codon position, and unloaded yet cognate dRNAs that fully match the codon but lack a duplicon. When no cognate dRNA is available, hungry duplisomes likely became more promiscuous as the day warmed, accommodating near-cognate dRNAs to avert premature termination at the cost of a nucleotide substitution. Unloaded dRNAs present a graver threat to the polynucleotide elongation cycle if they can freely enter the A-site. Conceivably, the duplisome somehow screened all dRNAs for the presence of a 5’ duplicon before allowing access to the codon. If not, an unloaded yet fully cognate dRNA could readily pair with the codon to dwell in the decoding center via its initial codon-anticodon toehold. It is unclear whether such dRNAs would eventually accommodate, or just get out of the way. *Thus, like a flail, the thermal momentum of the displaced duplicon, not*

present in the unloaded dRNA, might be needed to shake open the adjacent stem. If an unloaded dRNA did accommodate, it would presumably cause polynucleotide release.

Table 11-2 shows the 16 dRNA isoacceptors of triplet decoding with doublet addition based on U34 destacking (scrunch) in the P-site, and U34 superwobble in the A-site. To be definite, during dRNA loading the universal U34 pairs with the universal A1. During decoding, polynucleotide-dRNA U34 in the P-site yields to the incoming duplicon-dRNA N36 in the A-site, while the duplicon-dRNA U34 matches any nucleotide in the third codon position.

1st \ 2nd	U	C	A	G
U	UUAdRNAUAA	UCAdRNAUGA	UAAdRNAUUA	UGAdRNAUCA
C	CUAdRNAUAG	CCAdRNAUGG	CAAdRNAUUG	CGAdRNAUCG
A	AUAdRNAUAU	ACAdRNAUGU	AAAdRNAUUU	AGAdRNAUCU
G	GUAdRNAUAC	GCAdRNAUGC	GAAdRNAUUC	GGAdRNAUCC

TABLE 11-2. THE RNA CODE OF 16 ELONGATOR dRNAs

12. RNA duplication | repair & recombination

A bold conjecture about the dawn of life on Earth, the RNA world proposes to describe the phenomena, and explain the principles of life, in plain language of polymer chemistry. As in other attempts to understand adaptive systems and intelligent behavior *ab initio*, although the implementation of RNA life is simple, the divisions of labor between genes (hereditary polymers) from their products (intermediate and final polymers, metabolites) became clearer in later stages of polymer life (sections 13-18). These were the first examples of the emergence of new levels of selection that John Maynard Smith and Eors Szathmary called *major evolutionary transitions* (Maynard Smith & Szathmary 1995). Thus, earlier evolutionary stages have fewer levels and manifest simpler phenomena, while later stages have more levels and reveal clearer principles. So, we first examine the mechanisms of RNA life bottom-up from polymer chemistry, and later analyze their functions top-down into roles made clear and distinct in cellular life and yet later stages. In section 19 we frame the evolution of polymer life, and beyond, in the language of search processes. There what biologists call affordances, genome and epigenomes, respectively, are *mazes,* or interfaces of tests and actions with our surroundings, *maps,* or durable records of past maze exploration, and *plans*, or ephemeral marks on these maps for present exploitation.

The great insight of Charles Darwin and Alfred Russell Wallace was that living species explore immediate affordances, as well as invent intermediate ones, for no higher purpose than preserving and propagating these maps (ref). Our DNA genome and its epigenomic marks are examples *par excellence* of an evolving map and its working plans, respectively (Jablonka & Lamb 1995; Maynard Smith & Szathmary 1995; book). Following Darwin and Wallace, nineteenth century biologists skirted about the unknown nature of heredity in both the theory of natural selection, and particular theories of embryology, metabolism, physiology, behavior, *etc*. Opposing poles of search were widely recognized, *viz*. exploring novel affordances (evolution) versus exploiting familiar ones (biological regulatory mechanisms), but not the continuum between. Understanding the process of heredity has occupied biologists ever since. Even today, the creative and conservative functions of heredity, commonly identified

with horizontal and vertical gene transmission, are studied separately as evolutionary and regulatory biology, respectively.

The first, and easier task of modern biology has been to characterize the habitual plans of living species that exploit familiar affordances, *viz*. the multilevel regulation of metabolism, cell biology, development, immunity, physiology and behavior. These are all grounded in genome replication and gene expression, the original mechanisms derived from polymer life for preserving maps and following compiled plans to familiar goals. Here, the conservative functions of heredity, captured in the term *vertical gene transmission* (VGT), entail faithful genome copying, protection and repair, purifying selection to eliminate errors in gene replication, as well as quality control of gene products.

One principal conservative function of any catalyzed process of genome replication is to faithfully copy its own replicase (Eigen 1971). In our model of processive RNA copying this includes the duplisome, ribozyme P, and a working set of dRNAs. Whatever the actual copyase ribozyme(s), speculations about RNA life have focused on the speed and fidelity of copying its larger components, as well as folding and stability of the active replicase. Like medieval manuscripts before the invention of printing, long polynucleotides were premium goods in RNA life, copied slowly and accurately, and folded with care. There was strong selection, moreover, to prevent or repair common forms of RNA damage from spontaneous or biotic insults.

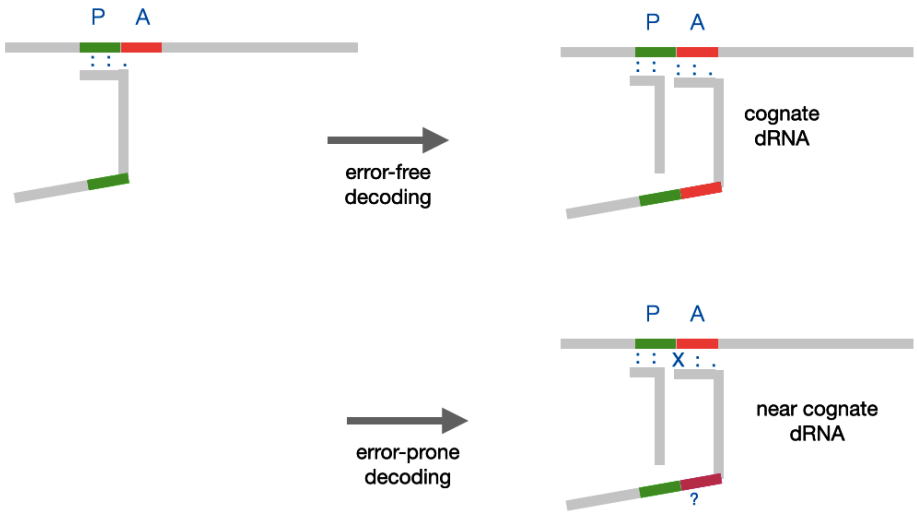


FIGURE 12-1. ERROR-PRONE dRNA DECODING

Processivity of RNA duplication no doubt took priority over fidelity. If the cognate dRNA were depleted, the decoding center likely accommodated a near cognate dRNA as the day warmed, hazarding a substitution error over indefinite delay, or premature termination (Figure 12-1). Encountering damaged or missing nucleobases in the template, the duplisome likely skipped over them to reinitiate at a downstream codon, producing a deletion but no break in the copy strand (Figure 12-2). In the absence of a proximate source of Gibbs energy for processive scanning, diffusion alone could only skip short distances (in perhaps either direction). Finally, as the duplisome unfolded the downstream template, it likely continued past frank breaks in the backbone by rethreading the nearest free 5’ end (Figure 12-2). The copy might suffer a point mutation, but an otherwise complete duplicate could be made. Although a gapped template could refold after copying, it was not covalently repaired by this means. Finally, whether a ribozyme assembled from several small chains, or folded as a single large chain, it was important to copy each polynucleotide as close to both ends as possible. There was strong selection to protect these ends from chemical erosion, and to restore any nucleotides shortened in copying, or degraded by wear.

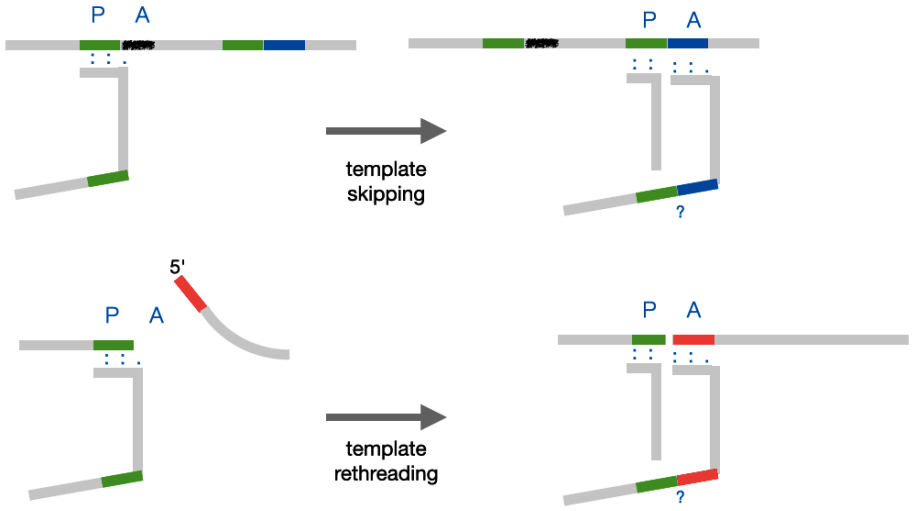


FIGURE 12-2. SKIPPING DAMAGED TEMPLATE & RETHREADING BROKEN END

In protein synthesis, the first (initiation) codon of the open reading frame directs the start of polypeptide translation, defining the reading frame, while the last (termination) codon directs polypeptide release. Initiation and termination are coupled to the cycle of ribosome subunit assembly and disassembly, respectively. In this way, mRNAs with long 5' or 3' UTRs are swapped in and out without translation from the beginning, or to the end. Such precise control of polypeptide ends requires various initiation, termination, and recycling factors, dedicated START and STOP codons, and the initiator $tRNA_i^{(f)Met}$ (section 16).

Viral mRNAs, organelles, and cellular stress have revealed simpler mechanisms of protein initiation and termination (Yamamoto et al 2016; Beck & Moll 2018; Huber et al 2019; Saito et al 2020; Leiva & Katz 2022). Leaderless mRNAs require only the START codon, not upstream elements such as Shine-Dalgarno sequences (prokarya), or 5' caps (eukarya). Under some conditions, these mRNAs can initiate on intact ribosomes in the absence of initiation factors. Although these examples appear significantly simpler than factor-dependent initiation, it is unknown whether they reflect a primitive mechanism of initiation, or are later adaptations (section 16). After polypeptide release, bacterial ribosomes scan the mRNA to reinitiate translation at a downstream ORF without subunit dissociation. The ability of intact ribosomes to initiate on leaderless mRNAs, or reinitiate at a downstream ORF, raises the question of whether early ribosomes ever needed to dissociate into large and small subunits. Indeed the mobile subunit interface works in engineered ribosomes where the two subunits are judiciously stapled together (Aleksashin et al 2019). Conceivably the first ribosome, and any progenitor, comprised a single rRNA wherein the 3' end of the small subunit rRNA was continuous with the 5' end of the large subunit rRNA.

[What is the actual distance in the ribosome between SSU helix h45 and LSU H1?]

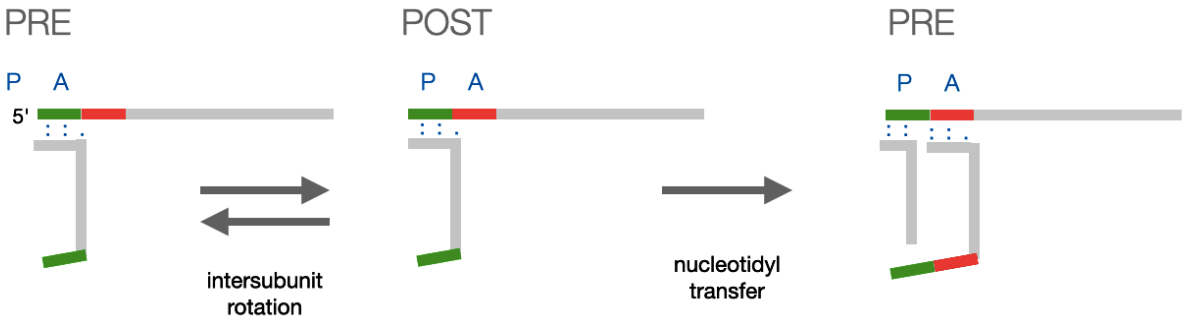


FIGURE 12-3. LEADERLESS INITIATION OF POLYNUCLEOTIDE DUPLICATION

Early duplisomes likely had few requirements for template recognition, and no preferred codons or dRNAs for initiation and termination. While the two duplication frames (even and odd) utilize alternate series of dRNAs, their polynucleotide products are identical. Perhaps the simplest model of initiation is threading the 5' end of the template into the entrance channel of the intact duplisome where the first codon selects a cognate duplicon-dRNA in the A-site (Figure 12-3). At peak warmth this codon-anticodon helix freely translocates to the P-site, but absent a nascent chain in the exit tunnel and a freed dRNA for cold closing and exit, there was no proximate source of Gibbs energy to ratchet this Brownian movement in one direction. On the next day, processive elongation commences when the second codon selects a cognate duplicon-dRNA for nucleotidyl transfer. In sections 15 and 16 we compare *primitive leaderless initiation* begun with a general polynucleotide elongator dRNA or polypeptide elongator tRNA in the A-site, to the modern forms of protein initiation begun with a dedicated initiator $tRNA_i^{(f)Met}$ in the P-site.

In the duplisome nucleotidyl transfer center, hydrolysis is followed by either chain condensation for polynucleotide elongation, or chain release for polynucleotide termination. Release was likely a default option in the continued absence of the A-site substrate for any of several reasons: (1) no intact codon due to a damaged template or just reaching the normal end of the template, (2) an intact codon but no cognate or near-cognate dRNA in the working set to read it, or (3) the dRNA selected and

accommodated has a damaged or absent duplicon. In the absence of nucleotidyl transfer, the nascent polynucleotide was released through its exit tunnel as the freed dRNA closed for its own exit. As each isoacceptor was needed on average once every 16 elongation cycles, or 32 nucleotides, it is unlikely that any dRNA could be spared for codon-directed polynucleotide termination, say by modifying it to prevent loading and transfer, or omitting it from the working set entirely.[15]

There were three special problems associated with duplication of the dRNAs themselves: First, their loading, decoding and transfer reactions required chemical definition of both 5' (duplicon) and 3' (anticodon) ends. Thus, if the 5' phosphate were lost from the free dRNA or duplicon-dRNA, the adaptor would no longer be capable of duplicon loading and nucleotidyl transfer, respectively. Damaged dRNAs were still useful as feedstock oligonucleotides using their 3' OH, but there was strong demand for their repair. Second, a working set of dRNAs must be complete and balanced, that is all 16 anticodons must be represented, and in comparable numbers.[16] When one or more isoacceptors went missing from this set, there was strong demand for some means of restoring them. Finally, some form of concerted evolution was needed to constrain the natural drift of dRNA sequences from one another, and to allow co-evolution with binding sites and catalytic centers of the duplisome and ribozyme P which interact with the entire isoacceptor set.

Interestingly, the problem of maintaining a complete and balanced isoacceptor set is connected to the problem of 3' end duplication. What happens at the end of the template likely depends on whether the final codon has all three, or just two nucleotides. In Figure 12-4 we show an odd-length template ending // GCAUG with AUdRNAUAU at the A-site just after nucleotidyl transfer. After translocation, the final G of the template occupies the first codon position in the A-site, but is too short itself to pair with a duplicon-dRNA for another elongation cycle. As a result, the duplicate ends // GCAU, that is, without the final G nucleotide of the template. Upon copying an

[15] In section 13 we explain the origin of tRNAs as genome guardians née parasites that affirmatively terminated duplication at STOP-codons that arise in nonself polynucleotides on average once every 64 elongation cycles, or 128 nucleotides.

[16] The minimal set of dRNAs with one of each isoacceptor, comprises nearly 600 nucleotides (576 nt = 16 x 36 nt).

odd-length template, this *end erosion* creates an even-length product, so that the polynucleotide community skews toward even lengths.

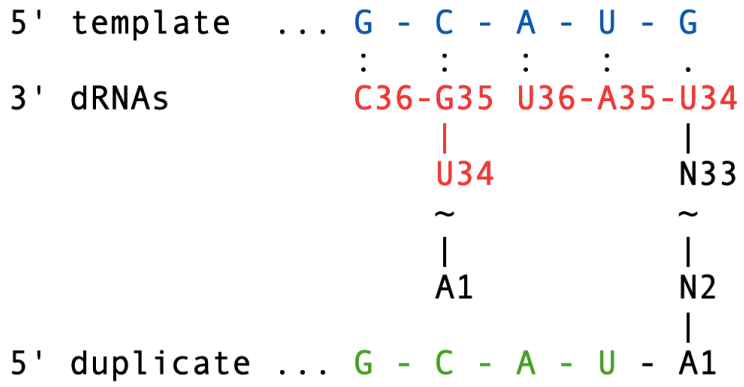


FIGURE 12-4. DUPLICATION OF ODD-LENGTH TEMPLATE WITH END EROSION

Unlike odd-length templates, where the final nucleotide is eroded in copying, we suggest the final two nucleotides of even-length templates were decoded, but with somewhat reduced fidelity (Figure 12-5). The result was a full-length duplicate without end erosion, but with occasional substitutions in either of the final two nucleotides. For RNAs generally, this error-prone copying of 3' ends, or *end wobble*, meant that the final two nucleotides were more variable than internal positions. This may have been important for primitive mechanisms of combination and repair. For dRNAs particularly, end wobble ensured any vacancies in the working set of isoacceptors were soon filled, and balance restored. It also allowed concerted evolution of the core stem and loop (nucleotides 1-34) so that the optimal distribution of dRNA length and sequence could adapt to changes in the duplisome, ribosome P, or physical environment.

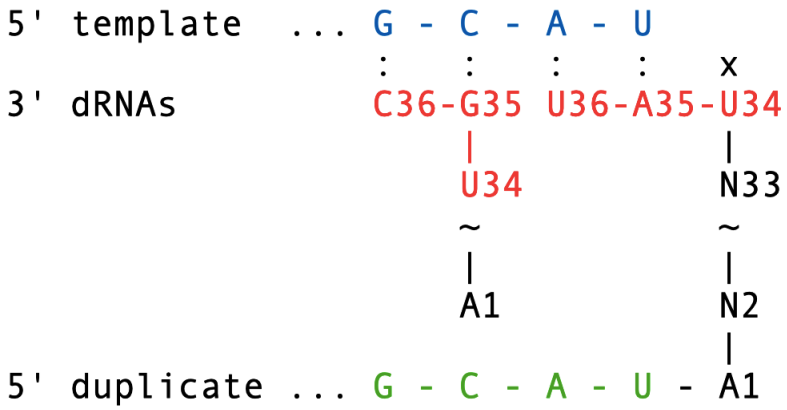


FIGURE 12-5. DUPLICATION OF EVEN-LENGTH TEMPLATE WITH END WOBBLE

In Figure 12-6 we depict the final intermediate in copying an unfolded dRNA template (without the 5' duplicon). In this example, the covalent intermediate is the tandem dRNAUCAdRNAUUG of length 36 + 36 = 72 nucleotides. After translocation to the P-site for hydrolysis (small blue arrow) and release, the product and by-product are a duplicate of the template dRNAUCA in the nascent polynucleotide exit tunnel, and the freed dRNAUUG that translocates to the E-site. In the absence of end wobble, these two dRNAs load complementary duplicons. For convenience, we have drawn both the template dRNA and the final dRNA in this nascent duplicate with the identical core sequence (nts 1-34). In general, these were drawn from pools of dRNAs that likely had some sequence variants at any time. In section 13 we propose that the first tRNA arose from one such tandem dRNA made as a copying intermediate.

FIGURE 12-6. TANDEM dRNA INTERMEDIATE

Having discovered the multilevel regulatory mechanisms of living things, *viz*. their lower-level search plans, the second, and arguably harder task of modern biology is to

explain the higher-level search plans that comprise the creative functions of heredity. Early critics of Darwinism noticed that the weak link in this theory of evolution was not natural selection, but natural variation. As Jacob Schurman put it, “Natural selection produces nothing; it only culls from what is already in existence. The survival of the fittest is an eliminative, not an originative, process,” or even more pithily, “The survival of the fittest, I repeat, does not explain the arrival of the fittest” (Schurman 1887 p78).

Darwin himself conjectured that natural selection acts on *small (quantitative) heritable variations* in all possible directions, whatever the physiological nature of these characters, and their constraints (Provine 1971). Proponents and opponents alike referred to these unplanned steps of exploration as *blind variation*. Plant and animal breeders reported occasional *large (qualitative) variations* known as monsters or sports. These were oftentimes lethal, sometimes heritable, and almost never advantageous in the natural environment. Optimists saw them as hopeful monsters, worthy gambles in an unplanned search for new functions, but pessimists just saw an unsustainable loss of vital functions.

The great mystery of heredity was the curious patterning of individual characters acquired and lost, amongst the vast number of species characters transmitted unchanged, in any line of descent. After Mendel’s theory of paired genetic elements, transmitted in one copy from each parent, and potentially expressed in their offspring, heritable variation was formulated in population genetics as a stochastic process of gene mutation and chromosome recombination (Provine 1971). In their models of variation and selection in sexual populations, the creative and conservative functions of heredity were strictly symmetrical, *viz*. fixing favorable alleles and haplotypes by hill-climbing was the obverse of eliminating unfavorable ones by purifying selection. Beyond the myopic, if not blind sources of variation, the creative functions of heredity were mainly to avoid the pratfalls of Muller’s ratchet and linkage disequilibirum (refs).

Before the theory of search complexity, no one could say whether blind variation and selection might find some proverbial needles in the haystack, or was a tacit appeal to miracles. Today we know that exploration, including the creative functions of heredity, entails greater search complexity than exploitation, including the conservative functions of heredity. Thus, in computer science and AI, various search problems

proved intractable as the worst-case performance of all known algorithms slows exponentially with natural measures of problem size (Cook, Karp). *Importantly, hard won solutions that are difficult to find, can be easy to verify and use*. This fundamental asymmetry of exploring and exploiting, or leading and following, can be seen in natural numbers whose prime factors are hard to discover by division, but easy to verify by multiplication, or Sudoku puzzles whose solutions are hard to find, but easy to verify once found.

In the evolution and function of biopolymers, problems of search complexity arise because sequences, 3D conformations, folding times, and activity landscapes all scale exponentially with polymer length. For local optimization of catalytic activity *in vitro*, or total fitness *in vivo*, by hill-climbing and purifying selection, a polynucleotide of length n need only be compared to 3n nearest neighbors, yet an exhaustive search, or global optimization has potentially 4^n sequences to compare. Put another way, the combined search space of two polymers of length m and n has 4^m x 4^n sequences, compared to just $4^m + 4^n$ sequences if each part contributes independently (aka *additively*) to the activity. Paired sequences that search their combined space in lock-step, *viz*. by *co-variation*, form an important special case of quasi-additive search complexity (section 3). Simple models of population genetics compared the fitness contribution of one allele to another, holding all other loci fixed. In more realistic models, the fitness of the whole was the sum of contributions from alleles of loci throughout the genome, corrected for pairwise or higher interactions, *viz*. dominance, epistasis, etc. In general, any additive improvements that do not dependent on combined search spaces are found sooner than later.

The conservative functions of heredity entail preservation and expression of knowledge, that is, faithful transmission of maps of past experience and plans of present behavior. These are cashed out in VGT as preservation and replication of the genome and its epigenomes. Focused on conservative functions, the models of steady-states and purifying selection in polymer evolution identified quantitative limits and perils, captured in the concepts of *genetic load*, *error catastrophe*, and *mutational meltdown*. In the theory of faithful vertical transmission as running in place, the main challenge were to replicate faster than degradation, to eliminate mistakes faster than

they were made, to protect and repair critical sequences, as well as to suppress selfish and useless elements. Even sex and recombination were to be explained as conservation and repair, not exploration and discovery (Maynard-Smith).

The creative functions of heredity acquire new knowledge through some combination of in-house discovery and HGT. These are not cashed out, as Darwin had supposed, by exploring small variations in all possible directions (Provine 1971). Indeed, the most common spontaneous mutations of nucleic acids, viz. single nucleotide substitutions and small indels, were so greatly oversampled that error-free repair pathways evolved to reduce their frequency (section 18). At the same time, sundry pathways evolved to catalyze more promising, yet otherwise improbable mutations. Just as it took time to realized that many details of lower-level search plans are not adaptive for survival and reproduction, it has taken time to realize the many details of these higher-level search plans are adaptive for genome evolution.

Discovered piecemeal by molecular biologists, the substrates and catalysts of creative heredity first seen as a kludge and clutter of genomic debris, are in fact simple realizations of planned trials in polymer life. These began as simple mechanisms for admixture and assortment of polynucleotide communities, as well as combination and fragmentation of these polymers. Today, throughout cellular and viral life, sundry ribozymes and enzymes rearrange existing genes, recombine related sequences, combine unrelated sequences, create duplication or deletions, and transpose or copy sequences to new chromosome locations altogether. In prokarya, these processes of natural variation for evolution range from regular sampling of mobile operons from the pangenome to unchecked movements of extremely selfish elements. In eukarya, they range from regular pairing and recombination in euchromatin to RNA-mediated duplication and shuffling in heterochromatin, and genuine infection by MGEs from outside the species (Hutter et al 2000).

Unlike meiotic recombination, which requires extended local homology as well as global chromosome synapsis, duplisomes were extremely short-sighted, matching nothing longer than anticodon triplets. Beyond rethreading the 5’ ends of broken templates, they likely played a role in creative combination of sequences by threading the 5’ ends of unrelated second templates. Thus, upon reaching the end of one

template, the intact duplisome might thread the 5’ end of new template analogous to the reparative rethreading of a broken template (cf. Figure 12-2). The finished product combines all of the first template sequence and all of the second sequence, perhaps with a small indel at their junction. There was another likely way of combining sequences when the duplisome stalled midway along a template without reaching its end. To wit, the duplisome might disassemble, allowing the large subunit with a polynucleotidyl-dRNA in its exit tunnel to reassemble with an empty small subunit to resume duplication on a fresh template. Between these two mechanisms, the final product combines all, or just the initial part of the first template sequence, with all of a second sequence.[17]

Besides break repair or novel product combinations from duplisome rethreading or reassembly during copying, there were likely older mechanisms for repair and combination of templates themselves, not just their copies. We suggest that the common reactive ends of polynucleotides in duplisome life were the 5’ phosphate and 3’ OH. Where these were missing, damaged or blocked, two primordial ribozyme, a 5’-*exonuclease* and a *polynucleotide phosphoryl ligase*, could restore them, discarding

---

[17] Here we notice a possible mechanism of polynucleotide initiation on a first template, or continuation on a second template, analogous to the rescue of bacterial ribosomes stalled at the ends of non-stop mRNAs (ref). To wit, a donor-template RNA with a donor-like 5’ sequence and template-like 3’ sequence might enter the A-site of an intact duplisome with no codon to be read. This could occur in two ways: (1) awaiting initiation when their was no template at all, nor polynucleotide-dRNA in the P-site, or (2) with a polynucleotide-dRNA in the P-site, awaiting termination at the end of the template, but no codon at all, or perhaps just one nucleotide, in the A-site. After nucleotidyl transfer to its dRNA-like 5’ end and translocation, RNA elongation continued on its template-like 3’ end. The product is …
We conjecture that trans template swapping on intact duplisomes using a donor-template RNA (dtRNA) as the second template gave rise to bacterial trans-translation using a transfer-messenger RNA (tmRNA) to rescue ribosomes stalled at the end of nonstop mRNAs. Whereas tmRNA rescues ribosomes stalled on damaged mRNA, the original dtRNA had a creative role in combining entire sequences end-to-end, not an eliminative role in quality control of bad templates.
… abductive, invasive, parasitic …
IDEA what precisely is the structure of the dtRNA? duplicon-dRNA-anticodon-template
[IDEA]?can ribozyme P load a donor-template RNA … sequences trailing the anticodon swung out of the way … in section 9 we proposed a scheme of dRNA loading ….
[IDEA] when used to initiate, only the template is duplicated … distinction between full-length ribogene and its working products ….
Despite some mechanistic analogy between donor-template RNAs with a regulatory role in polynucleotide replication and combination, and transfer-messenger RNAs with … rescue ribosomes with end stall problem …. ?is there an end stalled problem at all; idea tmRNA evolved from dtRNA for recombination not ribosome rescue?
Uncertain that there was any need for a dedicated mechanism of initiation or rethreading, much less this peculiar donor-template RNA, we leave the question for interested readers.

one (or more) terminal nucleotides. The exoribonuclease, ancestral to ribozyme P, hydrolyzed the 5’ nucleotide to prepare the 5’ phosphate polynucleotide, solving one grave threat to the elongation cycle of duplisome life, dRNA damage with loss of the 5’ phosphate. The polynucleotide ligase, ancestral to the duplisome nucleotidyl transfer center, hydrolyzed the 3’ terminal nucleotide to push condensation of its 3’ OH with the 5’ phosphate polynucleotide (Figure 12-7). Together these two workhorses of early RNA life prepared and combined ends of any two polynucleotides with the minimal sacrifice of one nucleotide from each to ensure undamaged ends and drive condensation. Before the evolution of processive copying, these ribozymes allowed quick and dirty repair of broken strands, as well as combination of unrelated RNAs. Evolved for terminal joins, the exit tunnel of the ligase P-region admitted RNA substrates with indefinite 5’ extensions, while the entrance tunnel of its A-region admitted RNA substrates with indefinite 3’ extensions.

[QUERY] repair an unloaded dRNA damaged at 5’ phosphate by removing N1? general ligase can lengthen, ribozyme P can trim back to N1 or to N-2? no need for ribozyme P to charge at N2, it remains a specialist for charging at N1

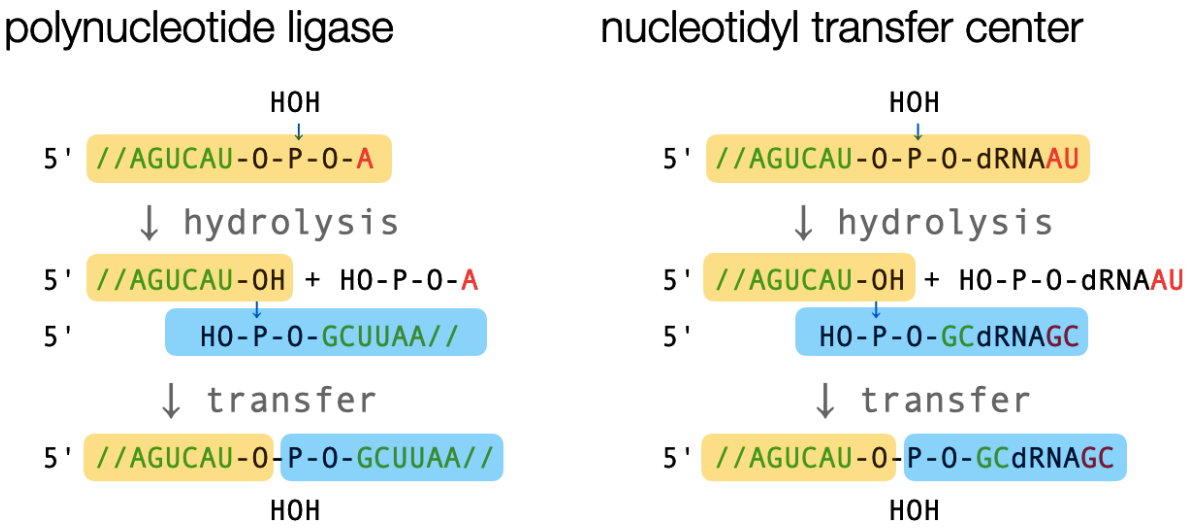


FIGURE 12-7. RNA LIGASE ORIGIN OF DUPLISOME NUCLEOTIDYL TRANSFER CENTER

Various covalent modifications might protect polynucleotide ends from damage, prevent their copying or combination, target them to polymer complexes or other

compartments, and otherwise regulate their activity (Orgel 1989). Anticipating aminoacyl tRNA synthetase ribozymes (section 14), we propose that a family of polynucleotide acyl ligase ribozymes, homologs of the polynucleotide phosphoryl ligase, esterified common metabolites to the 3' end of polynucleotides. These ribozymes, we suggest, accepted a variety of carboxylic acids, certainly amino acids, and possibly peptides, glucoronic acid, fatty acids, or other ligands, according to their affinity and availability. They likely favored ligand diversity over specificity, so that 3' tagged RNAs were not so much targeted to complexes and compartments, as selected by their affinity for them. Like the phosphoryl ligase, these acyl ligases sacrificed one or more terminal nucleotides to energize a sequential hydrolysis-condensation reaction between the penultimate O2'/O3' of the polynucleotide and carboxyl carbon of the metabolite.

Ancient ribozymes of site-specific recombination, self-splicing introns likely arose in the RNA world, and have continued to flourish in DNA life using reverse transcriptase and homing endonuclease enzymes (ref). Viewed as mobile genetic elements, self-splicing introns insert themselves into host RNAs by reverse splicing at target sites defined by pairing with intron guide sequences. There they need contribute little to enjoy whatever evolutionary prospects the host ribogene confers, so long as they can remove themselves precisely by splicing, restoring an uninterrupted host sequence, and making themselves available for reinsertion there or elsewhere in the genome. In abortive cycles of reverse-then-forward splicing at the same site, introns did no great harm, nor much good, mobile genetic elements with really nowhere to go.

Once there were two copies of an element at the same site in homologous RNAs, or at distinct sites in non-homologous RNAs, self-splicing introns could mediate site-specific recombination, a productive form of *trans*-splicing (Zaug & Cech 1985). Likely functions of self-splicing introns in RNA life range from conservative to creative including: (1) occasional exchange between homologous ribogenes to avoid mutational meltdown from Muller's ratchet; (2) concerted evolution of ribogene families that co-evolve together in one part and specialize apart in another; (3) accelerated evolution of ribogene families through recombination of parts; (4) transduction of flanking sequences from one insertion site to another by shifting splice-donor or splice-

acceptor sites between forward and reverse splicing; and (5) introduction of elements at non-homologous sites to promote novel sequence combinations. Interestingly, there are not one but two, ancient families of self-splicing introns. These two groups may have evolved in separate lineages, or one after the other within the same RNA community. In either case, there was likely some creative advantage to having both, perhaps reflecting differences in insertion-site recognition, or flanking-site transduction.

[compare GISSI intron tags for regulation of RNA duplication with Maizel & Weiner genome tags for RNA replication]

donor-trailer RNAs = dRNA intron template

leader-donor RNAs = template intron dRNA

13. RNA termination | constitutive tRNAs

Beyond exhuming prebiotic and ribozymatic processes of RNA copying, the origin of polypeptide translation is the greatest challenge to tracing polymer life back to an RNA world. Our canonical genetic code maps codons to amino acids, and by iteration, mRNAs to polypeptides. This map is many-to-one, that is any codon assigns (at most) one amino acid, while any amino acid may have six (L R S), four (A G P T V), three (I), two (C D E F H K N Q Y), one (M W), or no codons (non-proteinogenic amino acids) at all. The function is partially defined as three *stop* codons (UAA UAG UGA) map to no amino acid.[18] The brilliant mechanistic insight and experimental breakthrough was that this amino acid code is the product of two simpler relations involving the family of small adaptor RNAs known today as transfer RNAs (Crick 1955/1958; Crick et al 1957; Hoagland 1959; Zamecnik 1960; Fry 2022). Each tRNA, complexed with its cognate aminoacyl-tRNA synthetase (aaRS) enzyme, is charged at its 3' end with an amino acid, and then, complexed with the ribosomal decoding center, is matched via its anti-codon to an mRNA codon. These charging and matching relations together specify the amino acid code.

Untangling the origins of protein coding, we examine polypeptide translation *before any support from coded proteins* in section 15. At this breakout stage of protein life, all processes that today require coded proteins, e.g., ribosome (tRNA) biogenesis, rRNA (tRNA) modification, amino acid biosynthesis, tRNA charging, polypeptide initiation, elongation, and termination, and protein secretion, were catalyzed by (possibly extinct) ribozymes and random polypeptides, happened spontaneously, or did not happen at all. Under this assumption, we compare two scenarios for the origins of translation: First, ancestors of rRNAs and tRNAs had no definite functions in the RNA world, or second, these molecules functioned in RNA duplication more or less as sketched in sections 8-12. The first hypothesis places no constraints on primitive ribosomes and tRNAs beyond parsimony with modern ones, while the second hypothesis requires parsimony with duplisomes and dRNAs, their conjectured progenitors, as well.

---

[18] There are curious exceptions to these canonical stop codons in mitochondrial translation, as well as recoding of stops codons for selenocysteine and pyrrolysine (ref).

Along any evolutionary path from polynucleotide duplication to polypeptide translation there were key changes in (1) the structure and charging of adaptor RNAs, (2) the mechanism of decoding, (3) the chemistry of polymer transfer, (4) the process of translocation, and (5) the energetics of elongation. Duplisomes meanwhile remained the principal mechanism of RNA copying up until the invention of RdRP enzymes. If so, remodeling them for protein translation was akin refitting a ship at open sea, not in the shipyard. There were, we suggest, two key points of regulation, (6) one focused on the adaptor RNAs, and (7) another on the ribosome née duplisome, for orthogonal duplication and translation without costly redundancy or interference. We identify RNA changes associated with each step, and suggest an adaptive path from polynucleotide duplication to polypeptide translation. In apologetics of our just-so story of ribosome evolution, we propose explanations for hitherto curious features of cellular life, venture testable predictions, and raise unexpected questions.

Here we posit two intermediate stages along the path from polynucleotide duplication to polypeptide translation that make the breakout problem tractable, and our solution unique. We propose that *tRNAs and their charging ribozymes were invented to regulate polynucleotide synthesis, not direct polypeptide synthesis.* To wit, before any iterative translation of polypeptides, the duplisome acquired the means of codon-directed termination of nascent polynucleotides using plain tRNAs, and then the means of regulating these tRNAs by acylating specific ligands to their 3’ end. Thus, polynucleotide termination evolved in three stages: (1) An era of *default release* whenever the polynucleotide-dRNA underwent hydrolysis without any duplicon-dRNA accommodated in the A-site for transfer (section 12). This might occur if the duplisome reached the end of one template without threading another, the template had too much damage to skip over, no cognate or near cognate duplicon-dRNA was available, or its duplicon was damaged or missing. (2) An era of *codon-directed termination* using constitutive terminator tRNAs (this section). (3) An era of *codon-directed termination* using conditional terminator tRNAs (section 14).

Before discussing the origin of tRNAs as polynucleotide chain terminators, we introduce some regulatory principles of polymer life and the emergent distinction between gene replication and expression. Francis Crick contrasted genes (hereditary

polymers) and their products (final polymers) as the central dogma of polymer life. Here the conservative functions of heredity comprise processes that either sequester or copy full-length genes, while gene expression comprises processes that modify, cleave, truncate, or translate genes into active products, often through one or more irreversible intermediates. By LUCA, DNA replication was the example *par excellence* of conservative heredity, while RNA transcription and protein translation were examples of gene expression.

Nuclease ribozymes and bistable riboswitches likely were important regulatory mechanisms of RNA life. Actions of nucleases in creative heredity and gene expression are effectively irreversible. Whether used by a parasite to attack, or by their host to parry, RNases could selectively cleave non-clique ribogenes so these no longer function as full-length templates. In the regulation of gene expression, RNases could activate ribozymes by removing an inhibitory sequence, or inactivate them by cleaving their catalytic center. Unlike nucleases, riboswitches regulate reversible as well as irreversible events. Modern riboswitches have aptamer sequences that monitor the presence of fluoride or magnesium ions, amino acids, nucleotide derivatives, or key proteins, as well as guide sequences that recognize cognate RNAs (RNPs) to evaluate their conformation or physiological status. Primitive riboswitches likely promoted (sequestered) ribogenes for copying, or activated (inhibited) ribozymes for catalysis, in response to the presence or absence of their cognate ligands, or pairing to their guide sequences.

In cellular life, assorted cap and tail modifications (1) protect the ends of RNAs, and (2) regulate their functions. An even greater variety of covalent modifications are found at internal positions that (3) stabilize the folded RNA, and (4) protect critical sequences from physical insult, biotic nucleases, or antibiotics. In some cases, these modifications (5) balance RNA interactions with multiple partners, such as anticodon pairing with every cognate codon while rejecting all others, (6) provide checkpoints in biogenesis for quality control, or (7) regulate the activity of RNPs, toggling them ON or OFF, or from some original (default) function to a derived function.

Covalent modifications of rRNAs and tRNAs are concentrated at important sites of substrate interaction and catalysis. Many of these modifications trace back to LUCA,

and some perhaps to protein life, or even earlier. Enzymes for rRNA modification use families of ancient guide RNAs (sRNAs archaea, snoRNAs eukarya) for cleavage, 2'-O-methylation, or pseudouridinylation at specific nucleotides. These guide RNAs may predate their enzymes, working with lost ribozymes or spontaneous chemistry.

One of the great insights of evolutionary biology is that virtually any process of cells and multicellular organisms can be co-opted by infectious agents, *viz*. MGEs or organized parasites such as encapsidated viruses, cells, and multicellular organisms (Burt & Trivers 2006; Werren 2011; Burroughs & Aravind 2016; Koonin et al 2020b). Thus, many of the regulatory mechanisms found in any host have been honed through back-and-forth arms races of infection and immunity, and likely many were first introduced by parasites in forms of HGT. The primary object of any infectious agent is to overcome any host defenses, and divert host resources to parasite reproduction. At this level of selection, any cost or benefit to an individual host is of little importance. But at a higher level of selection, the structure and preservation of the host lineage is of great importance to co-evolution of the parasite lineage as captured in the theory of density-dependent selection (Reznick et al 2002; Bertram & Masel 2019).

Perhaps the simplest way to understand how the availability of hosts seen as environmental affordances shapes the evolution of parasite features is the theory of r- versus K-selection (MacArthur & Wilson 1967; Pianka 1970). The parameters 'r' and 'K' are borrowed from the logistic equation of ecology … where r is the maximal intrinsic rate of natural increase given an excess of hosts, and K is the carrying capacity given an excess of parasites. Under low ratio of parasites to hosts, or r-selection, … small parasites, rapid maturation, short lifespan, large broods, semelparity …. Under high ratio of parasites to hosts, of K-selection, large parasites, slow maturation, long lifespan, small broods, iteroparity. Over time, a parasite that successfully invades a naive population of hosts changes the parasite/host ratio from r-selection to K-selection. For viruses, for example, the shift from lytic rampages with destruction of hosts, to lysogeny or other forms of dormancy and reproduction when the host individual itself is stressed and unlikely to survive.

Many simple mechanisms of parasitism have two principal components, *viz*. restriction/modification or toxin/antitoxin systems in bacteria. Thus, any *toxin*, viz.

negative regulator of an essential gene or product, that spreads by HGT, can potentially partner with an antitoxin to form a two-component protection racket. For example, guide sequences for RNA modification that protect critical sequences against spontaneous damage, may have first spread parasitically among RNA communities as part of restriction-modification systems that promote the same damage as biotic insults. Though these spread at low ratio of parasites/hosts as you-must-pay-for-protection rackets, wherein the host is better off without the parasite, at higher ratio of parasites/hosts these confer immunity, an insurance policy against the ravages of similar invaders. Below we explain the invasion and domestication of terminator tRNAs as a one-component system that infects new populations at low ratio parasites/hosts and co-evolves with an entire ribocommunity of duplisomes and ribogenes ….

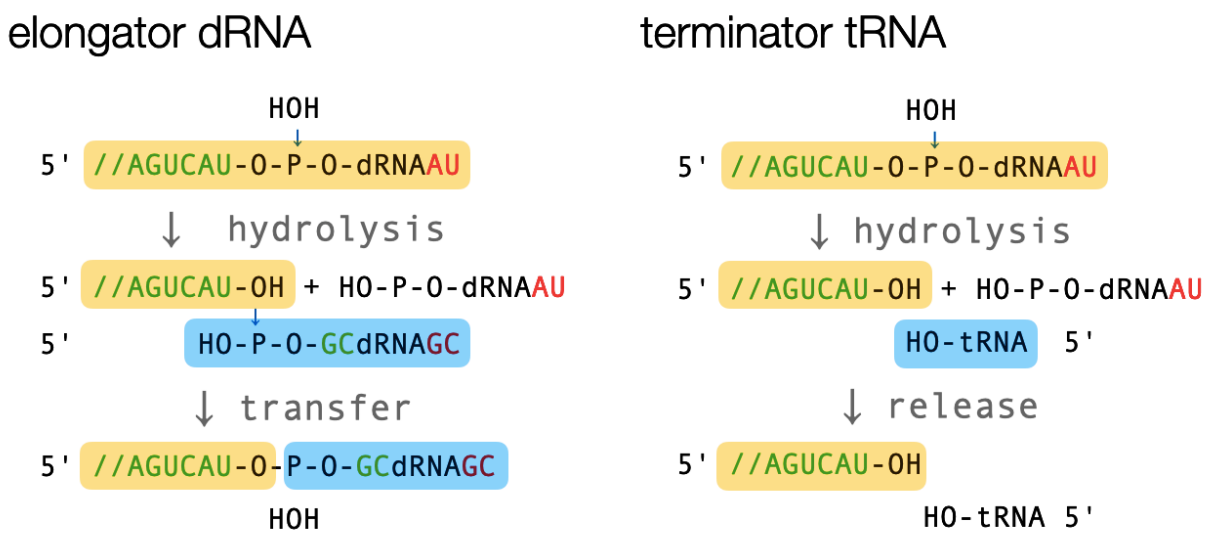


FIGURE 13-1. CODON-DIRECTED POLYNUCLEOTIDE ELONGATION VERSUS TERMINATION

Our proposed reactions of the duplisome nucleotidyl transfer center in codon-directed polynucleotide elongation and termination are shown in Figure 13-1. The key difference is whether the A-site substrate is an elongator dRNA or a terminator tRNA. All 16 dRNAs are needed too often for polynucleotide elongation to spare even one of them as a terminator. To be useful for termination, a tRNA must be highly selective, reading rare STOP codons without misreading common codons to cause premature termination. We propose that terminator tRNAs were matched by Watson-Crick pairing

in their third codon position, not just the first two positions. In a random template with equal frequencies of all four nucleobases, the frequency of inframe STOP codons would be 1/64 x 1/2, so an open duplication frame (ODF) would be 128 nucleotides on average. For longer polynucleotides, problematic STOP-codons could be eliminated by substitution of any of the three codon positions, or any odd-length indel upstream to shift the reading frame.

We suggest that primitive tRNAs arose as *killer genes* that caused premature termination of susceptible RNAs, thus freeing their duplisomes to copy fresh templates, including the killer tRNA. In a simple steady-state picture of duplisome life, the community of ribogenes ranges in length from short (dRNA) to long (ribozyme P, rRNA), all with fair access to duplisomes for initiation of copying. With one initiation event per template per 30 days, we expect about one duplisome on a 60 nucleotide template, and five duplisomes on a 300 nucleotide template. When a killer tRNA infects this community, it causes premature terminations, affecting especially longer ribogenes, that free up duplisomes for new initiations, and shifts the overall size distribution of nascent RNAs toward shorter products. In a nutshell, the immediate selfish advantage of CODON-directed polynucleotide termination to killer tRNA is freeing up duplisomes to initiate on new templates, including itself. This is a simple example of the parasite diverting host resources to favor their own reproduction. There is nothing particular about the initiation events that favors the killer tRNA over a neutral freeloader, nor indeed a useful host ribogene, but the termination events disfavor longer ribogenes because their likelihood of harboring the cognate RIBOSTOP-codon increases with length.

. When the parasitic tRNA infects an RNA community, it can terminate polynucleotide duplication at any cognate triplet codon. Thus, it proffers a clique of favored ribogenes, *viz*. those without critical RIBOSTOP-codons, at the expense of all those with them. In a simple form of variation and selection, disfavored ribogenes may still contribute a copy to this favored clique through an odd-shift in reading frame upsteam of the STOP codon. As this junior partner harbors an off-frame STOP-codon, its own copies are vulnerable to termination from upstream frameshifts during duplication. Hence, after purging all overt STOPs, there was continued, albeit weaker selection to

remove cryptic STOPs as well. If a community survives the invading tRNA, enlarging its clique of RIBOSTOP-free genes to encompass all essential functions, it not only becomes resistant to the parasite, but is now protected from foreign ribogenes by this same guardian née killer tRNA.

The rampages of killer tRNAs in populations of host communities, and their transformation from parasites to guardians, were more or less immediate. Meanwhile, emergent roles of terminator tRNAs in conditional regulation of gene expression required a novel mechanism for reversible activation or inhibition of tRNAs (section 14).

FIGURE 13-2. NEGATIVE SELF-REGULATION OF tRNA RIBOGENE

As parasites, the onus was on any killer tRNA to compete well-enough for decoding against cognate dRNAs to create a favored clique in the intra-community competition of ribogenes. Thus, at any template sequence of the form $_1$NNNNNN$_6$, the cognate tRNA terminator for RIBOSTOP-codon $_4$NNN$_6$ preempts the cognate elongator dRNA for ribosense codon $_3$NNN$_5$. As the favored clique enlarged to encompass the entire community of essential ribogenes, not least those of processive RNA copying itself, a tentative partnership of duplisome and tRNA emerged. Now inter-community selection favored the duplisome and tRNA that worked together to prevent premature off-STOP polynucleotide termination. This selected for improvements in both the decoding center and the tRNA anticodon arm, as well as limits on tRNA activity. One simple

mechanism of negative self-regulation of tRNA copy number is a cognate STOP-codon (xxx) positioned early within the tRNA ribogene itself (Figure 13-2).

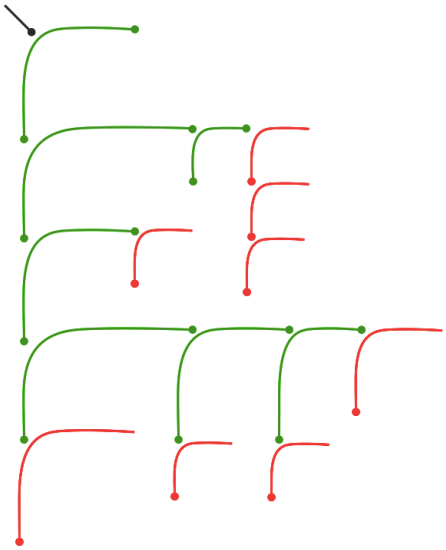


FIGURE 13-3. PEDIGREE OF PREMATURE TERMINATION OF RIBOGENE DUPLICATION

In cellular life, there is clear separation of gene replication and expression, and in multicellular life, close coordination of genome replication, as well as patterning of gene expression, amongst cells. In RNA life, the distinction between gene replication and expression was less sharp, and copying less coordinated. Figure 13-3 depicts the pedigree of a ribogene after 15 bouts of polynucleotide duplication. Seven green bouts escape termination to make full-length copies, while eight red bouts undergo STOP codon-directed termination to discard nascent 5' fragments. Thus, the sixteen terminal branches of the pedigree show eight full-length ribogenes, including the original template, and eight discarded 5' fragments. This simple pedigree does not show the relative progress, and possible coordination of duplication along different branches. Finally, for two or more duplisomes on the same template, at any elongation cycle, one duplisome could terminate while the others continue elongation. We conjecture that subunit disassembly *tout court* allowed terminated duplisomes to redeploy elsewhere without interfering with flanking ones that continue elongation.

Here we conjecture that the first tRNA arose from the tandem dRNAs that form as final intermediates during the duplication of dRNAs. Sequencing yeast $tRNA^{Ala}$ (IGC),

Robert Holley and colleagues proposed the four-way junction, or *cloverleaf model* of tRNA secondary structure (Figure 13-4). One alternative was an imperfect, double hairpin that lengthens the D- and T-arms at expense of the acceptor and anticodon stems. In various *tandem hairpin models* of tRNA evolution, fusion of two hairpins gave rise to the primitive tRNA (Eigen & Winkler-Oswatitsch 1981a,b; Di Giulio 1992, 2004; Dick & Schamel 1995; Schimmel & Ribas de Pouplana 1995; Nagaswamy & Fox 2003; Widmann et al 2005). Besides the lengths of hairpin stems and loops, these models vary whether the stems were nearly perfect, weakly paired, or bulged, and whether the two hairpins were nearly identical, highly diverged, or entirely unrelated. More importantly, the models vary in their proposed function of the ancestral hairpins in RNA replication or polypeptide synthesis, in when, how and why hairpin fusion occurred, and in when, how and why the family of tRNAs radiated. We leave to interested readers some models of the origin of cloverleaf tRNAs from simple repeats other than tandem hairpins (see Agmon 2022).

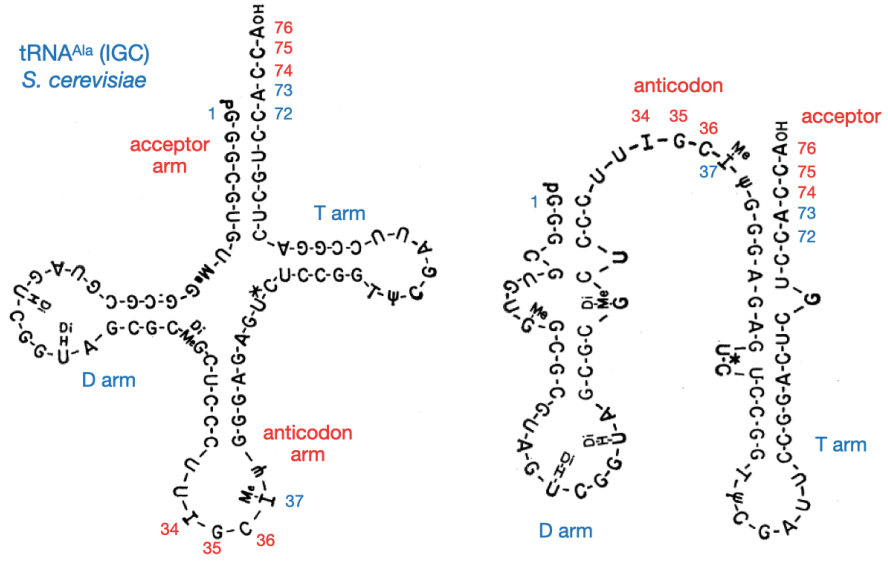


FIGURE 13-4. CLOVERLEAF & DOUBLE HAIRPIN tRNA STRUCTURES (HOLLEY ET AL 1965)

We illustrate all and sundry tandem hairpin models of tRNA origin with just one example: Massimo Di Giulio conjectured that two identical RNA hairpins gave rise to the D-arm, and the T-arm, respectively, with both halves contributing equally to the acceptor and anticodon stems (Di Giulio 1992, 2004). For sake of discussion, he proposed an ancestral hairpin of 38 nucleotides, comprising a perfect stem of 12

basepairs, a loop of 10 nucleotides, and a 3' trailer of 4 nucleotides (Figure 13-5). In his model, the hairpins were actually a family of sequences each with a presumptive anticodon triplet immediately followed by a common 3' trailer $_{35}$DCCA$_{38}$ homologous to the 3' trailer $_{73}$DCCA$_{76}$ of mature tRNAs. Although he was silent about how (and how many times) hairpin fusion occurred, Di Giulio inferred that the tRNA variable arm arose from the 3' hairpin either during or after the fusion event, and that the amino acid identifier (ID) sequence of the tRNA acceptor arm arose from the four nucleotides of anticodon (ANT) and discriminant (D).

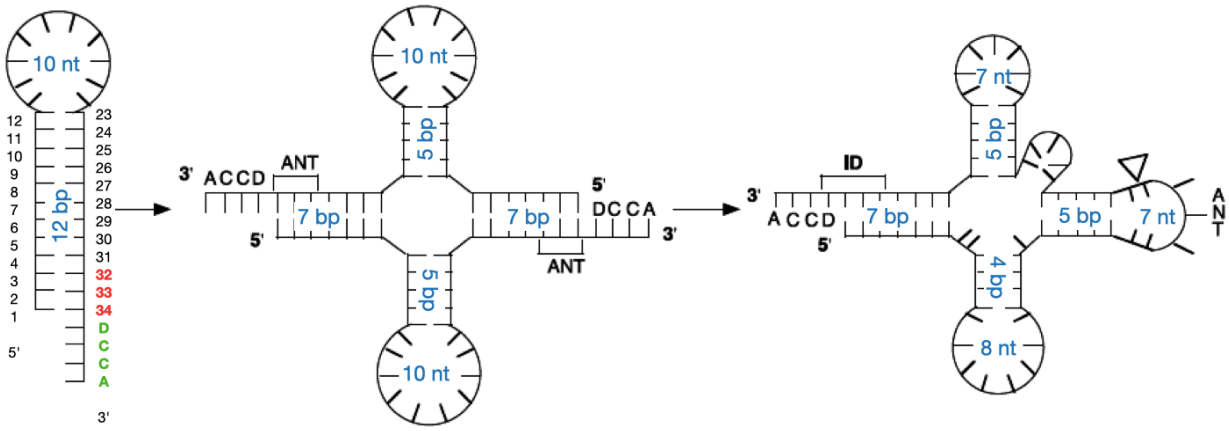


FIGURE 13-5. TANDEM HAIRPIN ORIGIN OF tRNA (AFTER DI GIULIO 1992, 2004)

Several aspects of the biophysics and biochemistry of modern tRNAs have been taken as evidence for or against a tandem hairpin origin. Double hairpins have not been reported as major folding intermediates in tRNA biogenesis, nor as major unfolding intermediates of mature tRNAs *in vitro.* Under various salt conditions, after the elbow separates, D-stems are generally the first, and T-stems the last to melt (see Privalov 2012). The first suggestion of double hairpins as an actual intermediate *in vivo* came from discovery of *copia* retrotransposons in *Drosophila melanogaster* that use the 5' fragment of tRNA^{iMet} (CAU), not the 3' end of the intact tRNA, to prime reverse transcription of their minus strand (Kikuchi et al 1990). When tested *in vitro*, initiator pre-tRNA^{iMet} and at least two elongator pre-tRNAs from *Drosophila* can be cleaved *in vitro* after the anticodon by bacterial RNase P (Hori et al 2000; Tanaka & Kikuchi 2001). This internal cleavage, or *hyperprocessing*, gives biochemical evidence that these pre-

tRNAs assume a double hairpin as a minor conformation *in vitro* (Figure 13-6). *In vivo* cleavage of $tRNA^{iMet}$ at nucleotides N39↓N40 likely uses the retrotransposon-coded RNase H, not the household RNase P of pre-tRNA maturation. More recently, a large literature has emerged on the processing of pre-tRNAs and mature tRNAs into a variety of fragments (tRFs) with possible regulatory functions (ref).

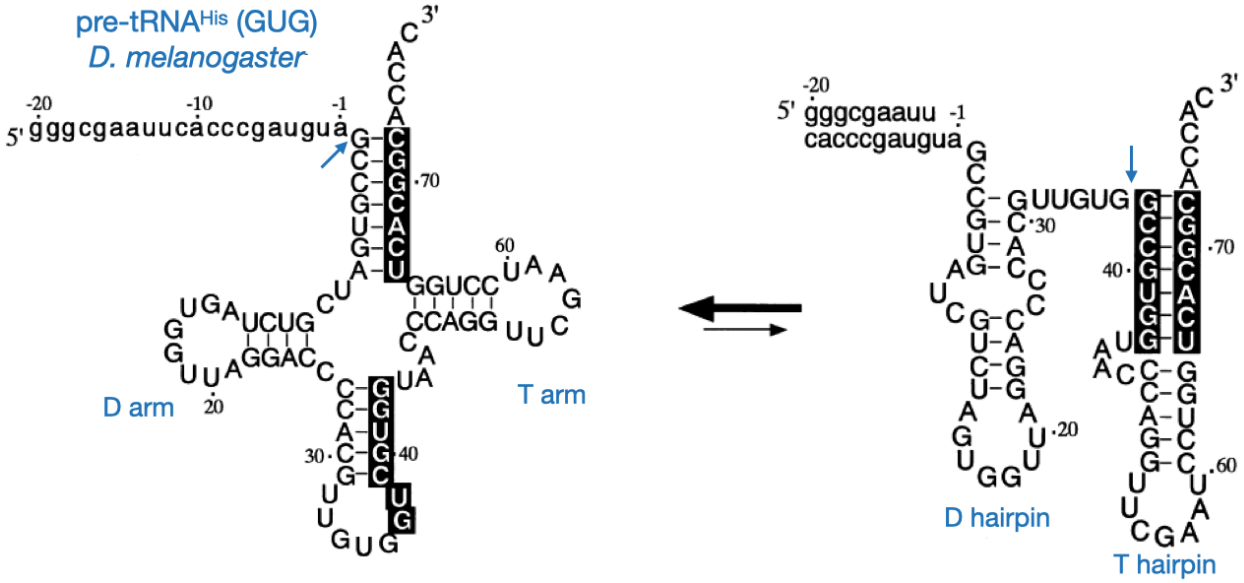


FIGURE 13-6. INTERNAL CLEAVAGE OF tRNA BY E. COLI RNAse P RNA (TANAKA & KIKUCHI 2001)

Our proposed origin of tRNAs from tandem dRNAs differs from other tandem hairpin models in three ways: First, whereas those theories have either (1) no particular function in mind for the ancestral hairpin, (2) a genome tag function for recognition of RNA templates by the replicase ribozyme or primer, (3) an acceptor function in random polypeptide synthesis, or (4) an anticodon function in primitive decoding, we propose that (5) dRNAs functioned in duplisome-mediated RNA copying. Second, the functional requirements of duplicon loading, template decoding, nucleotidyl transfer, and translocation, constrained the dRNA structure, so our primitive tRNA must be parsimonious with this progenitor hairpin, not just modern tRNAs. Third, whereas most tandem hairpin models are agnostic about how and when hairpin joining occurred, we suggest tandem dRNAs were ordinary intermediates in the duplication of dRNAs when the original tRNA was exapted for STOP codon-directed polynucleotide termination.

The tandem dRNAs of the final intermediate are drawn from the common pool of dRNAs, the 5’ dRNA as the duplicated template and the 3’ dRNA as the final substrate. The only necessary sequence relation between them is that the anticodon of the 5’ dRNA is *one-and-the-same dinucleotide* as the duplicon of the 3’ dRNA. If both dRNAs happen to be identical (excepting their complementary anticodons), the anticodon and acceptor stems of their cloverleaf correspond in length and sequence to the distal stem of their dRNA progenitor (Figure 13-7). More generally, the two halves of the cloverleaf would be similar, but not identical in sequence, and might pair together more perfectly, or less perfectly, than either progenitor hairpin. Thus, all tandem dRNAs formed as duplication intermediates could likely fold as double hairpins, but only rare combinations could also fold as stable cloverleafs. This is contrary to the situation in modern tRNAs where all sequences fold as stable cloverleafs, but only rare tRNAs can also fold as double hairpins.

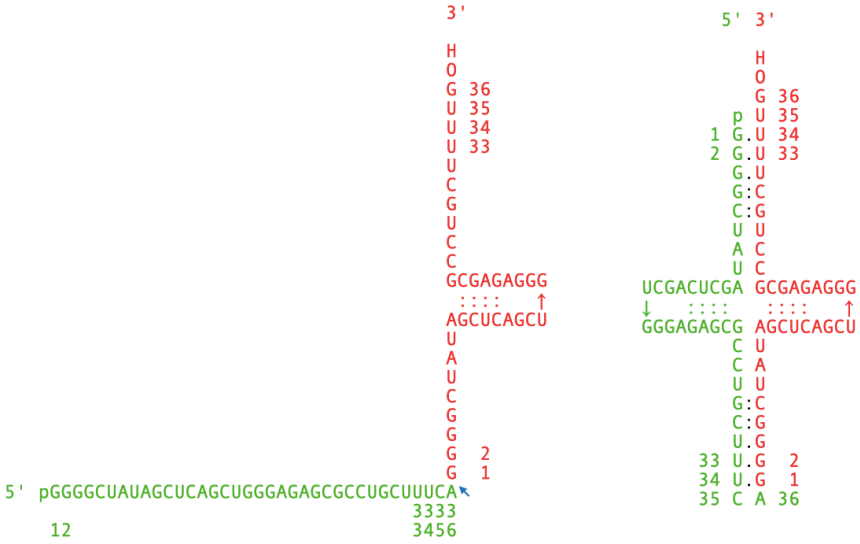


FIGURE 13-7. TANDEM dRNA INTERMEDIATE FROM FIGURE 12-6 FOLDED AS A CLOVERLEAF

In modern tRNAs, tertiary interactions between the modified D- and T-loops at the elbow stabilize the L-shaped fold, or orthogonal arrangement of their stems (Zhang & Ferre-D’Amare 2016; Roovers et al 2021). Here we propose that a rudimentary elbow structure was the *sine qua non* of the primitive tRNA. Cloverleaf secondary structure was at best a minor conformation for some tandem dRNA intermediates (Figure 13-7). Released prematurely from the duplisome, *viz*. before the final hydrolysis (small blue

arrow) in the nucleotidyl transfer center, these intermediates remained substrates for hyperprocessing by ribozyme P in their major, double hairpin conformation. By stabilizing the cloverleaf secondary structure, the elbow suppressed the double hairpin alternative, preventing internal cleavage of the primitive tRNA by ribozyme P, the usual fate of tandem dRNA intermediates released intact from the duplisome.

We suggest that the tRNA cloverleaf arose from a tandem dRNA intermediate whose tandem hairpins opened up to pair together from head to tail except for their loops and proximal stems. Evolution of its crucial elbow began with specialization of the loop sequences. Today the conserved D-loop (née 5' dRNA) and T-loop (née 3' dRNA) sequences are A14 R15 Y16 Y17 G18 G19 N20 R21, and U54 U55 C56 R57 A58 N59 Y60, respectively (Roovers et al 2021). Unmodified tRNA forms the elbow structure, presenting the flat faces of its central Watson-Crick basepair G19 : C56 to the solution (Bryne et al 2010). Giving the D- and T-loops their respective monikers, modifications of nucleobases U16 U17 to *dihydrouracils*, and U54 U55 to *thymine* and *pseudouracil*, make the elbow more rigid and heat stable.

Unlike the critical functions of their 3' ends in amino acid charging and peptidyl transfer, no function of the 5' end of tRNAs seems to warrant precise cleavage by an ancient ribozyme. The duplisome hypothesis explains the tRNA maturation by RNase P as a vestige of dRNA loading. While ribozyme P remained the means of dRNA loading by oligonucleotide charging and duplicon trimming, its role for primitive tRNAs was limited to clearing any 5' leader that might interfere with their polynucleotide chain terminator function (this section), or the emergent amino acid charging and peptidyl transfer functions of their 3' end (sections 14-16). The tRNA elbow not only suppressed internal cleavage of double hairpin folds, but allowed evolution of the specificity domain S of ribozyme P with its molecular ruler mechanism to trim adventitious extensions of the tRNA 5' end (ref). The interdigitated double T-loop motif (IDTM) of this S domain was just the first of several IDTMs and other motifs that arose in charging ribozymes and rRNA to recognize the new tRNA elbow (sections 14-15).

At the breakout of codon-directed polynucleotide termination, the terminator tRNA was a fortuitous mimic of an elongator dRNA in the duplisome A-site. Under r-selection, its initial fate hung on its ability to agressively terminate copying of host

polynucleotides, releasing more duplisomes into the pool for initiation. Under K-selection, its eventual fate hung on its selectivity and restrain. However poorly at first, duplisome centers, evolved for dRNAs, decoded and accommodated this primitive tRNA nearly twice their size. There were three immediate differences between dRNA and tRNA decoding. First, STOP-codon recognition by the terminator tRNA was based on thermodynamic stability of the codon-anticodon helix, not kinetic competition of the codon with the duplicon for an anticodon toe-hold (section 11). Second, mimicking an open dRNA, the acceptor arm of the primitive tRNA accommodated quickly without awaiting the midday warmth. One final difference in tRNA decoding, we suggest, was that its anticodon trailer, notably N37, clashed with the P-site dRNA (Figure 13-8). With the even, or scrunched reading frame $_{3}NNN_{5}$ occluded, the tRNA anticodon favored the odd frame $_{4}NNN_{6}$.

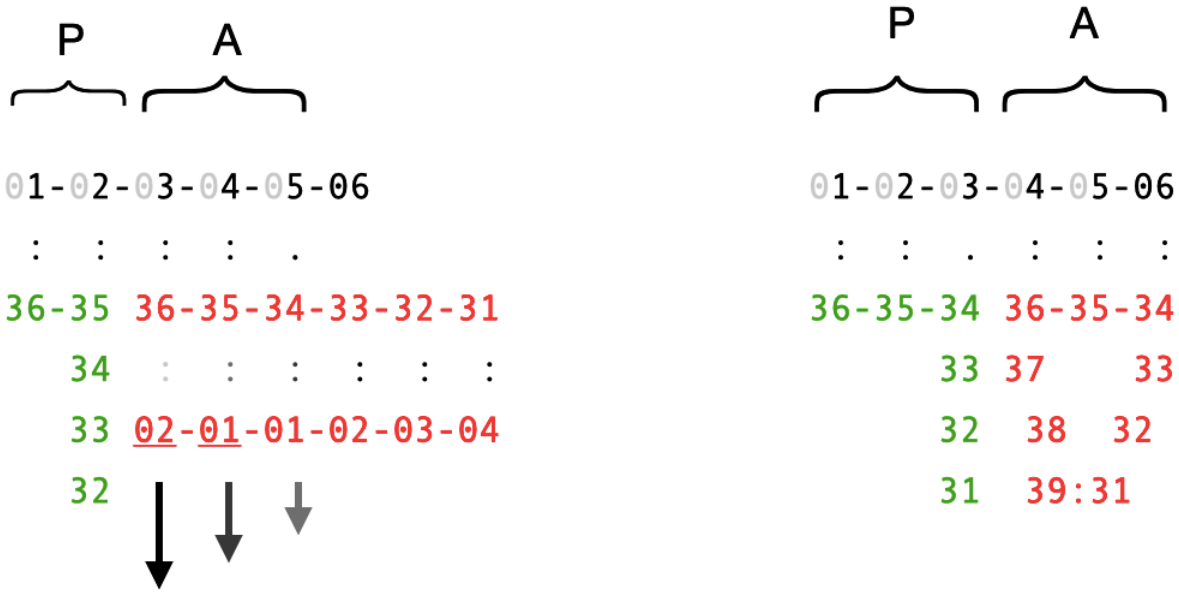


FIGURE 13-8. DECODING ELONGATOR dRNA VERSUS TERMINATOR tRNA

To exploit codon-directed polynucleotide termination, and avert conflicts with polynucleotide elongation, the primitive tRNA and duplisome centers underwent a number of rapid changes in no particular order. The choice of the odd, or unscrunched frame for RIBOSTOP codons was no doubt sloppy at first. However, there was immediate selection for the duplisome and tRNA to adopt one reliable frame to prevent premature

off-STOP termination. In sections 15-16 we suggest this nearly accidental choice of the odd frame for CODON-directed polynucleotide termination favored a triplet translocon at the breakout of polypeptide translation, leading to the elaborate mechanisms of reading frame defense found in protein translation today. Along with the change from polynucleotide to polypeptide products, this change from a doublet translocon in polynucleotide duplication to a triplet translocon in polypeptide translation is the most dramatic event in the duplisome origin of the ribosome.

Meanwhile, any tight squeeze to accommodate the tRNA acceptor arm in the nucleotidyl transfer center A-site was relaxed without compromising accommodation of dRNA duplicon-leaders. In the ongoing arms race between rival duplisome communities to domesticate killer tRNAs and turn them into genome guardians, terminators with new STOP-codon specificities drove improvements in orthogonal decoding of tRNA versus dRNA. In short order, the A-site of the terminating duplisome became adept at decoding $_4$NNN$_6$ and accommodating the cognate tRNA for polynucleotide termination, or more often, after no tRNA matched, decoding $_3$NNN$_5$ and accommodating the cognate duplicon-dRNA for polynucleotide elongation.

Several ingrained assumptions about origins of the genetic code are thrown into question by our conjecture that tRNAs arose as parasites of duplisome life, and were fully domesticated, before any association of tRNAs with amino acids, must less polypeptide translation. To make itself felt, a toxin or killer gene need only benefit (that is, less harm) some particular clique. But for a terminator tRNA to guard its host community, the costs of inadvertent, or off-STOP termination must be less than the benefits of on-STOP termination of non-self ribogenes. To duplicate 1000 nucleotides without an off-STOP termination, the duplisome must reject all non-cognate terminator tRNAs, while awaiting the cognate elongator dRNA, in 500 decoding sessions. That is, the rate of premature termination, over all and sundry codons, and their sequence contexts, should be less than 1 in 500.

In general, genome conflicts reach fitful steady states that only truly end when some major transition creates a higher-level agreement. Until then, the coevolution, or *arms race*, of say endemic parasites and their hosts is better understood as two, coupled competitions: one amongst parasites for improved virulence, and another

amongst hosts for improved resistance. Thus, parasites and hosts undergo cycles wherein a parasite variant invades some population of vulnerable hosts, and spreads through the population until some host variant emerges that can parry its attack.[19] As the fraction of resistant hosts rises, and the population achieves *herd immunity*, the numbers of active parasites decrease, allowing hosts to relax their immunological vigilance, setting the stage for a new epidemic, or reappearance of an old one. In the small world dynamics of infection, parasites make frequent jumps to closely related hosts, and rarer jumps to distantly related ones.

With accurate tRNA decoding and STOP-free cliques of all essential ribogenes, the terminator tRNA became a fairly cheap defense against foreign ribogenes. If one tRNA was good, were two tRNAs not better? The benefit of tRNA surveillance rose with each new STOP-codon, but the cost of maintaining a STOP-free clique, and the opportunity cost of foregoing certain triplets, or constraining them to the unguarded even frame, rose as well. We suggest that an RNA community employed at most a handful of guardian tRNAs at any time, but occasionally rehired a guardian from the pangenome whose contract had lapsed, or more rarely, encountered a killer tRNA never before seen in that population of hosts. Any invasion ended with extinction of the parasite, of the host population, or some sustainable cooperation of one host community and guardian tRNA. Through this macroevolutionary arms race of repeated infection and improved resistance, terminator tRNAs, and their accurate decoding, became the commonplace of duplisome life.

The parasite origin of tRNAs suggests that polynucleotide STOP-codons and their tRNA anticodons were simply those triplets that could be accurately read, and were common in the original host, and now, foreign ribogenes. As in other explanations of the origins of the genetic code, these first tRNA anticodons might favor Gs and Cs for the greater stability of C:G and G:C pairs (see Knight et al 2004). Or, considering the importance of limiting the frequency of STOP-codons, as well as averting off-STOP termination, these tRNA anticodons might avoid Gs and Us that form wobble U.G and G.U, as well as canonical C.G and A.U pairs. In any case, the coverage of the

[19] In the two-level model of VGT and HGT in RNA life, we reserve the term *population* for collections of *individuals*, or hosts, each of whom is a *community* (aka mutualism, clique) of ribogenes defined by some physical, organic, or sequence-based compartment (section 1).

canonical 64 triplet codons by some terminator tRNAs in the pangenome would never be more than partial, and the working coverage by terminator tRNAs in any RNA community much sparser yet.

The terminating duplisome had two decoding problems, terminator tRNA versus elongator dRNA, that required not only concurrent (aka *orthogonal*) solutions, but some level of coordination. As suggested above, most odd frame codons had no terminator tRNA, allowing the even frame codon to be read by an elongator dRNA. Because polynucleotide copying was relatively tolerant of occasional near-cognate (missense) dRNA, or small indels from shifts of the reading frame, dRNA decoding required only moderate accuracy. But because premature polynucleotide termination disrupts ribogene duplication and ribozyme expression, tRNA decoding required higher accuracy.

In the provisional pairing of mRNA codon and tRNA anticodon, the ribosome decoding center provides a dynamic environment for rapid and accurate selection and rejection of proffered matches (Ogle et al 2001, 2002). The key rRNA elements, helices h44 and h18, test the quality of codon-anticodon matching through domain closure. In this test, a combination of hydrogen bonds and isosteric fit determines whether nucleobases are correctly matched (cognate), nearly matched (near-cognate), or incorrectly matched (non-cognate). In the case of near-cognate matches, there is debate whether closure is prevented entirely, or is allowed, but destabilizes the codon-anticodon helix by forcing non-canonical pairs into less favorable Watson-Crick geometry (Demeshkina et al 2012; Khade et al 2013).

IDEA - the importance of accuracy in immune effector drove reliable 3 WC decoring

ribosome decoding center

first codon-anticodon basepair monitored & locked by A1493

second codon-anticodon basepair monitored & locked by A1492 & G530

all three SSU rRNA nucleobases are universally conserved

[Khade et al 2013] all of the hydrogen bonds from rRNA decoding center to codon-anticodon duplex are sequence independent/balanced

obvious for 2'-OH groups

true for contacts with nucleobases of mRNA

2’-OH of A1493 bonds O2 if pyrimidine and N3 if purine

() A-minor interactions in the decoding center to make first and second positions reliable again , recover the reading fidelity sacrificed The ribosome decoding center uses domain closure to ensure WC pairing in the first and second codon positions … somewhat more choice in the third position …

if 36 35 are WC and 34 is GU wobble or modified 34 pair then latch

ADD rRNA helix 44 latch and up the difficulty of accommodation (recover read 16)

E-site mRNA -3 -2 -1
tRNA 36 35 34
rRNA G963 uS7 G926

P-site mRNA 1 2 3
tRNA 36 35 34
rRNA A790 C1400

A-site mRNA 4 5 6
tRNA 36 35 34
rRNA A1913 A1493 A1492 G530 C1054
h44 A1492 A1493
h18 G530
Are this decoder nucleotides paired, bulged or looped?

[MOVE DOWN to polypeptide initiation] anti SD sequence at 3’ end of SSU rRNA beyond helix h45 is 1530GUAG**ACCUCCUUA1542**

QUESTION ….SD…AUG NNN ….

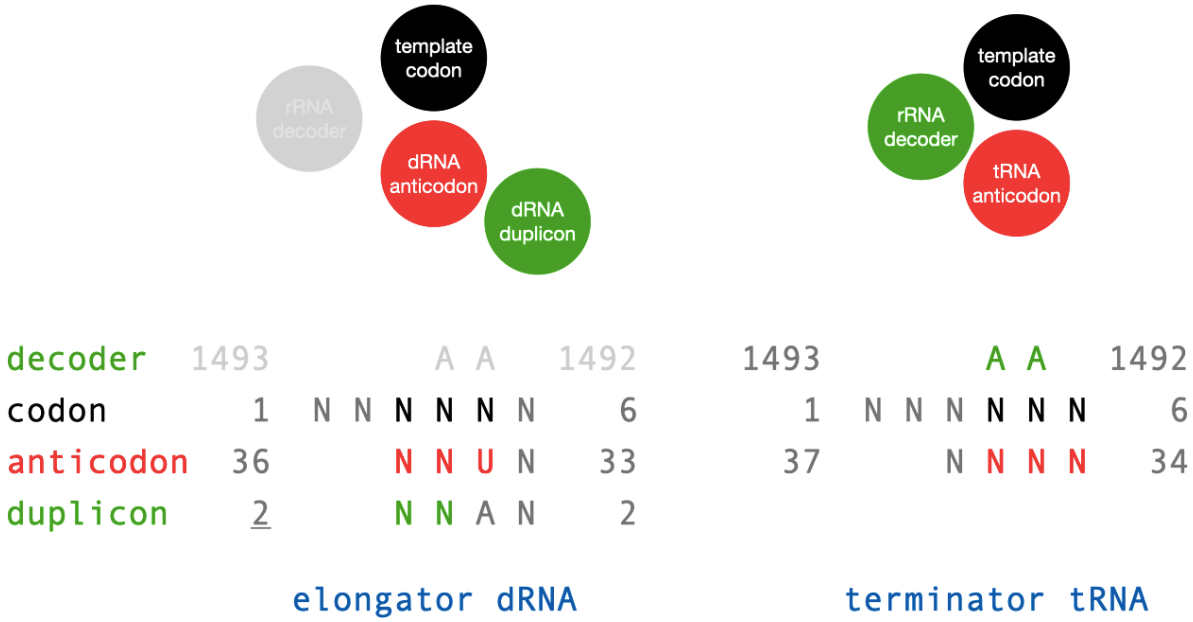


FIGURE 13-9. dRNA & tRNA DECODING IN THE TERMINATING DUPLISOME

We propose that the domain closure test for tRNA decoding arose in the terminating duplisome to reduce premature polynucleotide termination. In Figure 13-9 we depict how a common decoding center tests both tRNAs for polynucleotide termination and dRNAs for polynucleotide elongation. Both decoding mechanisms must approximate three polynucleotide stands, *viz*. codon-anticodon-duplicon in the toehold-opening test for dRNAs, and codon-anticodon-decoder in the domain-closure test for tRNAs. The anticodons of dRNAs and tRNAs occupy equivalent positions in the decoding center except that tRNAs are shifted by +1 nucleotide relative to the template RNA, reading the $_4NNN_6$ codon rather than the $_3NNN_5$ codon. The decoder strand of the duplisome is positioned opposite the duplicon strand of closed dRNAs so that tRNA decoding does not occlude dRNA decoding. The kinetic picture is that rapid elective tRNA decoding of codon $_4NNN_6$ preempts slow default dRNA decoding of codon $_3NNN_5$. tRNA decoding needs high selectivity of all three positions, dRNA decoding requires ignoring the nucleobase identity of the superwobble position.[20]

[20] Because of the difference in reading frame, as well as the decoding mechanism, any similarity of superwobble decoding of the third codon position in conjectural dRNAs and the third codon position in actual tRNAs is a formal analogy, not an evolutionary homology.

Beyond the domain closure test for tRNA selection, the accuracy of tRNA decoding including reading frame definition has been improved through modifications of tRNA anticodon arm (Agris et al 2017). Sundry changes in tRNA decoding along the way to the canonical genetic code of protein translation are challenging to reconstruct as new tRNAs were introduced and existing tRNAs changed their codon specificities both to read new codons or ignore old ones (sections 14-16). In protein life, translational GTPases EF-Tu allows the domain closure test to be applied twice, once on initial selection with accommodation blocked, and once again after GTP hydrolysis, with accommodation allowed (sections 16). In general, molecular recognition with essentially the same discrimination repeated twice, separated by an exergonic step, to nearly square the accuracy, is known as kinetic proof-reading (ref).

Besides irreversible cleavage of ribogenes or their folded ribozymes by endonuclease ribozymes, and RIBOSTOP codon-directed polynucleotide termination, there were likely ancient mechanisms to regulate polynucleotide elongation more-or-less reversibly (*cf*. Buskirk & Green 2017). Here we suggest that the signal recognition particle (SRP) RNA arose in RNA life as a duplisome-bound riboswitch that monitored nascent polynucleotides, or other signals, to regulate polymer elongation. Nearly one-third of the cellular proteome is targeted for membrane insertion or secretion. Already LUCA had phospholipid cell membranes, and a mechanism for co-translational secretion based on the ribosome-bound SRP that scans emerging polypeptides for a hydrophobic segment, either an N-terminal signal sequence, or the first transmembrane α-helix, which cytoplasmic proteins lack (Voorhees & Hedge 2015). When it engages a hydrophobic segment, SRP pauses polypeptide elongation until it docks with SRP receptor (SR) on the cell or ER membrane. There it hands the ribosome nascent chain complex over to the membrane translocon (prokarya SecY, eukarya ER Sec61) to resume translation, now coupled to secretion.

SRP has an RNA, and a secretory GTPase (prokarya Ffh, eukarya SRP54), while its membrane receptor has an homologous GTPase (prokarya FtsY, eukarya SRα). SRP RNA has two domains, called S and Alu, that recognize the hydrophobic polypeptide segment, or arrest elongation, respectively. The S domain presents the SRP GTPase to the polypeptide exit tunnel where this protein scans for emergent hydrophobic

segments. Upon recognizing the signal polypeptide, the SRP GTPase on the ribosome docks with the SR GTPase on the membrane, where they activate one another to allow ribosome handover. From signal recognition to ribosome handover, the Alu domain pauses polypeptide elongation by blocking the factor binding site at the ribosome subunit interface. In *Bacillis subtilis* this domain interacts directly with rRNA stalk-base helices H43 H44 and the α-sarcin-ricin loop on H95 (Beckert et al 2015). Elongation arrest is less important in the small bacterial cell where the ribosome-SRP complex quickly encounters its receptor on the cell membrane (Wild et al 2020). In *Escherichia coli*, the Alu domain has been lost entirely, leaving only the S domain.

We conjecture that SRP RNA originated in duplisome life to regulate polynucleotide elongation, and was later exapted for polypeptide elongation. For sake of discussion, we suggest that the ribosome factor binding site for translational GTPases arose as a duplisome binding site for Alu-related riboswitches that enabled or disabled elongation (Ahl et al 2015). An early form of regulation for gene replication and expression, there are several likely roles for Alu-riboswitches with some form of sensor domain, e.g., an RNA guide sequence, or a polypeptide recognition factor. These include (1) pausing polymer elongation to synchronize growth, force polymer release, or divert resources from elongation to initiation; (2) blocking elongation of polymers with specific identifier sequences; and (3) blocking polymers with misfolded, pathogenic, or foreign segments.

14. RNA termination | conditional tRNAs

We proposed that tRNAs arose in duplisome life as polynucleotide chain terminators that invaded naïve populations as killer genes, wrecking havoc on susceptible hosts, until adapted as genome guardians of resistant hosts. Here we propose that in the heyday of constitutive termination, some tRNAs were exapted from roles as clique champions to new roles in metabolite-regulated termination. To wit, ribogene duplication and expression were regulated by decisions at RIBOSTOP-codons whether to terminate or continue polynucleotide elongation. Thus, tRNAs DOWN-regulated ribozyme expression by terminating early, or UP-regulated expression by terminating after the catalytic domain, but before an inhibitory domain. The activity of these regulatory tRNAs was conditioned, we propose, on the presence or absence of specific metabolites esterified at O2'/O3' of their terminal ribose. It is easy to see how a bulky ligand, attached to the 3' end of the acceptor arm, might preclude tRNA selection for polynucleotide termination. In the absence of that ligand, say a nutrient, ribozymes needed for its catabolism could be DOWN-regulated by cognate RIBOSTOP-codons positioned early in their ribogenes.

There are three mechanistic differences between our conjectural regulatory tRNAs and the familiar riboswitches and regulatory proteins of cellular life: First, whereas ligands bind these cellular factors non-covalently and reversibly, they were covalently attached to terminator tRNAs. Second, whereas ligand specificity and affinity are determined by ligand-recognition domains intrinsic to riboswitches and regulatory proteins, the 'smarts' for matching ligands to tRNAs came from tRNA-specific metabolite-specific acyl ligase ribozymes. Finally, metabolic states were evaluated over a longer timescale in duplisome life. The crucial decisions at STOP-codons, whether to accept a terminator tRNA, or await an elongator dRNA, were entrained to the diurnal polynucleotide elongation cycle itself. Moreover, several RIBOSTOP-codons distributed over the length of the ribogene, using the same or different regulatory tRNAs, tested conditions, and integrated the response, over the entire course of duplication. Thus, STOP-codons positioned upstream within the ribogene allowed timely redeployment of duplisome resources to other ribogenes, while STOP-codons positioned downstream

allowed the ribozyme to be activated or aborted based on the latest metabolic conditions, or to complete a 3' inhibitory domain, creating a full-length ribogene that could be copied itself, or cleaved to form the active ribozyme.

The differences of conditional tRNAs from riboswitches in their mechanism of gene regulation affected the evolvability of regulatory circuits. First, we suggest, ligand specificity of regulatory tRNAs was determined by their cognate acyl ligase ribozyme accessing an aptamer space for natural ligands. Second, the actions of regulatory tRNAs were genome-wide, not tied to one ribogene as for *cis*-acting riboswitches found in mRNA leaders. Thus, adaptation of the regulatory circuit was as simple as introducing (or eliminating) the cognate STOP-codon at a suitable position along the target ribogene. Third, in any individual ribocommunity, the termination code of STOP-codons and metabolites, defined by the working set of conditional-terminator tRNAs and their charging ribozymes, was limited to a few isoacceptors. The pangenome of all ribocommunities, meanwhile, became a rich repository of different combinations of terminator tRNA, specifying the target codon, and its charging ribozyme, recognizing the ligand. We suggest below that these ribozymes and their cognate tRNAs were likely transmitted together as modular ribogenes (see Figure 14-2).

Few things about the origin of protein coding have been more puzzling than tRNA charging, catalyzed by twenty-some aminoacyl-tRNA synthetase (aaRS) enzymes, one for each proteinogenic amino acid, with curious exceptions (Rubio-Gomez & Ibba 2020). Each charging enzyme recognizes one or more tRNA isoacceptors by some combination of positive and negative identifier nucleotides, found not just in the anticodon, but within the acceptor and variable arms as well (Giege & Eriani 2023). Besides a cognate tRNA, these enzymes require ATP and the free amino acid as substrates, proceeding through the aminoacyl-adenylate intermediate bound non-covalently.

Averting the chicken-or-egg paradox of needing an entire family of aaRS enzymes to make any coded protein, early molecular biologists looked for direct chemical affinities of amino acids and oligonucleotides to explain tRNA charging *before support of coded enzymes*. Fitting individual amino acids into the crevices between adjacent basepairs of duplex DNA, George Gamov launched the quest for stereochemical fits

between amino acids and code words, or *codons* of two or more nucleotide letters (Gamov 1954). Besides combinatoric objections to Gamov's proposed map from DNA to polypeptide sequences, Crick doubted that van der Waals interactions with the hydrophobic faces of nucleobases could discriminate between amino acid sidechains (Crick 1955).

After elucidation of the codon triplets of the genetic code, as well as the anticodon triplets that read them, the search for stereochemical fits shifted from duplex DNA to mRNA and tRNAs (Nirenberg 19##; Holley 1965). Some looked for direct binding of amino acids to tRNAs, or just their anticodons, while others looked for spontaneous tRNA charging from aminoacyl-adenylates, or other activated forms of amino acids. These searches failed to uncover any clear stereochemical fit between tRNAs and amino acids underlying the canonical genetic code (Schimmel & Ribas de Pouplana 1995).

Under the RNA world hypothesis, the search for a primitive mechanism of tRNA charging shifted, from spontaneous reactions exploiting direct fits between tRNAs and amino acids, to reactions mediated by ribozymes with aptamers for aminoacyl-AMPs or another activated form of the amino acids. Today a variety of artificial ribozymes can catalyze RNA aminoacylation. For experimental convenience, as well as research objectives, these models depart from the enzymatic tRNA charging reaction in their substrates. Supplying the electrophile (carbonyl carbon), chemically activated amino acids include aminoacyl-adenylates (Illangasekare et al 1995, 1999), CoA thioesters (Li & Huang 2005), cyanomethyl and 3,5-dinitrophenol esters (Murakami et al 2006; Ohuchi et al 2007), and oxalones (Pressman et al 2019; Liu et al 2020; Janzen et al 2022). Supplying the nucleophile (ribose alcohol), polynucleotide substrates include full-size tRNAs, minihelices, NCCA, or O2' of an internal ribose. In some models, the same RNA acts as catalyst and nucleophile donor, allowing direct selection of catalytically active sequences via the self-aminoacylation product.

Because entropic catalysis of bimolecular reactions uses substrate concentration and position, there is no bright line between catalytic and structural RNAs that approximate reactants through some combination of basepairing and aptamers. In model reactions using tRNA mimics, the carbonyl carbon of aminoacyl-5'-phosphate of

one RNA is attacked by the 3' OH of another RNA approximated through their basepairing, either directly to one another, or via a bridging RNA intermediate (Tamura & Schimmel 2003; Wu et al 2021). With suitable reactants the product can be extended from the aminoacyl-ester to peptidyl- and dipeptidyl-esters. One curious result is that the D-ribose of modern RNA favors charging with the L-amino acid of modern proteins (Tamura & Schimmel 2006). Those authors suggested that as RNA communities settled on D-ribose, they put the imprimatur of nucleotide chirality onto amino acids and peptides through this charging reaction.

Seeking a practical reagent to charge tRNAs with non-cognate, or even non-proteinogenic amino acids, Hiroaki Suga and colleagues perfected a short (46 nt) artificial ribozyme that acylates virtually any natural tRNA using 3,5-dintrophenol esters of α-amino- or α-hydroxy-acids (Lee et al 2000; Murakami et al 2006). This reagent dubbed the *flexizyme* recognizes only the 3' terminal NCCA of the acceptor arm, not the L-shaped fold, much less discriminates among tRNAs. Designed and selected for substrate promiscuity, the flexizyme says little about likely substrates or mechanism of extinct ribozymes for charging tRNAs, or acylating RNAs more generally. More recently, ribozymes have been selected for self-aminoacylation at internal 2' OH using biotinyl-tyrosyl-oxazolone (Pressman et al 2019). Starting from three of these ribozymes as seeds, a spectrum of ribozymes selected to utilize oxazolones of F I L M V W show positive correlation between specificity and activity (Janzen et al 2022).

If artificial ribozymes are doubtful models of tRNA charging in the RNA world, T-box riboswitches, found in leaders of bacterial mRNAs, are likely vestiges of the lost aaRS ribozymes (Grundy & Henkin 1993; Suddala & Zhang 2019; Ishida et al 2020; Zhang 2020). These riboswitches have now been found in Gram-positive bacteria for tRNA isoacceptors of all twenty proteinogenic amino acids. T-box riboswitches monitor individual amino acids through the peculiar lens of tRNA aminoacylation to regulate anabolism. Thus, they recognize individual tRNAs and test their aminoacylation status to regulate the transcription or translation of their aaRS enzyme, as well as enzymes for biosynthesis and transport of the cognate amino acid.

The 5' domain of a typical T-box riboswitch comprises an interdigitated double T-loop motif (IDTM) to recognize the tRNA elbow, and a triplet *codon* to specify its

anticodon (Zhang & Ferre-D'Amare 2016; Suddala et al 2023). Pairing with the terminal NCCA of the uncharged tRNA, the 3' domain of the riboswitch effectively tests whether an amino acid is present, though not whether it is cognate for that tRNA. Different T-box riboswitches read-out the absence of an amino acid on tA76 in either of two ways: masking the terminator to allow transcription to continue, or unmasking the Shine-Dalgarno box to allow translation to commence.

It bears repeating that T-box riboswitches are likely vestiges of tRNA charging ribozymes in polypeptide life that recognized the tRNA elbow and anticodon to selectively charge the acceptor arm. In proof of concept, Suga and colleagues have engineered the riboswitch from the *Bacillus subtilis* glyQS gene to charge its cognate $tRNA^{Gly}$ (GCC) using phenylalanyl-cyanomethyl ester (Ishida et al 2020; Lu et al 2024). Whether or not T-box riboswitches descend from tRNA charging ribozymes, this chimera of a natural riboswitch and an artificial ribozyme demonstrates the feasibility of a genetic code based purely on RNAs. As these modular riboswitches recognize well-separated features of the L-shaped tRNA from elbow to anticodon to acceptor NCCA, it is easy to see that the matching of anticodons and amino acids in the original code may have been purely conventional, with no subtle stereochemical predispositions. As the recognition elements for tRNAs and amino acids are independent, their peculiar combinations were likely just accidents of pairing or ligation.

In constituitive-terminator tRNAs, elbow interaction with the S domain of ribozyme P thwarted undesirable reactions of tRNA cleavage into halves, and duplicon loading at the 5' end. In conditional-terminator tRNAs, a new family of ribozymes acylated ligands to the 3' end. Here we propose that each tRNA charging ribozyme had two parts: an elbow and anticodon recognition domain, surviving today in the 5' domains of T-box riboswitches, and an older polynucleotide acyl ligase domain (section 12). These tRNA-specific metabolite-specific acyl ligase ribozymes, dubbed *ribozymes T*, hydrolyzed the tRNA tail to drive condensation of the freed 3' OH with small metabolites. Based on the 5' recognition domains of T-box riboswitches, and the minimal ribosome peptidyl transfer center, ribozymes T were perhaps 300-400 nucleotides overall, or four-fold larger than their tRNAs.

We suggest that conditional tRNAs and their metabolite-charging ribozymes coordinated all aspects of replication and metabolism in duplisome life through the patterning of STOP-codons in ribogenes. Much of the logic of gene regulation found in cellular life has possible analogs in duplisome life mediated by simple variations on conditional polynucleotide termination. For example, conditional tRNAs likely allowed RIBOSTOP-directed termination and reinitiation of the intact duplisome to make multiple products from one template. Moreover, the same tRNA might have two or more charging ribozymes with opposing effects.

Besides amino acids, tRNA ligands likely included esters of other carboxylic acids, e.g., glucoronic acid, fatty acids, peptides, etc., arising in RNA life. Acyl ligases of ribozymes T discriminated between these metabolites in the tRNA charging reaction, but the role of these ligands for tRNA activity in the duplisome was less specific. For UP-regulation of catabolic pathways, a variety of ligands could likely interfere with tRNA entry, or accommodation of the acceptor end. For DOWN-regulation of anabolic pathways, duplisomes favored the charged tRNA over its uncharged form. In particular, the nucleotidyl transfer center was adapted to accommodate the acylated acceptor arm. We suggest one set of ligands, L-α-amino acids, with uniform structure, modest size and positive charge, as well as their charging ribozymes, became prominent regulators of anabolism.

In Figure 14-1 we present a simple scheme for how ligands were attached to tRNAs, and later removed, as well as how they disabled, or enabled entry and accommodation in the duplisome. First, a tRNA-specific metabolite-specific acyl ligase, or ribozyme T, charged the tRNA with the cognate ligand if available, else released the cleaved tRNA uncharged. Thus, the tailed substrate of the charging reaction **tRNA NNNNNN…** differed from its two alternative products, *viz*. the truncated tRNA either with or without the ligand. *In the jargon of computer science, the charging reaction performs a test on metabolite availability.* The readout of this test was different according to whether the free tRNA (left), or the charged tRNA (right), was accommodated in the duplisome. The charged tRNA eventually underwent spontaneous or ribozymatic deacylation. Based on aminoacyl-tRNAs, the half-lives of various acyl-tRNAs from spontaneous hydrolysis likely ranged from hours to days.

Thus, the tempo of acylation and deacylation was compatible with the diurnal elongation cycle of duplisome life.

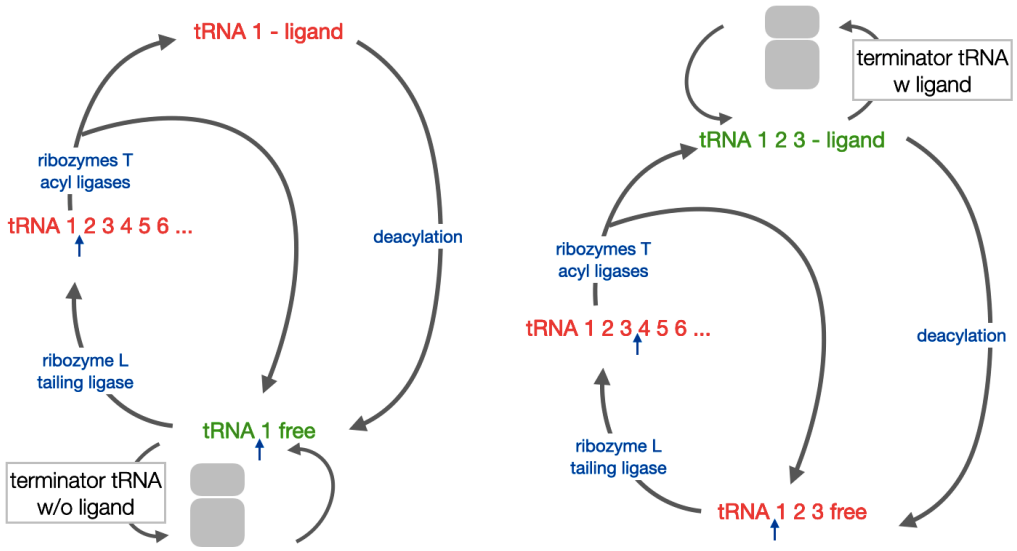


FIGURE 14-1. tRNA CHARGING CYCLE FOR OFF-LIGANDS (LEFT) & ON-LIGANDS (RIGHT)

In Figure 14-1 we depict slightly different sites of cleavage and acylation for down-regulated and up-regulated tRNAs, respectively, where the object is to exclude the charged form of tRNA from the duplisome, or exclude the free form, respectively. Finally, free tRNAs were not substrates for ribozymes T until a polynucleotide ligase restored their 3' tails. A single *tRNA tailing ligase* dubbed ribozyme L, descendant of the primordial polynucleotide ligase ribozyme (section 12), might work on all tRNAs. Its site for sequential hydrolysis-condensation demarks the core tRNA from the tail proper. In section 16, we relate the biogenesis and novel functions of the universal tRNA CCA tail in protein life to this lost tRNA tailing reaction within the regulatory cycle of conditional-terminator tRNAs in RNA life.

Useful combinations of metabolite-specific acyl ligase ribozymes and their cognate tRNAs likely moved through the pangenome as ribogene units. For sake of discussion, we show one such arrangement in Figure 14-2 where the full-length polynucleotide is cleaved into an active ribozyme T and its cognate tRNA by ribozyme P (small blue arrow). Here, before the clear divisions of labor between gene replication and expression in later polymer life, activation of the products entails destruction of a full-length template.

FIGURE 14-2. RIBOGENE OF RIBOZYME T AND COGNATE tRNA

We suggest that radiation of the family of tRNAs began in the RNA world. For constitutive terminators, there was strong selection for tRNA variety compatible with precise decoding of codon triplets in the odd duplication frame. Preserving their common interactions with the duplisome centers and ribozyme P, these guardian tRNAs initially diversified through single nucleotide substitutions in the anticodon. New opportunities for tRNA diversity emerged with conditional terminators where elements beyond the anticodon helped recognition and charging by cognate acyl ligase ribozymes.

For sake of discussion, we propose that the self-splicing intron in the tRNA anticodon arm arose in duplisome life, where it proved useful to explore the opportunities of regulatory tRNAs. A variety of tRNA genes in bacteria and plastids have a group I self-splicing intron inserted within or immediately after the anticodon (Kuhsel et al 1990; Xu et al 1990; Reinhold-Hurek & Shub 1992). The canonical non-autonomous introns found between N37 and N38 in archaea and nuclear tRNA genes, removed by tRNA splicing endonuclease (TSEN) and ligase enzymes, are likely vestiges of these self-splicing introns (ref). Thus, a subset of tRNAs carried introns in this position in LUCA tracing to perhaps one insertion earlier in polymer life. We suggest this seed event of reverse splicing into the anticodon arm of a terminator tRNA allowed rapid prototyping of new tRNAs through mix-and-match of identifiers in the two halves. Finally, the variable arms of tRNAs have identifiers for aaRS enzymes, but no other clear function today (ref; cf tmRNA). From their position after the anticodon arm, they may be vestiges of imperfect tRNA *trans*-splicing events. Conceivably, they had an aptamer role in duplisome life to help enrich or hold metabolites for acylation of regulatory tRNAs.

Before we consider the breakout of polypeptide translation in duplisome life in section 15, it is worth reviewing how greatly we have reframed the problem. First, tRNAs themselves arose for polynucleotide termination, and were not at first aminoacylated, much less had a role in random polypeptide synthesis. Second, amino acids were part of a wider set of metabolites used to regulate tRNA actions as polynucleotide terminators. Finally, there is no reason to think that more than a fraction of the 64 triplet codons were read by terminator tRNAs at this stage. *What pre-adapted aminoacyl-tRNAs for polypeptide translation was these tRNAs worked via accommodation, rather than exclusion, of their charged acceptor arms in the duplisome.*

15. Breakout of polypeptide translation

Making one crucial change to the polynucleotide elongation cycle in Campbell's duplisome hypothesis (Campbell 1991), plus our conjecture on the origin of tRNAs as polynucleotide chain terminators, we distinguish eight eras of polymer life: (1) spontaneous RNA copying up to the first ribozymes for RNA recombination and repair; (2) processive RNA copying mediated by the duplisome and its dRNAs; (3) constitutive polynucleotide termination using plain tRNAs; (4) conditional termination using chargeable tRNAs with ligand-specific tRNA-specific acyl ligase ribozymes; (5) polypeptide translation without assistance from coded products; and (6) protein translation incorporating coded proteins and enzymes into all aspects of bioenergetics and metabolism, including translation *tout court*. Finally, there was an era of polymerase enzymes leading to (7) retirement of the duplisome; and (8) genome takeover of DNA life.

Despite the moniker *RNA world*, simple amino acids and random polypeptides were no doubt abundant and useful to RNA life (Noller 2004; Cech 2009; Frenkel-Pinter et al 2020). The former included primary amino acids with prebiotic feedstocks, as well as secondary amino acids first made with ribozymes. The latter included low complexity, intrinsically disordered polypeptides that assembled with polynucleotides to form fluid coacervates or structured ribonucleoproteins, and with lipids to form membrane pores (ref). During the long *conviviencia* of noncoded polypeptides with templated polynucleotides, there was selection for ribozymes that improved the syntheses of useful amino acids and polypeptides. To focus on the origin of protein coding, we stipulate aspects of amino acid and polypeptide synthesis that do not appear to favor the duplisome over other, conjectural replicase ribozymes at the breakout of polypeptide translation (Orgel 1989).

The changes in polypeptide and protein synthesis were little less remarkable than the changes in nucleic acid synthesis.

(1) prebiotic, spontaneous, random: dry-down condensation … leaving groups? primary amino acids

(2) biotic ribozymatic non-coded: peptide or amino acid carriers & activation; secondary amino acids

- creative proposals in both (1) and (2) for activation of amino acids or nucleotides by combination with one another (refs). Like the earlier question of xenogenomes and ante-RNA life, we leave the … many of these suffer the borrow from Peter to pay Paul …. identify and solve a problem, only to elide it entirely from the palmpesest of cellular life … streetlight effect of interesting chemistry rather than necessary chemistry …

(3) breakout ribosome coded: aminoacyl tRNA and peptidyl tRNA [section 15]

- ribozyme T charges tRNAs and peptidyl transfer center

(4) modern ribosome with protein support of translation [section 16]: aaRSs , translational GTPases, amino acid biosynthesis

There are various proposals for abiotic high-energy substrates handed over to ribozymes, as well as downstream intermediates in these catalyzed pathways to polypeptides (Liu et al 2019, 2020). Extrapolating from modern cells, the most important carriers of amino acids and polypeptides in late RNA life were likely mono- or polynucleotides, activated as higher-energy acyl-phosphate mixed anhydrides and lower-energy acyl-esters. The free energy of hydrolysis for aminoacyl-phosphates is about 4 kcal / mol higher than aminoacyl-esters, which in turn, is about 8 kcal / mol higher than peptide bonds (Carpenter 1960). Thus, aminoacyl-phosphate anhydrides of RNA carriers were likely abiotic inputs, or early intermediates in ribozymatic peptide synthesis (Leman et al 2006).

energy comparison of acyl-phosphate to pyrophosphate

kinetic comparison of acyl-phosphate to pryophosphate [lacks second negative charge and thus more reactive to nucleophiles]

At the breakout of polypeptide translation in the late RNA world, there were already brisk markets for random polypeptides of simple compositions. Useful polypeptides were subject to depletion, creating pent-up demand for any mechanism that better supplied them. Coming at the heyday of random polypeptides, the first coded polypeptides did not need any great variety of residues, nor precise control of their sequence, to be versatile and useful. Thus, a rudimentary code of just two or three classes of amino acids could supply current demands, as well as hock new wares (ref).

Among their virtues, the first coded polypeptides likely had greater, more reproducible length, regular repeats, and distinct domains marked by abrupt transitions from one composition or sequence to another. Moreover, using different mRNAs, or one mRNA with different starts and reading frames, polypeptides of different size, composition and sequence could be made concurrently from the common pool of amino acids.

Given the terminating duplisome with 16 dRNAs for polynucleotide elongation, and two or more aminoacyl-tRNAs for polynucleotide termination, we propose that the polynucleotide elongation cycle was co-opted for polypeptide elongation. The mechanisms of tRNA charging and decoding carried over unchanged, so that *the mature amino acid code for polynucleotide termination became the first genetic code for polypeptide elongation*. Contrasting the continuity of tRNA decoding, polypeptide translation required an abrupt, if catalytically modest, change in the polymer transfer reaction. In Figure 15-1 we compare polynucleotide termination and peptide bond formation. Both reactions use an aminoacyl-tRNA in the A-site, but polynucleotide termination has a polynucleotidyl-dRNA in the P-site, while polypeptide elongation has a peptidyl-tRNA in that site. Only small changes in substrate positioning likely were needed for this center evolved for nucleotidyl transfer and polynucleotide release to now catalyze peptidyl transfer and polypeptide release. All of this changed the energetics of elongation, both the covalent chemistry of substrates and products, and the Brownian movements of the duplisome and its thermal ribomotors.

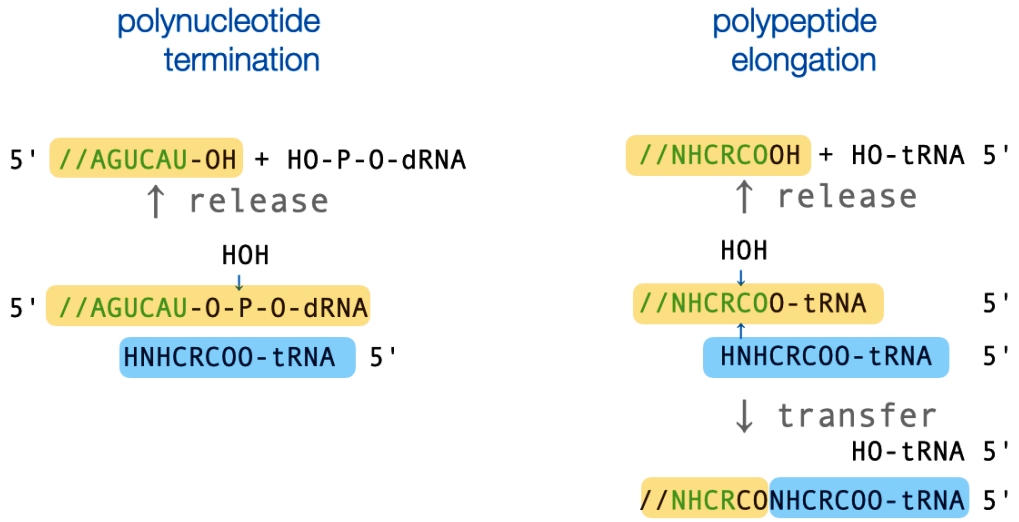


FIGURE 15-1. FROM POLYNUCLEOTIDE TERMINATION TO POLYPEPTIDE ELONGATION

Perhaps the simplest model for initiation of polypeptide translation comes from the presence of two (or more) adjacent STOP-codons for terminator tRNAs charged with amino acids (Figure 15-2). With a polynucleotidyl-dRNA in the P-site, and an aminoacyl-tRNA in the A-site, hydrolysis releases the nascent polynucleotide through the exit tunnel. Translocation, driven by cold closing and exit of the freed dRNA through the E-site, brings the aminoacyl-tRNA into the P-site pending duplisome disassembly. But now the second "STOP-codon" has entered the A-site, although there is no more polynucleotide to terminate. If this codon selects a cognate aminoacyl-tRNA, the stage is set for polypeptide initiation with the formation of the first peptide bond. With acceptor arms of aminoacyl-tRNAs in both the P- and the A-sites of the transfer center, the nucleophile (α-amine nitrogen) of one attacks the electrophile (carbonyl carbon) of the other. Some subtlety of relative substrate positions, or perhaps chemical protection of the initial residue, analogous to formylation of methionine of the initiator tRNA in bacteria, favored the A-site α-amine to form this first peptide bond. Whatever the cause, once symmetry is broken, so that the proximal exit tunnel carries a nascent peptide or dipeptide from the P-site tRNA, the direction of transfer is sterically constrained in all further elongation.

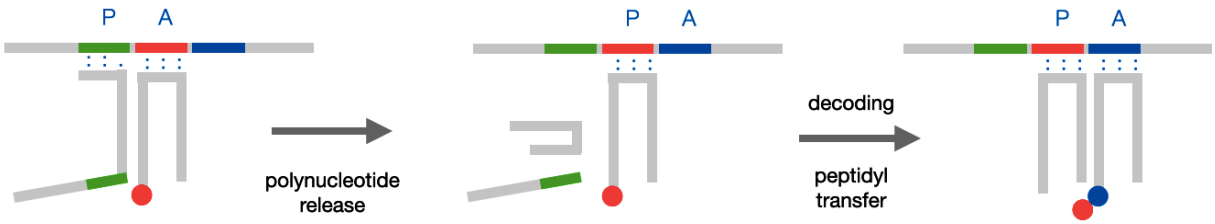


FIGURE 15-2. POLYPEPTIDE INITIATION COUPLED TO POLYNUCLEOTIDE TERMINATION

In the scheme above coupling polypeptide initiation to polynucleotide termination, RIBOSTOP-directed polynucleotide release immediately initiates translation. Uncoupling polypeptide translation from polynucleotide duplication, *primitive leaderless initiation* from the A-site likely carried over from the duplisome to the ribosome (*cf*. Figure 12-3). Here leaderless mRNA and initiating aminoacyl-tRNA enter the A-site with no adaptor

or template in the P-site at all (Figure 15-3). Although the favorable Gibbs energy of phosphoester bond hydrolysis and cold closing of the freed dRNA from polynucleotide termination are absent in leaderless initiation, the greater stability of acylamide than acylester bond still drives peptidyl transfer.

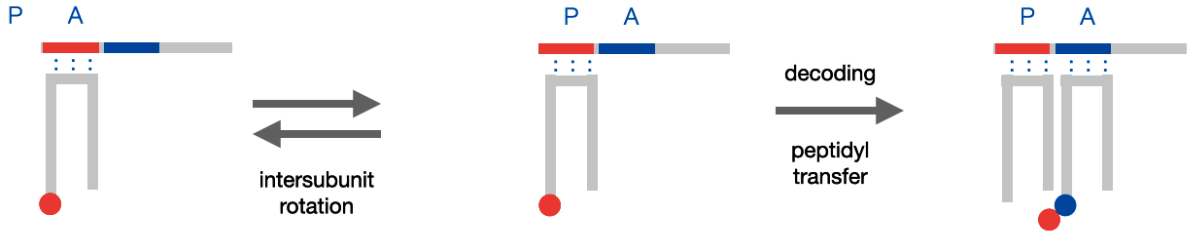


FIGURE 15-3. PRIMITIVE LEADERLESS INITIATION OF POLYPEPTIDE TRANSLATION

Like polynucleotide duplication, there were special problems, as well as opportunities for regulation, associated with the initiation and termination of polypeptide translation. Before any assistance of coded proteins, these events were either spontaneous, or mediated by RNA and noncoded polypeptides. At the breakout of polypeptide translation, the ribosome née duplisome likely had little or no control over the site of initiation, and the question of whether to translate or to duplicate likely had precedence over just where to start translation. Moreover, the reading frame of polypeptide translation was more important than the position of the initiation site along the mRNA. Thus, the breakout ribosome likely had no dedicated initiator codons and tRNAs for P-site initiation, but rather any codon in the A-site, internal or leaderless, could initiate translation from its cognate aminoacyl-tRNA. The evolution of peptide START-codons and initiator tRNAs that bypass the A-site allowed programmed initiation and reinitiation of translation at internal sites in the mRNA (section 16).

For termination of peptidyl-tRNA paired to its mRNA codon in the ribosome P-site, there are a number of possibilities for the A-site occupant.[21] Thus, peptidyl-tRNA likely underwent slow hydrolysis whenever elongation was stalled by absence of aminoacyl-tRNA in the A-site because the cognate tRNA was depleted, or just having reached the

[21] We assume that dRNAs had no role in polypeptide termination. Besides the constraints of the diurnal elongation cycle on dRNA decoding, peptidyl-tRNA in the P-site may have occluded their selection in the A-site. Later some change in the ribosome née duplisome itself, such as the incorporation of 5S RNA (discussed below) or specialization of the L12 stalk, may have precluded their entry entirely.

3’ end of the mRNA. Before the discovery of class I release factors, early molecular biologists suspected that one or more special tRNAs might trigger a fast affirmative decision to terminate (Scolnick et al 1968; Brenner 1971). Indeed under non-physiological conditions, e.g., 30% acetone, codon-directed release factors can be replaced by deacylated-tRNA, or just the tRNA mimic CCA (Caskey et al 1971; Bao et al 2022). In codon-directed tRNA-dependent termination, polypeptide terminator tRNAs might either be constituitively unchargeable, or conditionally uncharged, affording a new level of regulation.

There have been numerous experimental attempts to separate the ribosomal mechanisms for peptide bond formation and polypeptide release, distinguishing catalysis *simpliciter*, from possible conformational switch of the peptidyl transfer center from aminolysis to hydrolysis. Several observations appear pertinent: First, polypeptide elongation is much faster than termination, *viz*. bacterial ribosomes catalyze 15-20 peptide bonds per second while factor-mediated polypeptide release is 10-40 fold slower (Katunun et al 2002; Zavialov et al 2002). Second, in both bacteria and eukarya, certain fungal antibiotics inhibit peptide bond formation and stimulate peptide release (Caskey et al 1971; Polacek et al 2003). Third, several conserved nucleotides in the inner shell of the peptidyl transferase center are essential for polypeptide release, but not peptide bond formation (Youngman et al 2004; Amort et al 2007).

There is an obvious mechanistic analogy between polynucleotide release in the duplisome and polypeptide release in the ribosome, as well as significant disanalogy. On the one hand, hydrolysis of the phosphoester of the polynucleotide-dRNA in the duplisome P-site is necessary for polynucleotide release, but is also the obligate first step in polynucleotide transfer. On the other hand, hydrolysis of the acylester of the polypeptidyl-tRNA in the ribosome P-site is necessary for polypeptide release, but competes with polypeptide transfer by aminolysis (Polacek & Mankin 2005). Molecular modeling of polynucleotide-dRNA in the ribosome transfer center might suggest whether or not hydrolysis of this conjectural substrate in the duplisome could be homologous to hydrolysis of actual peptidyl-tRNA.

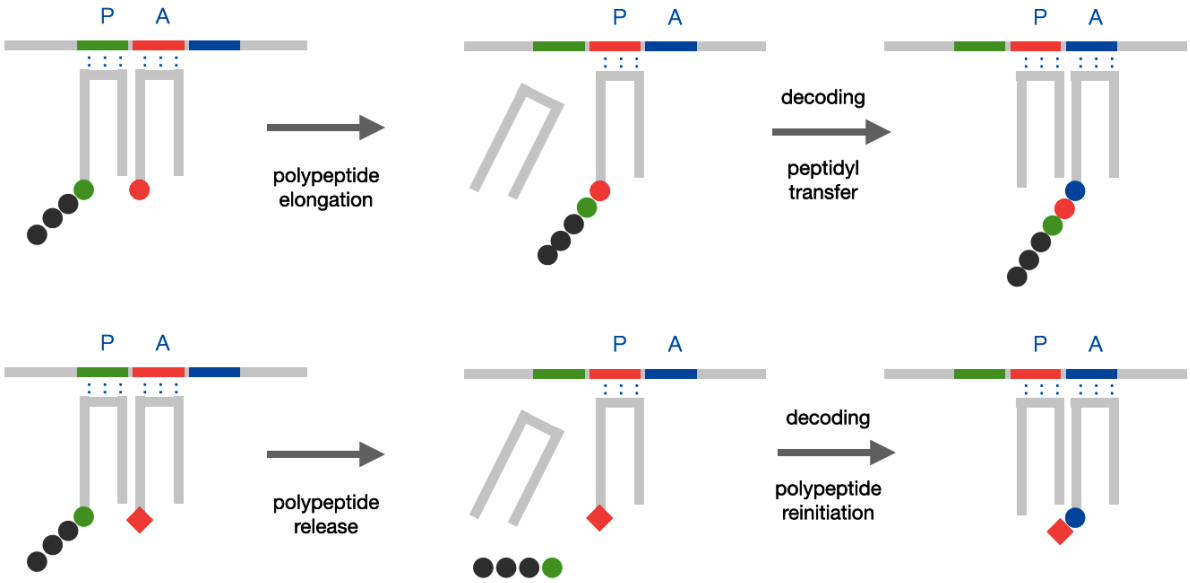


FIGURE 15-4. POLYPEPTIDE REINITIATION COUPLED TO POLYPEPTIDE TERMINATION

Polypeptide reinitiation could be coupled to polypeptide termination in polycistronic mRNAs using an aminoacyl-tRNA with blocked α-amine, as in bacterial formyl-methionine initiator tRNA. In Figure 15-4 we compare two successive steps of polypeptide elongation with the corresponding two steps of polypeptide termination and reinitiation, where the red circle depicts an unblocked amino acid, and the red diamond depicts an amino acid with blocked α-amine. Terminating one polypeptide to initiate the next, one and the same mRNA triplet first acts as a polypeptide STOP-codon read at the A-site by the blocked (aka reinitiator) tRNA, and then as a polypeptide START-codon when translocated to the P-site. Table 15-1 summarizes the emergent roles of RNA adaptors in polynucleotide duplication and polypeptide translation.

	■■ dRNA loaded	tRNA empty	tRNA ● charged	tRNA ◆ blocked
POLYNUCLEOTIDE				
initiation	(1)			
elongation	(1)			
termination		(2)	(3)	
POLYPEPTIDE				
initiation			(4)	(5)
elongation			(4)	
termination		(4)		(5)

TABLE 15-1. RNA ADAPTORS IN POLYNUCLEOTIDE DUPLICATION & POLYPEPTIDE TRANSLATION

We conjectured that translocation in polynucleotide elongation was driven by dRNAs acting as thermal motors, notably the cold closing of freed dRNA in the P-site after nucleotidyl transfer. Locked in an open conformation, L-shaped tRNAs were weak thermal motors at best. Absent the work of hairpin closing, a new source of Gibbs energy was available to drive translocation in polypeptide elongation, *viz*. the greater chemical stability of the acylamide (aka peptide) than acylester bond (Krayevsky & Kukhanova 1979; Leung et al 2011). Decoding and accommodation of open tRNAs, as well as peptidyl transfer itself, were relatively fast, so once a polypeptide was initiated, multiple peptide bonds could be made isothermally in a short while compared to the diurnal hot-cold cycle of polynucleotide elongation. *If so, the breakout ribosome née terminating duplisome had two modes of elongation, slow and careful polynucleotide duplication using duplicon-dRNAs, or quick and dirty polypeptide translation using aminoacyl-tRNAs.*

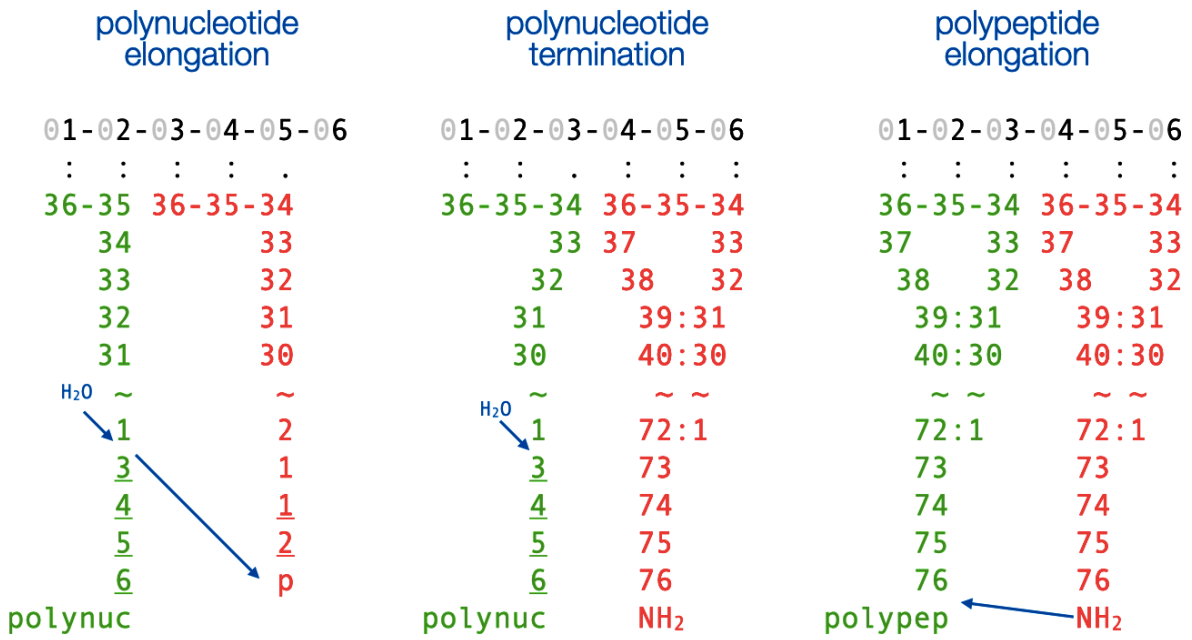


FIGURE 15-5. CHANGES IN DECODING & TRANSFER FROM DUPLISOME TO RIBOSOME

The breakout ribosome had one brilliant new trick, polynucleotide-directed polypeptide synthesis, borrowed wholesale from the terminating duplisome. Beyond the transfer reaction itself, the most striking change from polynucleotide to polypeptide elongation was the increased template movement from two to three nucleotides in each elongation cycle. We suggest this larger translocon followed directly from the +1 shift of reading frame introduced when a terminator tRNA preempts the elongator dRNA for the A-site (Figures 13-8 &15-5). Like the reading frame itself, translocon size was no doubt sloppy at first. Over time the advantages of longer polypeptides with greater variety and more accurate placement of amino acids, selected for reliable decoding and translocation of non-overlapping triplet codons.

Examining features of the ribosome and tRNAs that define and maintain the reading frame, we can distinguish ones that may trace to polypeptide life, from ones that arose later in protein life (Figure 15-6). In the A-site, modifications of tN37 promote cross-strand stacking. In the P-site, modifications of tN34 defend the wobble pair against encroachment or scrunch from the A-site tRNA. Finally, conserved modifications in the

small subunit rRNA itself formed the wobble seal.[22] These several tRNA and rRNA modifications require enzymes today, but some may trace to ribozymes in polypeptide life. Additional improvements in reading frame defense continued in protein life (section 16).

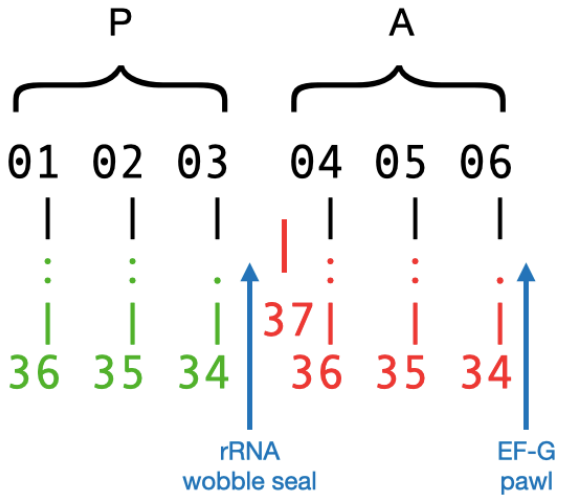


FIGURE 15-6. DEFENDING THE READING FRAME {?MOVE PAWL TO FIRST/SECOND PAIRS?}

No matter how useful the new technology of coded polypeptides, this early stage of protein life, which we have dubbed *polypeptide life*, relied on continued RNA duplication. In the parlance of synthetic biology, RNA duplication and protein translation were orthogonal functions.[23] During their long *conviviencia*, the speed, accuracy, and versatility of polypeptide translation greatly improved without compromising the well-oiled machinery of polynucleotide duplication. There is a continuum of possibilities for orthogonal evolution of ribosomes and duplisomes. At one pole, ribosomes diverged early and completely from duplisomes. At the other pole, one new factor arose to switch the common core from polynucleotide elongation to the new mode of polypeptide elongation. This regulatory factor might act irreversibly during ribosome biogenesis, or bind reversibly to the mature duplisome to switch it

[22] RNA modifications … wobble seal … yeast …. $m^1acp^3\Psi$1191 1-methyl-3-alpha-amino-alpha-carboxyl-propyl pseudouridine in hairpin loop of h31 (refs).

[23] There are several examples from synthetic biology of orthogonal ribosomes (Orelle et al 2015; Schmied et al 2018; Carlson et al 2019; Aleksashin et al 2020).

from polynucleotide to polypeptide elongation, and back. Obvious possibilities are an rRNA modification, a riboswitch, or a polypeptide.

5S RNP in the central protuberance just above the peptidyl transfer center, between the L1 stalk and factor binding site, has no catalytic role in the mature ribosome, and has been lost entirely in mitochondrial ribosomes (Koripella et al 2020). 5S RNP docks midway during large subunit biogenesis and acts as a wrench-like chaperone for rRNA domain D5 (Zhou et al 2019; Micic et al 2020). At the nucleolar stress response checkpoint in eukarya, accumulation of 5S RNP due to problems in large subunit maturation sequesters E3 ubiquitin ligase MDM2, allowing TP53 mediated cell cycle arrest and apoptosis (Bohnsack & Bohnsack 2019; Castillo Duque de Estrada et al 2023).

For bold discussion, we conjecture that 5S rRNA (~120 nt) arose in polypeptide life to toggle biogenesis of the polymer transfer center from polynucleotide to polypeptide elongation. Two curious observations may be pertinent: First, antibiotics that bind domains D2 and D5 of the large subunit rRNA rescue peptide bond formation in the absence of 5S rRNA (Khaitovich & Mankin 1999). To work, these drugs must be present during large subunit biogenesis, but can then be washed out with ethanol without abolishing activity of the folded subunit. Second, in engineered ribosomes with circularly permuted 5S rRNA fused to 23S rRNA, the large subunit adopts either a normal fold that supports translation, or a novel fold that shifts H89 and trailing nucleotides 2490-2505 as much as 30 Å (Huang et al 2020). In this shift, C2498 C2499 U2500 pair with the H80 P-loop G2251 G2252 G2253, precluding its pairing with the 3’ CCA extension of the P-site tRNA for peptidyl transfer.

Improvements in polypeptide translation from the breakout ribosome entailed a small number of changes in tRNAs and rRNAs, in no necessary order, tuning the system from polynucleotide termination to the new function of polypeptide elongation. These changes included: (1) gating the ribosome P-site to admit tRNAs; (2) adjusting the polymer transfer center to position and orient the 3’ ends of both tRNAs for peptidyl transfer; and (3) lubricating the nascent polymer exit tunnel for peptide bond formation, and polypeptide exit. The exit tunnel at the breakout of polypeptide translation was likely shorter and wider than the tunnel of modern ribosomes (Fritch et

al 2018; Dao Duc et al 2019). The new polymer and its transfer reaction changed the mechanical forces of elongation. and the exit tunnel now had to accommodate polypeptides of simple sequences and compositions.

Both immediate advantages of coded polypeptides, and emergent advantages of self-folding protein domains with catalytic centers, selected for a greater variety of amino acids and greater control of mRNA translation. Alongside tuning their interactions with the common machinery of translation, the working set of tRNAs expanded beyond the original polynucleotide termination code. On one hand, diversification of tRNAs and their charging complexes enlarged the portfolio of proteinogenic amino acids, even as individual complexes became more selective about their chargeable amino acids. On the other hand, diversification of tRNAs and improvements of the ribosome decoding center, expanded to the total coverage of codons.

The requirements of decoding were different for polypeptide elongation than for either polynucleotide elongation, or polynucleotide termination. Polynucleotide elongation required a complete and balanced set of dRNAs to faithfully duplicate any RNA template. Processivity was crucial, but only moderate fidelity was needed. The solution was a working set of 16 dRNA isoacceptors using the third codon position for stability, not sequence information. Occasional near-cognate duplicons or one nucleotide frameshifts introduced single nucleotide substitutions or indels, respectively, in otherwise full-length duplicates. Codon-directed polynucleotide termination had to be more accurate, and RIBOSTOP-codons comparatively rare, to produce full-length duplicates without premature terminations. The solution was a small number of tRNAs with triplet codons selected for accurate decoding. In short, all even-frame doublets were read by 16 cognate elongator dRNAs, but only a few of the 64 possible odd-frame triplets were preempted by some terminator tRNA.

There were two great tensions in the evolution of protein translation, one toward longer polypeptides, and the other toward more careful selection of residues. For longer polypeptides, the ribosome née duplisome had to read all triplets in the ORF without stalling or hesitation. There are two ways, not mutually exclusive, to achieve this: First, restricting mRNAs to a small set of sense codons such that the sequence

space of ORFs was a proper subset of the full ribogene sequence space. Two obvious suggestions are that early ORFs used only Gs and Cs, at most 8 codons, or that these were restricted to homopolymer and dinucleotide repeats.[24] Second, expanding the coverage of aminoacyl-tRNAs to more sense codons, either by relaxing the codon specificity of existing polynucleotide terminators, or by introducing entirely new polypeptide elongators (cf. Lehman & Jukes 1988).

The problem of longer polypeptides was ultimately solved by some combination of expanding the repertoire of elongator née terminator tRNAs from singlet to quartet codon boxes, and introducing entirely new elongator tRNAs, to cover hitherto unreadable triplets with no cognate aminoacyl-tRNA. Whereas the codon preferences of tRNAs used in polynucleotide termination had to be unique, third position degeneracy was allowed in the new role of polypeptide elongator tRNAs. Here it was important that all codons were readable, not that they were completely discriminated from one another. *As a problem of orthogonal evolution, in this piecemeal expansion of the genetic code, relaxing old tRNAs or adding new ones, it was still important not to prematurely terminate polynucleotides.*

With the polypeptide length problem effectively solved, there was selection to increase the variety of amino acids, and the selectivity of tRNA charging ribozymes. If elongator tRNAs had been under selection to generalize their codon boxes, with the table of codons now mostly assigned, they were now under selection to specialize them from quartets to duets to singlets. This entailed both modifications of tRNA anticodon arms, and changes in the ribosome, to improve the fidelity and variety of decoding, as well as to defend the reading frame. In cellular life, tRNA decoding and reading frame defense require enzymatic modifications of tRNAs and rRNAs. For example, splitting quartets into duets requires 5-methylation or other modification of U34 to suppress superwobble.

Exploiting the coding principle in polypeptide translation entailed concurrent piecemeal changes in the genetic code, introducing new amino acids, discriminating between hitherto interchangeable residues by giving them separate tRNAs and

[24] Today the four homopolymers encode homopolymers of G F K P while the six dinucleotide repeats [ACACAC]n [AGAGAG]n [AUAUAU]n [CGCGCG]n [CUCUCU]n [[GUGUGU]n include codons for A C E H I L R S T V Y.

charging ribozymes, and culling some hitherto proteinogenic amino acids entirely. The first aaRS enzymes, more so the ribozymes that preceded them, were no doubt promiscuous (Weber & Miller 1981). Changes in anticodon arms and decoding center allowed split codon boxes to squeeze in new amino acids, or give former roommates their own tRNAs. It is possible that some extant subfold fragments of proteins trace to a pre-modern genetic code with fewer tRNAs and more promiscuous charging ribozymes or enzymes (ref).

Early molecular biologists pondered when the canonical genetic code was established, and how malleable it remains today. The actual numbers of tRNA genes vary widely among species. There are some 58 tRNA genes in the core genome of *Escherichia coli*. After removing near duplicates (isodecoders), there are 43?46 distinct tRNAs (isoacceptors), each charged with exactly one of 22 proteinogenic amino acids (isotypes). Thus, a complete working set with one of each tRNA isoacceptors comprises well over 3000 nucleotides, comparable to the rRNA content of one ribosome particle. In the streamlined genetic code of eukaryotic mitodchondria, the number of tRNAs has been cut in half to 24 in *Sacchromyces cerevisiae* and 22 in *Homo sapiens*. Although there are many ribosome particles in the bacterial cell, and many equivalents of the working set, they are present in roughly 1:1 ratio, not say a 10 or 100 fold or greater excess of tRNA sets to ribosomes that we might expect for the ratio of small substrates to their large catalytic centers.

Under purely VGT, the question of when any two innovations arose along the same or different branches of a lineage has a clear and simple answer. Based on this assumption, the canonical genetic code was fully formed in LUCA. The questions of code evolution are more nuanced today: First, the idea that the genetic code was settled before prokarya split into the kingdoms of bacteria and archaea has been questioned. Second, different sets of tRNAs can realize the same genetic code, or map from codons to amino acids, that is, the code itself is more durable than its particular implementations.

On evolutionary timescales, all species undergo extensive HGT, both within the relatively brief window of speciation itself, and the longer periods after clear phyletic split. Frequent introductions of genes from near relatives, and less frequent

introductions from distant ones, give rise to selective sweeps and fixation, or more often, rejection of the foreign gene. Besides introducing new genes, HGT may displace existing ones in what are called non-orthologous gene displacements (NOGDs).

It has been proposed that many rival innovations in protein translation were yet unsorted at the split of bacteria and archaea, and continued with robust HGT across the nascent cellular kingdoms. Like the original guardian tRNAs, polypeptide translation using aminoacyl-tRNAs was a double-edged sword, allowing cells to import useful mRNAs or invite potential parasites. In the short run, there was immediate selection for divergence of genetic codes, or at least codon frequencies, to defend against MGEs and more organized parasites. But in the long run, there was selection for convergence of genetic codes to explore opportunities of one canonical protein code as the *lingua franca* of HGT (Vetsigian et al 2006; Koonin 2017). All in all, the benefits outweighed the risks, and successful lineages retained or converged on the canonical genetic code.

Whereas the coopetitive structure of evolutionary search itself favored a common code of protein life, the canonical code settled upon reflected additional factors of thermodynamics and kinetics of decoding, error minimization, pathway co-evolution, and no doubt accident (Grosjean et al 2010; Koonin 2017). Since the discovery of the genetic code, there have been earnest attempts to reconstruct its evolutionary history, using chemical principles, comparative genomics and combinatoric arguments (Ribas de Pouplana 2002). Skepticism of exact histories seems warranted until there is more agreement on both the *explananda*, or things that need to be explained, and the *explanans*, or pertinent principles and evidence. For the record, we trust that someday a more persuasive history of the genetic code may be reconstructed, including the additions and losses of tRNAs, changing specificities of charging ribozymes and later enzymes, and sculpting of the codon preferences.

The effect of frame shift during elongation is very different for polynucleotide duplication and polypeptide translation. Before terminator tRNAs, the impact of -1 or +1 frameshift on nascent polynucleotides was strictly local. Although different sets of elongator dRNAs are used in the original (even) and shifted (odd) frames, the products are identical, excepting the 1 nucleotide indel at the site of shifting. But with the advent

of terminator tRNAs, the frameshift would eliminate any downstream RIBOSTOP-codons and might introduce new ones (section 13). The impact of -1 or +1 frameshift on polypeptide translation is if anything, more catastrophic: not only does translation generally terminate prematurely, but the translated downstream sequence has no obvious similarity to that of the 0 frame. However, for simple repeats, or other low complexity polypeptides at the breakout of translation, it is possible that frameshift products were useful repeats themselves, or that random shifting between frames produced a spectrum of polypeptides of useful compositions and sequence.

The full value of faithful maintenance of reading frame emerged *pari passu* with the refinements of tRNA charging and decoding. With a half dozen or more tRNAs, and ORFs of a dozen or more residues, the search space of polypeptide sequences to explore was much larger, and more interesting. Whereas globular proteins have few unnecessary cavities, folded RNAs have large “breathing spaces” that are often filled with polypeptides in natural RNPs. Thus, the first coded polypeptides likely included intrinsically disordered sequences that folded with the ribosome and other ribozymes to form more stable or active RNPs. As ORFs got longer, and the positional constraints on sidechains at key positions became greater, the rewards drove biosynthesis and charging, refinement and fidelity of decoding, and defense of reading frame. Finally, we notice that a frameshift error during polypeptide translation spoils the nascent protein, but a frameshift error during gene replication, or mRNA transcription spoils any protein made from them. Thus, the improvements in protein translation, and increased cost of ribosome frameshift errors, selected indirectly for better frame preservation in RNA copying, and later, DNA replication.

Molecular biologists have used tRNA evolution as a proxy for the evolution of protein translation itself. Here we suggest the primitive tRNA underwent a number of structural changes along the evolutionary path from polynucleotide termination to polypeptide translation. Some were heritable sequence changes, while others were regular (nonheritable) modifications, mediated by ribozymes or enzymes of tRNA biogenesis and repair. *Pari passu* with the changes in tRNA structure, there were corresponding changes in the ribozymes, and later enzymes of tRNA charging, as well as ribosomal centers of decoding and peptidyl transfer. Meanwhile, the family of tRNAs

radiated from just one terminator tRNA to some twenty chargeable isotypes and forty isoacceptors.

For each innovation in tRNA structure and charging, we seek a parsimonious explanation of when, how, and why that change arose, consistent with our explanations of other features. Here we arrange five events in likely order: (1) D and T loops specialized to form the elbow, distinguishing constituitive terminator tRNA from tandem intermediates in dRNA duplication (section 13). In the ribosome and charging ribozymes, a number of tRNA-RNA interactions evolved to recognize this elbow. Thus, the elbow became an important recognition element at all three tRNA sites in the ribosome large subunit, interacting with A-site finger helix H38, P-site helix H84, and E-site IDTM in the L1 stalk. Meanwhile, in remarkable examples of convergent evolution, both ribozymes P and T acquired an IDTM to recognize the tRNA elbow (sections 13-14). (2) Extending and charging the tRNA 3’ end, the tRNA tailing ribozyme and various acyl ligase ribozymes T created conditional terminator tRNAs (section 14). (3) Insertion of the canonical group I intron in the anticodon arm facilitated shuffling of tRNA halves and the radiation of conditional tRNAs (section 14). (4) Using extant enzymes, and perhaps, extinct ribozymes, sundry tRNA modifications helped to stabilize the L-shaped fold, split codon boxes, and defend reading frame. (5) Aminoacyl ligase ribozymes T were displaced by tRNA charging enzymes, and the common tRNA tailing ribozyme L was displaced by enzymatic addition of the universal 3’ CCA tail (section 16).

16. Protein-supported protein translation

Following the breakout of polypeptide translation, its speed and fidelity improved through immediate changes in rRNAs, tRNAs, and other ribozymes or riboswitches (section 15). More gradually, and more profoundly, coded polypeptides and self-folding proteins invaded all aspects of bioenergetics and metabolism, including RNA replication and protein translation. In this evolutionary bootstrapping, proteins became catalysts and regulators of their own biosynthesis. Many of the later improvements in translation entailed ribosome and tRNA biogenesis factors and their RNA modification enzymes, tRNA charging enzymes, as well as ribosomal proteins and translation factors. In another remarkable transition, enzymes invented to improve feedstocks of random oligonucleotides were exapted as template-directed polymerases, replacing spontaneous or ribozymatic RNA copying with enzymatic replication and new levels of regulation (section 17).

To focus on the breakout of polypeptide translation, we stipulated feedstocks of amino acids and random polypeptides in the late RNA world without regard to their composition and synthesis. Such agnosticism made sense in so much as our models of tRNA charging (section 14) and codon-directed polypeptide elongation (section 15) were based on ribozymes T and the duplisome, not conjectural peptidyl ligase ribozymes, nor unknown prebiotic mechanisms of peptide condensation (Lambert 2008).

Random polypeptides form extended chains, or adopt useful secondary structures, such as α-helix, β-strand, or reverse turn, through co-assembly with other polypeptides, polynucleotides or membranes. After the breakout of polypeptide translation, coded polypeptides emerged that augmented and then supplanted these random polypeptides. Some of the coded sequences were poised to adopt supersecondary topologies, such as α- and β-hairpins, β-meanders, or $\beta\alpha\beta$-elements. It is thought that self-folding protein domains arose through duplication or combination of such poised sequences whose hydrophobic residues packed against one another in a globular core (Alva et al 2015; nobel prize 2024). In one narrative, a sampler of the evolutionary progression from extended polypeptides to poised elements to self-

folding proteins is found frozen in the ribosome as one moves from the older core to the younger periphery of this great RNP (Kovacs et al 2017; Lupas & Alva 2017).

For least assumptions about prebiotic chemistry, and most parsimony with polymer life as we know it, we have based processive RNA copying on oligoribonucleotides, the material *sine qua non* of RNA life, and sunlight, arguably the most certain and universal source of Gibbs energy in any conjectural RNA world (sections 8-12). Notwithstanding nucleotide-derived enzyme cofactors, or artificial polymerase ribozymes that utilize nucleotide triphosphates, there is no clear evidence that NTPs, not just oligonucleotides, trace to the RNA world, much less were used in polynucleotide synthesis (Liu et al 2020). For bold discussion, we propose that NDPs and NTPs first appeared in significant concentrations in protein life, made by ancestors of their extant enzymatic pathways. Before we put this case, we notice our position is defeasible, and plausible prebiotic sources of NDPs and higher polyphosphates have been proposed, *viz*. nucleoside 5' tetraphosphates from NMP and trimetaphosphate, or ATP from ADP and acetylphosphate (Lohrmann & Orgel 1973; Pinna et al 2022).

Whatever the environmental sources of Gibbs energy in RNA life, and their domesticated intermediates, metabolism in protein life consolidated around a common currency of NTPs as Gibbs intermediates (Vetter & Wittinghofer 1999). A panoply of new enzymes used this currency to drive reactions uphill, impose direction on otherwise reversible reactions, or choose between alternative states or products. Enzymatic reactions that use NTPs can be classed according to which phosphate (α, β, γ) provides the electrophile phosphorus atom, and which group (inorganic, nucleotide) transfers with the other leaving.

RNA polymerases are arguably the most important enzymes in the evolution of polymer life (section 17). These, and the related *nucleotidyl transferases*, catalyze the attack of an alcohol or water on the NTP α-phosphate to transfer or release NMP, with inorganic pyrophosphate leaving. Prompt hydrolysis of pyrophosphate makes these reactions nearly irreversible *in vivo*. RNA polymerases likely arose when the pools of activated nucleotides were predominantly NDPs, not NTPs. Indeed some template-directed DdRPs and DdDPs that normally utilize (d)NTPs can also use (d)NDPs as

substrates, albeit with higher K_m and lower V_{max} (Gottesman & Mustaev 2019). For comparison with RNA exonucleases, we show the RNA polymerase reaction as depolymerization (pyrophosphorylation), reverse its usual direction in cellular life (Table 16-1).

Enzyme	Reaction	Folds
hydrolase	... NNN OH + H_2O → ... NN OH + NMP	RNase II/RNB
phosphorylase	... NNN OH + p_i ⇌ ... NN OH + NDP	RNase PH
pyrophosphorylase	... NNN OH + pp_i ⇌ ... NN OH + NTP	Polβ-like, 2xDPBB, RRM-Palm
pyrophosphatase	$pp_i + H_2O \rightarrow p_i + p_i$	

TABLE 16-1. FROM 3'-to-5' EXONUCLEASE TO 5'-to-3' POLYMERASE

The family of *polynucleotide phosphorylase* enzymes have NDPs as their natural substrates or products (Grunberg-Manago et al 1956). Unlike hydrolytic exonucleases that degrade RNA to NMPs, these phosphorolytic exonucleases degrade RNA to NDPs from 3' to 5' by attack of an oxyanion of inorganic phosphate on the terminal phosphodiester bond, with O3' leaving. The reaction is reversible, *viz.* under high concentrations of inorganic phosphate and low NDPs, polynucleotides shorten with release of NDPs, while under low phosphate and high NDP, polynucleotides lengthen with release of phosphate. These phosphorolytic exonucleases form the catalytic core of ancient machinery for regulated maturation and degradation of RNAs from 3' to 5' known as the RNA degradosome in bacteria and eukaryotic mitochondria, or the RNA exosome in archaea and eukaryotic nucleus (Mitchell et al 1997; Viegas et al 2020).

Other than oligoribonucleotides themselves, the environmental sources and domesticated forms of phosphate in the RNA world are unclear (Pasek 2020; Nicholls et al 2023). We speculate that the emergence of polynucleotide phosphorylase was part of a general shift in metabolism, from RNA life with inorganic phosphate limiting for

polynucleotide synthesis, to protein life with phosphate in excess of the backbone needs of ribogenes and ribozymes. Overcoming phosphate insecurity, this inorganic anion was now available for intermediary metabolism, membrane lipids, and signaling, as well as the new energy currencies of NDPs and NTPs. In this new phosphate-intensive cellular metabolism, polynucleotide phosphorylase built up NDP stores when phosphate was abundant, reverting to random oligonucleotides when it was scarce.

Like polynucleotide phosphorylase, we suggest that primitive RNA polymerase enzymes utilized NDPs as substrates. Their success allowed later improvements, including mixed NDP/NTP substrate pools, and eventually, exclusively NTP pools, with hydrolysis of pyrophosphate to drive irreversible transfer. In protein life, at least three families of polymerase enzymes evolved to transfer NDPs (or NTPs) to the oligonucleotide 3' OH, with the release of inorganic phosphate (or pyrophosphate) (Koonin et al 2020a). These were (1) Polβ-type polymerases, (2) DPBB-type polymerases with two double-psi beta-barrel domains contained in one subunit or two separate subunits, and (3) RRM-type polymerases with an RNA recognition motif (aka Palm domain). Template-directed polymerases arose within each family, but the first RNA polymerases likely had little nucleotide specificity, and no template requirement.

We conjecture that a promiscuous terminal nucleotidyl transferase enzyme took over the provisioning of random oligonucleotides, bypassing and replacing older spontaneous or ribozymatic reactions. In particular, this enzyme supplied the random oligonucleotides added as dRNA leaders in the polynucleotide elongation cycle by phosphoryl ligase ribozyme P, and added as tRNA trailers in the polypeptide elongation cycle by phosphoryl ligase ribozyme L.[25] This primitive nucleotidyl transferase was likely a founding member of the Polβ superfamily that includes various terminal nucleotidyl transferases (TENTs), or tailing enzymes, *viz*. mRNA poly(A) polymerase, tRNA CCA tailing enzyme, and DNA terminal transferase.

Despite their ancient radiation, the tRNA family is constrained by common interactions with biogenesis and charging factors, as well as ribosomes and translation factors. Beyond purifying selection to slow their drift and divergence, common positive

[25] In section 12 describing these conjectural reactions, we propose that RNase P RNA descends from ribozyme P, while the ribosome peptidyl transfer center has a common ancestor with ribozyme L.

selection on tRNAs drove concerted changes and convergence. Here we propose a parsimonious explanation of when, how and why the $_{74}$CCA$_{76}$ tail at the 3' end of tRNAs arose. This tail provides a universal handle on tRNAs for amino acid charging and peptidyl transfer. In some bacteria, including *Eschericia coli*, the terminal CCA is transcribed from the gene itself, followed by maturation of the transcribed trailer by a 3' to 5' exonuclease enzyme such as RNase PH (refs). In other bacteria, and all archaea and eukarya, the transcribed trailer has no particular sequence, but is cleaved precisely at $_{73}$NpN$_{74}$ by RNase Z, an endonuclease enzyme with metallo-β-lactamase fold, and then CCA is added to the 3' OH of N73 by the CTP/ATP nucleotidyl transferase enzyme (ref). Whether CCA is transcribed or added later during tRNA biogenesis, RNase Z and CCA adding enzymes repair mature tRNAs that have lost one or more nucleotides from the 3' end. *Parsimony suggests that enzymatic biogenesis was the primitive state in LUCA, while genomic transmission and transcription of the CCA tail found in some bacteria is derived.* Biogenesis and repair of the tRNA 3' tail by common enzymes solved the problem of drift in length and sequence of this crucial handle used for charging multiple isoacceptors from one aaRS enzyme, and more generally, positioning any two tRNA acceptor arms in the common ribosome peptidyl transfer center.

The tRNA CCA tail interacts with three different RNAs of ancient origin, suggesting it might date to the RNA world itself, or was even present on the original hairpin that formed the 3' half of the primitive tRNA (*pace* Di Giulio 1992). In bacteria, the CCA tail of pre-tRNAs basepairs with G292 G293 in loop L15 of the S domain of RNase P RNA (Reiter et al 2010). In what has been called an RNA ruler, this interaction flays the pre-tRNA 3' NCCA trailer from its 5' leader, reinforcing the IDTM anchor of the tRNA elbow that roughly positions $_{1}$NpN$_{1}$ at the catalytic site. In archaea and eukarya, pre-tRNAs undergo leader removal before RNase Z cleavage and enzymatic addition of the 3' CCA. Consistent with this, the S-domain of their RNase P RNA has only the elbow-recognition IDTM, not the CCA-recognition sequence. Thus, *CCA recognition by bacterial RNase P RNA likely arose after the encoding of CCA in DNA life.*

The CCA tails of mature tRNAs basepair with the discriminators of T-box riboswitches in bacterial mRNAs to regulate their transcription or translation. Whereas the 5' recognition domains of these riboswitches likely descend from the charging

ribozymes T of primitive tRNAs, their 3' discriminator domains may be younger than translation itself (section 14). *Thus, on the evidence of RNase P and T-box riboswitches, the tRNA CCA trailer might either predate or postdate aminoacylation of primitive tRNAs.*

Interactions of the tRNA CCA tail with the A- and P-loops of the ribosome peptidyl transfer center provide no stronger case for its antiquity.[26] In one account of ribosome evolution, these loops in the outer shell of the peptidyl transfer center are homologous elements, dating back to a tandem duplication event that formed the protoribosome (ref). Alternatively, they are convergent elements selected by interactions with the tail common to both tRNA substrates. In light of our duplisome hypothesis, the A- and P-sites of the nucleotidyl transfer center predate the tRNA and its CCA tail. If so, the A- and P-loop interactions with the tRNA tail evolved piecemeal anytime onward from the terminating duplisome up to the late ribosome. These interactions better positioned the electrophile (carbonyl carbon) of the peptidyl-tRNA and the nucleophile (amino nitrogen) of the aminoacyl-tRNA for peptide bond formation or peptide release, without compromising the transfer reaction of RNA duplication. *Whenever they arose, we infer these short loops are not structural homologs, but functional analogs from convergent evolution.*

After pondering arguments for antiquity of the tRNA tail, we are not convinced that this universal CCA predates protein life. Its base-pairing interaction with bacterial RNase P RNA is likely a derived, not primitive character. Similar interactions T-box riboswitches and ribosome peptidyl transfer center may be derived as well. Even if we conjecture that this feature arose early in tRNA evolution, absent some mechanism of site-specific addition in tRNA biogenesis, or site-specific recombination in tRNA copying, we doubt that purifying selection alone could maintain a common CCA tail against drift in length and sequence.

If the various tRNA-RNA interactions do not preclude the comparatively late acquisition of the CCA tail, and if early acquisition introduces the problem of perservation, the most parsimonious origin is a nucleotidyl transferase of the Polβ

[26] Helix H80 has a stem of 3 Watson-Crick pairs topped by a 7 nucleotide P-loop; H92 has a stem of 5 Watson-Crick pairs topped by a 5 nucleotide A-loop (refs Gregory et al 1994; Green et al 1998).

superfamily, ancestral to the tRNA CCA adding enzyme. If so, the tRNA tail arose in protein life using NDPs or NTPs as substrates. However, the question of some extinct tailing ribozyme, predating this extant tRNA tailing enzyme, is more nuanced. In section 14 we introduced a tailing ligase ribozyme L as part of the tRNA charging cycle, but placed no particular constraint on tail sequence, and only a minimum requirement for its length, as the mature lengths of their tails were determined by the tRNA-specific ligand-specific ribozymes T. There we suggested that the ribogenes of aminoacyl tRNAs used as polynucleotide terminators ended with tRNA N76, while their common tailing and specific charging ribozyme cleaved at $_{73}$NpN$_{74}$ and $_{76}$NpN$_{77}$, respectively (see Figure 14-1).

No longer added and cleaved in the ribozymatic charging cycle, tRNA tails added enzymatically in tRNA biogenesis and repair, became more important than ever for enzymatic charging, and substrate positioning in the ribosome peptidyl transfer center. The tailing enzyme likely afforded greater control of length and sequence of tail: (1) to adapt the acceptor end to the new tRNA charging enzymes, (2) to prevent interference between polynucleotide duplication and polypeptide translation, and (3) to assist peptide bond formation and peptide release by substrate positioning and proton shuttle. For sake of discussion, we suggest that the original tailing enzyme added four (or more) cytosines to tRNA N73, *viz*. the default product of the tailing ribozyme L in absence of the oligonucleotide co-substate. If this enzyme predated tRNA charging enzymes, then enzymatically tailed tRNAs, like ribozymatically tailed tRNAs, were still charged by ribozymes T. Originally sloppy in its product, by LUCA the tailing enzyme had settled on the fixed length and sequence CCA, *pari passu* with tuning of the recognition elements in RNase P, T-box riboswitches, ribosome peptidyl transder center, tRNA charging enzymes and EF-Tu that interact with it.

Following the early expansion of the genetic code (sections 14-15), tRNA charging ribozymes were supplemented, and then supplanted piecemeal, by charging enzymes, probably because proteins could discriminate between hitherto interchangeable amino acids. That is, the conjectural ribozymes T were better suited to reading one anticodon, than selecting one amino acid. Meanwhile, one enzyme could recognize and charge several tRNA isoacceptors with different anticodons. For example, *Escherichia coli*

LeuRS charges five isoacceptors with anticodons GAG, CAG, UAG, CAA and UAA, reading codons CUY, CUG, CUR, UUG and UUR, respectively. During this handover from ribozymatic to enzymatic tRNA charging, some amino acids were added and others culled from the set of proteinogenic amino acids. Even after all ribozymes T for amino acids were replaced by charging enzymes, there were likely further duplications of aaRS enzymes and their tRNAs to introduce new amino acids or discriminate hitherto interchangeable ones.

All tRNA charging enzymes have a two-step mechanism to use Gibbs energy from ATP to drive aminoacylation: (1) the free amino acid forms aminoacyl-AMP with pyrophosphate leaving, and then (2) this intermediate is attacked by the 2'/3' OH of tA76 to form the aminoacyl-tRNA with AMP leaving. Like other nucleotidyl transferase enzymes, aaRSs may have worked with NDPs originally, and later used the Gibbs energy of pyrophosphate hydrolysis to drive the reaction irreversibly. Twenty-four different aaRS enzymes, comprising two unrelated catalytic folds, are found in cellular life today (Rubio Gomez & Ibba 2020). Befitting their great age, there is a remarkable diversity of subunit and domain organization, activities for proof-reading and editing the charged tRNA, as well as regulation of multi-synthetase complexes, and various moonlighting functions.

Molecular biologists were disappointed that the genetic code held no deep meaning like our explanation of the periodic table in terms of atomic orbitals (Koonin 2017). Whereas the *stereochemical fit theory* proved unsupported, and we suggest unnecessary, other explanations have fared better (Schimmel & Ribas de Pouplana 1995). In the *error minimization theory*, chemically similar amino acids are assigned near cognate codons, reducing harms of occasional errors in charging or decoding (Woese 1967; Knight et al 2004). In the *coevolution theory*, anabolically complex amino acids first made by extinct ribozymatic, or extant enzymatic pathways, infiltrated the genetic code through duplication and specialization of the tRNA and aaRS ribozyme or enzyme of their metabolic precursors (Wong 1975). In cellular metabolism, no fewer than eight proteinogenic amino acids have proteinogenic precursors: S→C/W, V→L, T →I, E→Q→H, D→N, F→Y, and even more have precursors that may have been culled from an earlier genetic code. Over time, changes to the genetic code became

increasingly difficult as protein life relied on one agreed translation of more and more mRNAs, so codon assignments that were originally path-dependent accidents became effectively frozen in place (Crick 1968; Vetsigian et al 2006).

Ligand	Ligase	Signal	Sensor	Readouts
carboxylic acid	ribozyme T	acyl tRNA	duplisome	↑ RNA termination ↓ RNA termination
L-α-amino acid	aaRS	aminoacyl tRNA	T-box riboswitch	↑ RNA termination ↓ protein initiation

TABLE 16-2. REGULATORY tRNAs FROM DUPLISOME LIFE TO CELLULAR LIFE

No longer needed for tRNA charging after the advent of aaRS enzymes, the family of ribozymes T were co-opted to monitor uncharged tRNAs as a proxy for amino acid depletion (Table 16-2). While their recognition domains became sensor domains of T-box riboswitches, the effector domain (acyl ligase ribozyme) was modified, or replaced entirely, to yield the new discriminators. In the former case, we may seek vestiges of its common ancestor with ribozyme P; in the later case, we may seek vestiges of another ancient riboswitch. Whereas acylated tRNAs made by ribozymes T or their deacylated forms were free signals to the duplisome for polynucleotide termination, and aminoacylated tRNAs were free substrates to the ribosome for polypeptide elongation, these modern mRNA signals to terminate transcription or initiate translation were deacylated tRNAs *complexed with their T-box riboswitches*.

Onward from the breakout of polypeptide translation, ribosomal proteins and elongation factors improved the speed and fidelity of elongation. In this great race of protein life, ribosomes became much faster than any other ribozyme or RNP, completing about 20 elongation cycles per second, *cf*. 10^6 times faster than conjectural diurnal duplisome (section 10). The main events in the ribosome elongation cycle are mRNA decoding and peptidyl transfer, ending in the PRE-translocation state, followed by translocation ending in the POST-translocation state, ready once again for decoding. During this cycle a new aminoacyl-tRNA enters the A-site, the deacylated tRNA exits

the E-site, the nascent polypeptide is lengthened by one residue, and it moves one residue through the exit tunnel. The remarkable speed and fidelity of polypeptide elongation, as well as its enormous burden on the cellular economy, trace to enzymatic consumption both of ATP in the pathways of amino acid synthesis and tRNA charging, and GTP within the ribosome itself.

In *kinase* reactions, the nucleophile is commonly the oxyanion of an alcohol or water, attacking the γ-phosphate of NTP for phosphoryl transfer or release, with NDP leaving. Whereas tRNA charging enzymes and RNA polymerases likely utilized NDPs at first, kinases absolutely require NTP substrates, and therefore, were *johnny-come-latelys* of protein life. In cellular metabolism, ATP promotes NMPs to NDPs to NTPs through various nucleotide kinases. As for ATP itself, ATP synthetase enzymes use proton gradients to promote ADP to ATP in respiration and photosynthesis. These membrane enzymes and their proton gradients were no doubt later developments, and likely some earlier form of substrate-level enzymatic phosphorylation used high-energy intermediates to promote ADP (and perhaps other NDPs) to nucleotide triphosphate as the energy currency of protein life.

Representing the kinase reaction class, two $\alpha\beta\alpha$ sandwich superfamilies, the P-loop and Rossmann folds, are among the largest, and most diverse protein superfamilies (Leipe et al 2003; Longo et al 2020). Both folds are dubbed nucleotide-binding domains because their substrates and co-factors include ATP and other phosphorylated ribonucleosides. It is unclear whether these folds diverged from a common $\beta\alpha\beta$ polypeptide in the big bang of protein life, or represent convergent evolution. There are at least two additional NTPase folds represented by the protein kinases and HSP90 superfamilies. [actin fold?]

P-loop ATPases include selective nucleotide monophosphate kinases that promote their (d)NMP substrate to its (d)NDP at the expense of ATP, the promiscuous nucleotide diphosphate kinase that promotes various (d)NDPs to (d)NTPs, ATP-dependent helicases such as RecA, ABC transporters, and the motor proteins kinesin and myosin.[27] These motor proteins use the Gibbs energy of ATP hydrolysis to pull cargo

[27]The term *P-loop* is used to describe either a polynucleotide loop in the ribosome peptidyl transfer center, or an entirely unrelated, polypeptide loop in NTPase enzymes.

along microtubules and actin filaments, respectively. Tubulin (Rossman-fold) and actin (actin-fold) themselves are ancient NTPases that use the Gibbs energy of GTP or ATP hydrolysis, respectively, to locally control their polymerization and depolymerization.

Kinesin and myosin are *powerstroke motors* that push forward an uphill step that immediately follows NTP hydrolysis. Cell biologists and biophysicists appeal to other biomechanical metaphors for ribosome-associated GTPases that pull forward an uphill step that immediately precedes GTP hydrolysis, preventing the process from rolling back downhill (Frank & Gonzales 2010; Hwang & Karplus 2019). There are three, more-or-less equivalent ways to think of them: First, as *Brownian ratchets* that use GTP hydrolysis to lock-in the POST-state of some thermal transition. Second, as *bistable switches* that enable some transition when flipped ON (aka GTP-bound state), but not when flipped OFF (GDP-bound state). Third, as *unidirectional catalysts* that accelerate some forward reaction in their GTP-bound state, but take themselves out of the equation for the reverse reaction by GTP hydrolysis, forcing a more circuitous return. Whatever metaphor we use, the GTPase must exchange its guanosine nucleotide before another transition.

An early, perhaps original role of the P-loop NTPase, ribosome-associated GTPases use the Gibbs energy of GTP hydrolysis to improve the speed and fidelity of ribosome decoding, translocation, and co-translational secretion (Leipe et al 2002, 2003; Maracci & Rodnina 2016). The deepest branches among GTPases are represented by translational GTPases and signal recognition particle GTPases, respectively. The family of translational GTPases has further specialized for regulation and catalysis of initiation, elongation, termination of translation, as well as ribosome recycling and rescue. All of the translational GTPases compete for a common factor binding site, or GTPase-associated center (GAC), at the interface of the ribosome subunits. In general, they dock in this factor binding site in their GTP-bound state to test for the presence or absence of some crucial feature of ribosome state with their extended C-terminal domains. Known as *gatekeepers*, these features include mRNA and tRNAs themselves, as well as a number of accessory factors. If those tests are satisfied, intersubunit rotation brings the conserved sarcin-ricin loop (SRL), nucleotides 2653-2667 of the LSU rRNA, into the enzyme catalytic center to activate GTP hydrolysis. Concomitant

with a change of gatekeepers, dissociation of the **GDP•GTPase** allows a new **GTP•GTPase** to enter the factor binding site, and meanwhile, the freed GTPase can undergo spontaneous or catalyzed guanosine-nucleotide exchange.

The high activation barrier (20-25 kcal / mol) for ribosome transition between the PRE- and POST-translocation states likely evolved *pari passu* with a primitive elongation factor as the means of overcoming it (Schilling-Bartetzko et al 1992). An early ribogene duplication gave rise to two translational GTPases, one specialized for the POST-translocation to PRE-transfer transition (bacteria EFTu / eukarya eEF1A), and the other for the PRE-translocation (aka POST-transfer) to POST-translocation transition (bacteria EFG / eukarya eEF2). Coupling decoding and translocation to the free energy of GTP hydrolysis, these GTPases lower the energy barrier between ribosome states, and augment the favorable Gibbs energy of peptide bond formation, ratcheting the elongation cycle in one direction (Table 16-3).

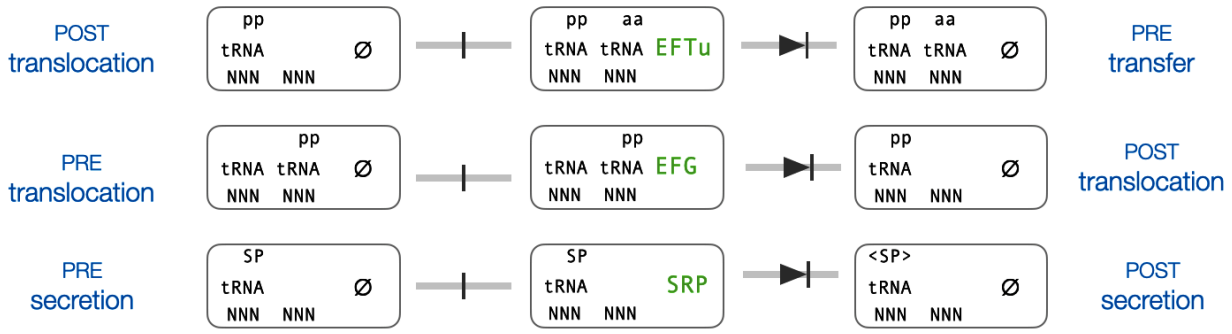


FIGURE 16-1. POLYPEPTIDE ELONGATION & SECRETION

In Figure 16-1 we show the gatekeepers and GTPases that regulate and drive polypeptide elongation and transmembrane secretion. In this, and similar figures (Figures 16-4 & 16-5), PRE-states and POST-states are shown on the left and the right, respectively. Association and dissociation events controlled by mass action are shown by a crossbar, while GTP hydrolysis is shown by an arrowhead. The thermodynamic problem of decoding is to frame the mRNA codon in the A-site, selecting a cognate aminoacyl-tRNA (**aa-tRNA**), while rejecting other tRNAs in the working set, based on the Gibbs energy of codon-anticodon pairing. The kinetic problem is to sample and test different candidate tRNAs quickly. By pre-assembling *ternary complexes* of **aa-**

tRNA with **GTP•EFTu** outside the ribosome, and queuing these candidates on the L12 stalk, repeated sampling became as fast and efficient as possible.

By separating decoding into two step, called *tRNA selection* and *proofreading*, EFTu allows minor groove inspection of the quality of codon-anticodon pairing in the A-site to be used twice. As misincorporation occurs only if both choices are wrong, such repeated choice, or *kinetic proofreading*, affords a nearly multiplicative improvement in decoding fidelity (Hopfield 1974; Ninio 1975). A a result, substitution errors of the bacterial ribosome are only ## per codon on average (ref). For any codon, an average of p^{-1} GTPs are consumed in proofreading, where p is the ratio of [cognate] / [cognate + near-cognate] tRNAs in the working set of ternary complexes. The lower bound on GTP consumption applies when the very first tRNA to undergo proofreading is accommodated as cognate, not rejected as near cognate.

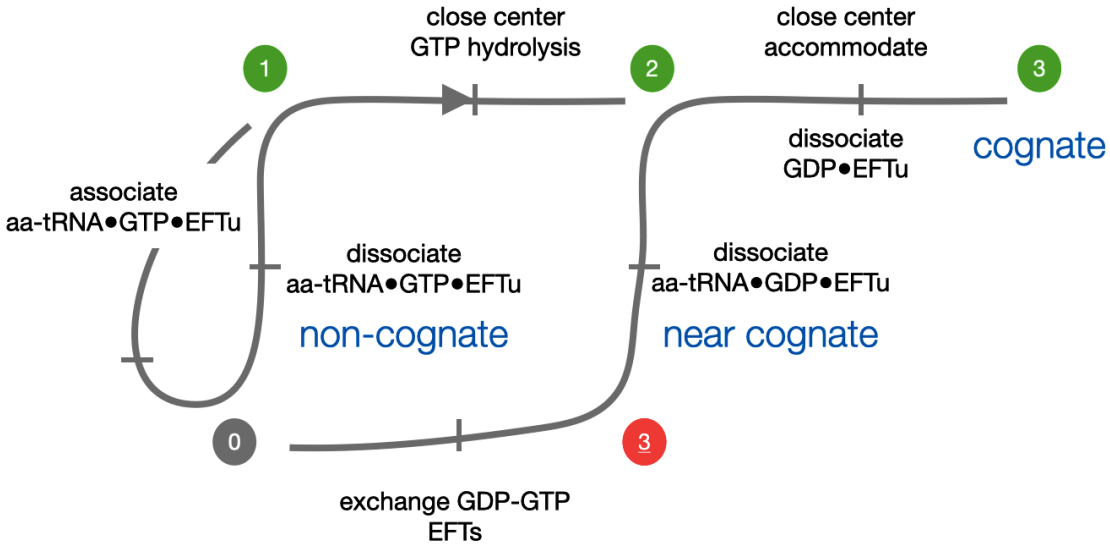


FIGURE 16-2. tRNA SELECTION & PROOFREADING FOR DECODING mRNA

In Figure 16-2 we show the principal runs of decoding as a place-transition diagram. In most non-cognate runs [0 1 0] a ternary complex **aa-tRNA•GTP•EFTu** briefly enters and then exits the decoding center without closure, or GTP hydrolysis. In most near-cognate runs [0 1 2 3 0] the ternary complex induces domain closure and GTP hydrolysis, but then is rejected upon proofreading, requiring the guanosine nucleotide exchange factor EFTs to swap out the GDP, and EFTu to reassemble with a

new GTP and aminoacyl-tRNA. In full cognate runs [0 1 2 3] the tRNA induces domain closure with GTP hydrolysis, and the **GDP•EFTu** undocks, allowing the second domain closure with accommodation of the **aa-tRNA** in the peptidyl transfer center. Peptide bond formation, viz. transition from this PRE-transfer state to the POST-transfer (aka PRE-translocation) state occurs rapidly.

The complete elongation cycle is shown in Figure 16-3. In this figure (as well as Figure 16-6 below), labelled nodes represent ribosome states defined by the presence or absence of specific gatekeepers, while labelled arcs represent transitions driven by GTP hydrolysis, as well as peptidyl transfer itself. The A-site gatekeepers are an empty tRNA site, a fully accommodated aminoacyl-tRNA (aa), or a peptidyl-tRNA (pp). The exit tunnel gatekeeper is a hydrophobic signal polypeptide (SP) at the tunnel mouth. In this elegant implementation of process logic, test-action operations are realized as short sequences of conformational and covalent steps involving the gatekeepers and their GTPases. These primitive operations are combined as *while*-programs of branching and cyclical processes implemented by the ribosome (Harel et al 2001).

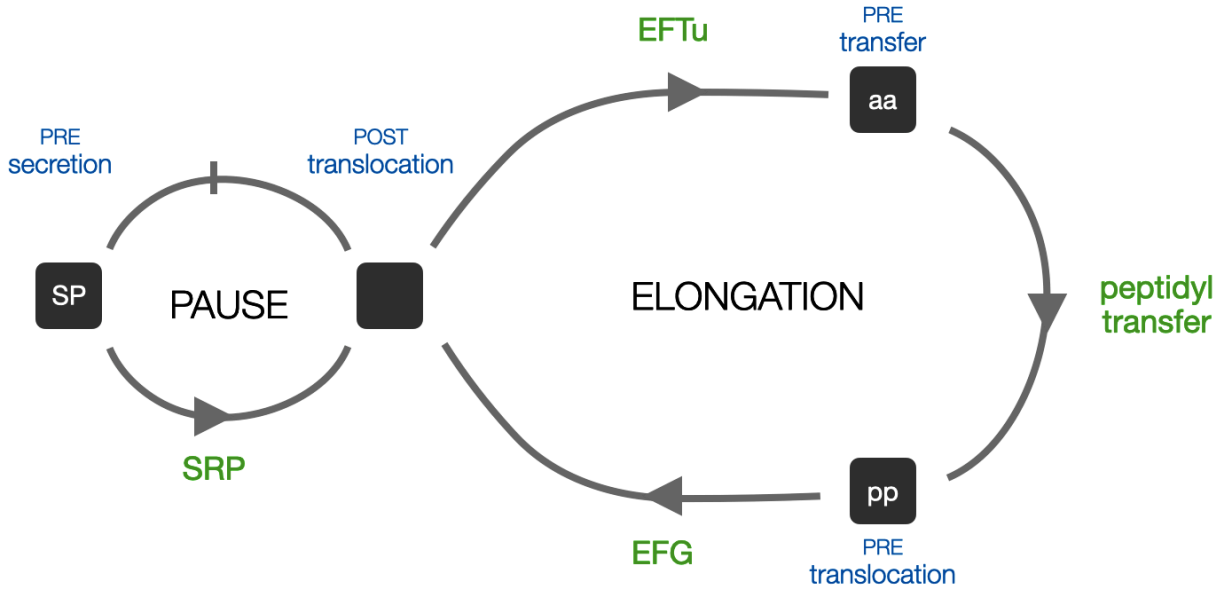


FIGURE 16-3. POLYPEPTIDE ELONGATION CYCLE IN LUCA

The thermodynamic problem of translocation is to move both tRNAs, their anticodons paired with the mRNA, one site in the ribosome, dragging the next codon

into the decoding center, despite slippery codon sequences, or adventitious secondary structure in the downstream mRNA. The kinetic problem is to keep up with the decoding step of the elongation cycle, which goes quickly when the first aminoacyl-tRNA sampled proves cognate, or more slowly with unlucky sampling, especially at rare codons. Translational GTPase EFG (archaea aEF2, eukarya eEF2) increases the translocation rate about 10^5 fold and reduces frameshifting (Peng et al 2019; Zhou et al 2019; Carbone et al 2021). Upon thermally driven intersubunit rotation, **GTP•EFG** docks in the factor binding site of the PRE-translocation ribosome where rigid extension of its C-terminal domains acts as a *doorstop* or *pawl*. Concurrent SSU head swivel and backward intersubunit rotation advances the $tRNA_2$-mRNA module, allowing the SRL to activate GTP hydrolysis. In what has been called a *compressed spring and trigger* mechanism, backward head swivel releases the inorganic phosphate and allows **GDP•EFG** to exit the factor binding site. Hydrolysis ratchets translocation as **GTP•EFG** cannot exit the factor binding site from the POST-translocation state.

In *Escherichia coli* and other gut commensal bacteria, EFG is replaced by an inactive GTPase variant called EFG2 in face of carbon starvation (Han et al 2023). This variant catalyzes translocation, but as it cannot hydrolyze GTP, it cannot drive the transition. As a result, it sustains translation at a much reduced rate, allowing a basal level of protein synthesis while conserving GTP stores. Although this hydrolytically inactive GTPase is derived from EFG, it hints at a primitive stage of protein life when translation was much slower, but more economical. [NB say something about (p)ppGpp alarmones versus GTP and GDP somewhere]

In sections 13 and 15 we discussed how primitive tRNAs defined and defended the reading frame in codon-directed polynucleotide termination on the heyday duplisome, and polypeptide elongation on the breakout ribosome, respectively. A variety of mechanisms, some more, some less conserved among the kingdoms of cellular life, have since evolved to better define and defend the reading frame during ribosome decoding and translocation. Today spontaneous frameshifting on bacterial ribosomes is less than 10^{-5} per codon overall (Kurland 1992).

Ribosomes are inherently prone to -1 slippage during translocation in the absence of EFG (Peng et al 2019; Zhou et al 2019). Acting as a pawl, the tip of EFG domain D4

uncouples the codon-anticodon duplex from the decoding center, and chaperones it during translocation to prevent slippage between mRNA and tRNA. Hydrogen bonds from EFG residues Q507 H583 to A37 N36, respectively, of the peptidyl-tRNA prevent back movement of the $tRNA_2$-mRNA duplex during translocation. Missense mutations in either of these pawl residues increase the rate of -1 frameshifting. In archaea and eukarya, the conserved pawl histidine[28] of EFG orthologs aEF2/eEF2 is modified to diphthamide (Djumagulov et al 2021; Milicevic et al 2023). Viable mutants of dipthamide biosynthesis in *Sacchromomyces cerevisiae* elevate -1 frameshifting, as well as ribosome drop-off from premature termination at out of frame STOP-codons (Shin et al 2023). The diphthamide residue is the target of ADP-ribosylation by diptheria and other bacterial toxins that prevent eEF2 from binding to the ribosome.

Whereas the pawl residues of EFG (aEF2/eEF2) interact with the peptidyl-tRNA during translocation, conserved features of the SSU rRNA and the uS9 ribosomal protein help stabilize the codon-anticodon pairing of deacylated tRNA in the P-site. In bacteria a modified guanine m2G plays the role of the hypermodified uracil $m^1acp^3\Psi$ [archaea??]. Conserved arginine in uS9 C-terminus stabilizes anticodon in P-site from the side. [Stm1?] In yeast and humans, a conserved uracil in the hairpin loop of helix h31 in the SSU head (U1191 *Sc*, U1248 *Hs*) is hypermodified to $m^1acp^3\Psi$.[29] This nucleotide interacts with tN34, and together with a conserved cytosine in the base of helix h44 in the SSU body (C1637 *Sc*, C1701 *Hs*), forms a lid on the wobble pair in the non-rotated PRE-translocation ribosome dubbed the *wobble seal* (Kisonaite et al 2022).

Whereas EFTu and EFG advance steps in the polypeptide elongation cycle, another ancient GTPase pauses elongation for co-translational secretion (Figure 16-3). Thus, when SRP GTPase, tethered to the S domain of SRP RNA, recognizes a hydrophobic sequence SP in the nascent polypeptide, the Alu domain of this RNA occludes the factor binding and A-sites, pausing elongation. Translation resumes when this GTPase docks with the homologous GTPase on the SRP membrane receptor, whereby SRP GTPase and SR GTPase activate one another, allowing handover of the ribosome

---

[28] ???(E.coli H583, S cerevisiae H699, human H715)

[29] $m^1acp^3\Psi$ denotes 1-methyl-3-α-amino-α-carboxyl-propyl pseudouridine

nascent chain complex to the membrane translocon for polypeptide secretion (ref). In section 12, we conjectured that the Alu RNA arose to regulate polymer elongation on duplisomes or ribosomes *before any support of GTPase enzymes*. Here we suggest that the first Alu-associated GTPases were activated by this RNA, much as translational GTPases are activated by the ribosome SRL. By LUCA, SRP was the sole extant Alu RNP, and the factor binding site was largely co-opted by translational GTPases.[30] Like pre-tRNAs, pre-SRP RNA is a substrate of RNase P, hinting an ancient nexus for their coordination.

In cellular life, initiation and termination of protein translation are key events in the regulation of gene expression. Unlike simple polypeptide repeats of indefinite length, it is crucial for globular protein domains that an entire sequence is present, and oftentimes, that there is no excess N-terminal leader, or C-terminal trailer. Thus, there was strong selection for precise control of polypeptide initiation and termination, as well as maintenance of the reading frame, to encode longer unique sequences, culminating in the self-folding domains with catalytic centers of extant enzymes. The antiquity of these mechanisms is still unclear as principal factors used for initiation and termination of protein translation postdate the split of bacteria and archaea some three billion years ago (Dever & Green 2012; Buskirk & Green 2017).

Here we précis key problems in initiation and termination of polynucleotide duplication and polypeptide translation, before suggesting a plausible evolutionary path of when, how and why, various solutions arose, and posing a few open questions. We conjecture that tRNAs arose for codon-directed termination of polynucleotide elongation (sections 13 & 14), and were co-opted for codon-directed polypeptide elongation at the breakout of translation (section 15). On the breakout ribosome, polypeptides likely terminated whenever elongation stalled with no aminoacylated-tRNA in its A-site. This happened at the end of the mRNA, of course, but also anywhere the template was damaged, or when there was no cognate or near-cognate aminoacylated-tRNA available to read the codon. As preservation of reading frame without an indel in the nascent polymer was far more important for polypeptide

[30] About 55 mya, the SRP Alu element has been co-opted by the eponymous SINE retroposon in primates to help steal the nascent RdRP enzyme of LINE mRNA (Ahl et al 2015; Al-Hashimi et al 2024).

translation than polynucleotide duplication, release of stalled polypeptides replaced the practice of duplisomes to forge ahead in polynucleotide copying, skipping past unreadable stretches of template.

In polymer biosynthesis, whether polynucleotide duplication or polypeptide translation, there were important tradeoffs between waiting patiently, forging ahead, or terminating hastily, for want of a charged elongator RNA in the polymer transfer center. Whereas ribosomes can idle awhile with an uncharged tRNAs in the A-site, then resume elongation on an aminoacylated-tRNA (Buskirk & Green 2017), we suggest that non-acylated tRNAs were affirmative gatekeepers in codon-directed termination of polynucleotide duplication (section 14). Besides slow default release on the breakout ribosome, we conjecture that polypeptides terminated more rapidly if a non-acylated tRNA accommodated in the A-site. Thus, on primitive terminating ribosome, sense codons and nonsense codons were read by aminoacylated and non-acylated tRNAs, respectively. The latter tRNAs might be constitutive terminators with no cognate charging ribozyme T, constitutive elongators inadvertently left uncharged, or conditional tRNAs that switched from elongators when charged, to terminators when uncharged. For sake of discussion, we conjecture the original role of EFTu was to monitor the aminoacylation status of consitutive elongator tRNAs to prevent premature termination from adventitious selection of their deacylated forms.

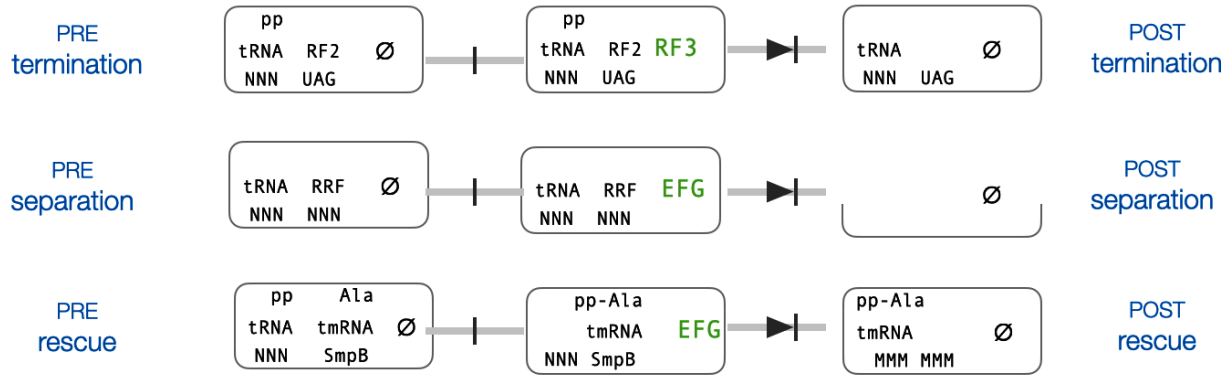


FIGURE 16-4. POLYPEPTIDE TERMINATION, RIBOSOME RECYCLING & RESCUE

Ribosomes seldom translate mRNAs to the end, but terminate when they encounter the first STOP codon, releasing their nascent polypeptide. They may then reinitiate on

another ORF downstream on the same mRNA, or recycle to initiate translation on a new mRNA entirely. STOP codons are recognized by class I release factors, tRNA mimics that sample the A-site codon in a compact form, opening into the peptidyl transfer center in an extended form upon codon recognition. In bacteria there are two class I release factors, RF1 and RF2, that read UAR and URA stop codons, respectively. Recognition of a stop codon results in stacking the third nucleobase of the stop codon on G530, and rearranges A1492 and A1493 into a termination-specific conformation. Packing the switch loop W319 in RF2? against A1492 and A1493 directs domain 3 into PTC for release. Accommodation positions the conserved GGQ motif of the release factor, which has a methylated amide on the glutamine side chain, in the peptidyl transferase center (ref). It is unknown whether this peptide motif activates hydrolysis of the peptidyl-tRNA directly, or acts indirectly by allowing solvent access.

Befitting their age, there are remarkable homologies as well as differences between termination factors in bacteria and eukarya (Buskirk & Green 2017). Although the sole class I release factor eRF1 is non-homolgous to bacterial RF1/RF2, it has the amidated GGQ motif of bacterial RF1/RF2 (remarkable convergent evolution?) (ref). In bacteria GTPase RF3 clears RF1/RF2 … post-termination ribosome. In eukaryotes, the GTPase eRF3 is evidently a paralog, not ortholog of bacterial RF3 …. bacterial RF3 spun off EFG, eRF3 spun off from EFTu. Spun off from EF-G, GTPase RF3 clears RF1/RF2 from the A-site of the post-termination ribosome to allow either read-through reinitiation or subunit recycling (2024).

Upon polypeptide termination, intact ribosomes can continue on the same mRNA to initiate a new polypeptide on the next ORF downstream (see below). However, swapping out one mRNA for another entirely, generally requires splitting apart the two subunits after termination. There are three likely reasons, not mutually exclusive, for splitting of primitive ribosomes, and perhaps earlier duplisomes, into two subunits: First, these great ribozymes could assemble on the template, and begin duplicating or translating from the middle, without scanning from the 5’ end. Second, they could terminate elongation and redeploy elsewhere without scanning to the 3’ end. Third, in polysomes, where several ribozymes translocate along the same template, individual

ones could initiate within the middle of the queue, or terminate and redeploy elsewhere without colliding with their neighbors, or idling uselessly.

In bacteria the gatekeeper for ribosome recycling is RRF, and this step is checked and driven by EFG itself (Janosi et al 1996). Genetic manipulations to lower available RRF can prevent normal recycling so that ribosomes pile-up near STOP codons; some are dislodged into the 3' UTR as POST-termination ribosomes, while others never reach the STOP codon, but back-up as PRE-translocation ribosomes (Saito et al 2020). Release of a natural mRNA with hairpin secondary structure in the 3' UTR requires subunit separation, but a truncated mRNA without this 3' trailer can be released from the intact ribosome (Wilmer 2022). Upon subunit separation, IF3 binds near the SSU platform to eject deacylated tRNA from the P-site, preventing premature subunit reassembly by occluding the B2 bridge with LSU H69 (Goyal et al 2017; Prabhakar et al 2017). Thus, the POST-separation ribosome comprises separate small and large subunits with deacylated tRNA and old mRNA cleared from the SSU so that a new mRNA may be translated. The mechanism of polypeptide termination is roughly analogous in eukarya and archaea, but the corresponding components are not homologous. termination uses an RF/GTPase complex. eRF1 is not cleared bu helps recruit the splitting factor (Rli1 yeast, ABCE1 mammals). ?ATPase …. Dom34 ….

Ribosomes stalled on aberrant mRNAs, or reaching the end without encountering a stop codon, are recognized and rescued by diverse quality control mechanisms that overlap normal mechanisms of termination and recycling (Buskirk & Green 2017; Korostelev 2021). Truncated mRNAs without stop codons can arise by gene mutation, interrupted transcription, or post-transcriptional cleavage by an endo- or exonucleases. In bacteria, *trans*-translation using an RNP of transfer-messanger RNA (tmRNA) and accessory protein SmpB is the primary mechanism to rescue non-stop mRNAs. In this coordinated process of tRNA selection and mRNA threading, Ala-tmRNA enters the A-site … acts as aminoacyl-tRNA in the peptidyl transfer, and as mRNA with ## codons ending with stop codon (Figure 16-4). Normal termination and recycling on the tmRNA stop codon….by RF?. The tmRNA.SmpB recruits the 3' exonuclease enzyme RNaseR to degrade the faulty mRNA, and meanwhile, the C-terminal tag AANDENYALAA targets the nascent polypeptide to the Clp protease

system. Reaching the end without canonical termination at a STOP codon mediated by RF1/RF2, and checked by RF3, bacteria employ various processes of *ribosome rescue*, *viz*. termination and recycling of ribosomes stalled midway along the mRNA, or at the ends of mRNA without stop codons.

Additional rescue pathways back-up trans-translation …. rescue pathways … [ArfA] [ArfB] [ref 79 endonuclease cuts A-site codon of arrested ribosome to allow tmRNA and alternatives to work]

After pondering the problems of polymer initiation, we suggest that *codon-directed initiation* was original to protein translation. For copying full-length ribogenes, the fundamental problems were to start as close as possible to the beginning of the template, and to finish as close as possible to its end. In sections 12 and 15 we suggested that at the breakout of translation, nascent polynucleotides and polypeptides were initiated on intact duplisomes, or ribosomes, respectively. In perhaps the simplest mechanism, dubbed *primitive leaderless initiation*, ordinary elongators, *viz*. duplicon-dRNA or aminoacylated-tRNA, were decoded in the A-site and advanced to the P-site through translocation *simpliciter* to initiate new polymers.[31]

The virtues of the A-site decoding center, *viz*. accuracy, generality, and frame defense, made this a poor place to invent codon-directed polypeptide initiation with a dedicated initiator tRNA and START codon. Rather the ability of elongators in the P-site to scan or skip past unreadable stretches based on a well-lubricated template channel, and the energetics and kinetics of anticodon-codon pairing, was co-opted to find polypeptide START codons. This required two new things: a dedicated initiator $tRNA_i$ that bypassed A-site decoding, as well as a means to exclude elongator tRNAs from the A-site that would compete in defining the start of translation. Thus, the factor IF1 occludes the A-site to prevent premature entry of the ternary complex **aa-tRNA•GTP•EFTu**.

Unlike primitive leaderless initiation from any codon that threads into the A-site, modern forms of initiation from the P-site look for a dedicated START codon near the 5' end of mRNAs (*leaderless initiation*), near the STOP codon of the upstream ORF

[31] [in working out the genetic code, ribosomes bound the 5' end of poly(U) and other synthetic polyribonucleoties, and started translation … or just bind of aatRNA? (Nirenberg)]

(*readthrough initiation*), or within longer stretches defined by specific mRNA features such as the Shine-Delgarno sequence (*canonical initiation*).

The thermodynamic problem of initiation is to find the start codon in a stretch of mRNA nucleotides using anticodon pairing, and other mRNA features forming its *ribosome binding site*. The kinetic problem is to quickly find either the first, or the next available start codon without skipping past it. In protein translation, a dedicated initiator tRNA and START codon in the P-site sets the reading frame as well as the polypeptide start (ref). The initiator $tRNA_i^{Met}$ (CAU) is charged by the same Met-ARS as elongator $tRNA^{Met}$ (CAU). In bacteria and mitochondria, this methionine is converted on tRNA to N-formyl-methionine (fMet) by the enzyme …. (ref). In bacteria, the anticodon CAU pairs with AUG as well as GUG START codons. Whereas elongator tRNAs form ternary complexes with **GTP•EFTu**, the charged initiator **fMet-$tRNA_i^{Met}$** forms a ternary complex with a different translational GTPase called IF2 (eukarya eIF5B). Spun off from EFG, IF2 checks the initiation complex with the dedicated initiator tRNA and START codon (Sprink et al 2016).

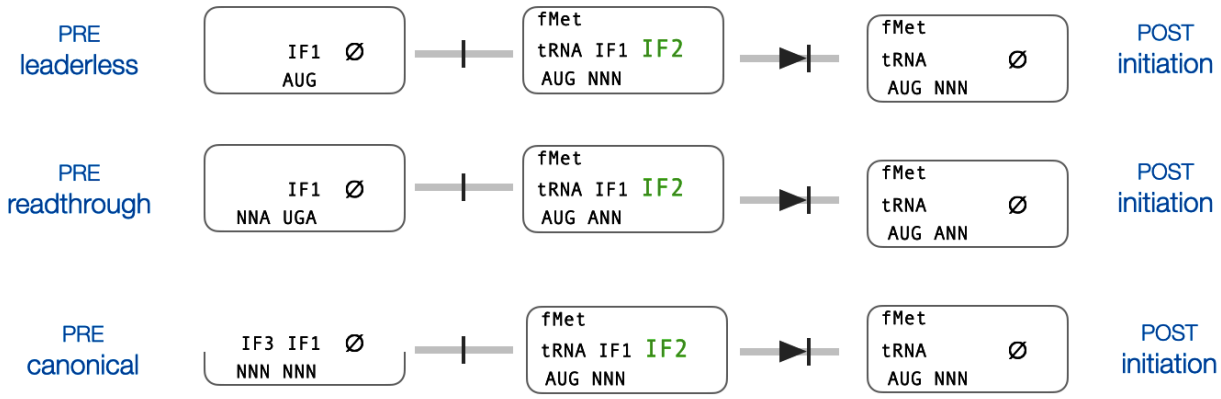


FIGURE 16-5. POLYPEPTIDE INITIATION

At least three distinct mechanisms of polypeptide initiation trace to LUCA called *leaderless*, *readthrough*, and *canonical* initiation (Figure 16-5). All forms require the dedicated initiator $tRNA_i^{Met}$ (CAU), gatekeeper IF1 and translational GTPase IF2. And, all three result in essentially the same POST-initiation ribosome, a special case of the POST-translocation ribosome with a “peptidyl-tRNA” of length one in the P-site and exit

tunnel. In bacteria, initiation on leaderless mRNAs with AUG at or very near the 5' end, the ternary complex **fMet-tRNA^{fMet}•GTP•IF2** docks in the factor binding and P-sites of the intact ribosome (Beck & Moll 2018). hydrolyzes upon successful mRNA threading with codon-anticodon pairing of the AUG of leaderless mRNA in the P-site. Several conditions favor leaderless initiation in bacteria including … the present of 5' p or ?5' OH from endonuclease cleavage from say the MazF toxin rather than 5' ppp of transcription, alarmones (p)ppGpp, or low temperature (Leiva & Katz 2022). Under stress, leaderless initiation requires fMet-tRNAi without IF2 or other coded initiation factors. Although these mechanisms hint at a primitive form of codon-directed initiation on intact ribosomes, these mechanisms are not presently well-understood, and should be presumed derived unless proved to be primitive.

Polycistronic mRNAs are common in prokarya, but their molecular biology still has many surprises (Kushner 2018). Underlying the phenomenon of *translational coupling*, initiation on a downstream ORF can immediately follow termination of the upstream ORF on the same mRNA. Like leaderless initiation, this readthrough initiation, or termination-reinitiation, likely occurs on intact ribosomes without subunit separation or recycling (Dever & Green 2012; Yamamoto et al 2016; Gunisova et al 2017; Inokuchi et al 2023). The strength of translational coupling falls off rapidly with intergenic distances beyond 10-25 nts. In unidirectional gene pairs from archaeon *Haloferax volcanii* and bacterium *Escherichia coli* that are candidates for readthrough initiation, AUG/GUG START codons typically overlap UGA or UAA STOP codons by -4 nt (AUGA) or -1 nt (UAAUG, UGAUG, UAGUG) (Huber et al 2023).

In *canonical initiation*, ribosome subunits are first recycled, or separated from one another, so that the SSU can associate with an entirely new mRNA. Beginning with the POST-separation ribosome, new mRNA binds the ribosome small subunit independently of initiation factors. IF3 prevents premature subunit reassembly and IF1 excludes premature entry of the ternary complex **aa-tRNA•GTP•EF-Tu** in the A-site. Meanwhile, the ternary complex **fMet-tRNA^{fMet}•GTP•IF2** places the anticodon in the P-site of the ribosome small unit for recognition of the mRNA START codon. Canonical initiation requires several seconds in bacteria, compared to some 20 elongation cycles that occur in one second. In prokaryotes, an RNA duplex between the SD (Shine-Dalgarno)

box, and the anti-SD sequence at the 3' end of the small subunit rRNA, helps position the AUG in the P-site (Shine & Dalgarno 1974). Validating this PRE-initiation complex of small subunit rRNA, mRNA, and initiator tRNA, IF2 allows subunit assembly, so that large subunit SRL activates hydrolysis of GDP and release of **GDP•IF2** and inorganic phosphate (refs). In eukaryotes, the ribosome small subunit recognizes the capped 5' end of the mRNA in proximity to the PABP factor on the 3' poly(A) tail, and scans the 5' UTR downstream with the anticodon of the charged initiator Met-tRNAiMet in the P-site for the first start codon.

The most profound difference between the initiator tRNA and elongator tRNAs is that it enters the P-site immediately without previous decoding, accommodation, and translocation from the A-site. However, it is interesting to ask why methionine was adopted to charge the initiator $tRNA_i$ (Bhattacharyya & Varshney 2016; Lahry et al 2024). Methionine was likely one of the last proteinogenic amino acids incorporated into the genetic code. It has no reported abiotic synthesis, and its enzymatic syntheses via homocysteine are costly in ATP consumption, as well as one-carbon metabolism. Methionine is important in cellular metabolism as a methyl donor (S-adenosyl methionine), and perhaps a fair proxy for amino acid availability generally. In bacteria, conversion of Met-$tRNA_i$ to fMet-$tRNA_i$ using N^{10}-formyl-tetrahydrofolate adds an additional layer of regulation. In bacteria and eukarya?, the initial methionine is routinely removed … allowing mature proteins to begin with any residue, and perhaps recycling this important amino acid. Although the same Met-ARS is used to charge initiator $tRNA_i^{Met}$ and elongator $tRNA^{Met}$, there is some doubt that $tRNA_i^{Met}$ spun off $tRNA^{Met}$ and not some other elongator tRNA (ref). Unique conserved features of the acceptor arm and anticodon arm of the $tRNA_i^{Met}$ help define its interactions with Met-ARS, preference for the P-site, and interactions with GTPase IF2 or other initiation factors.

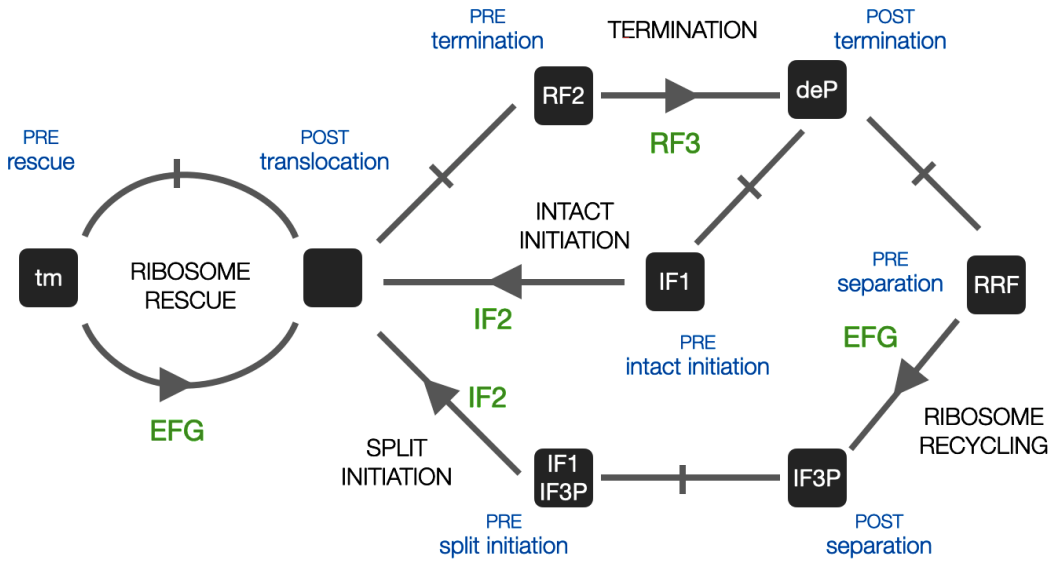


FIGURE 16-6. POLYPEPTIDE TERMINATION & INITIATION, RIBOSOME RECYCLING & RESCUE

In Figure 16-6, we complete the formal description of translation processes as a concurrent *while*-program of test-action operations implemented by the ribosome. To recognize various processes as parts of one flow graph, this figure can be combined with Figure 16-3 by identifying their common POST-translocation state, and with Figure 16-2 showing cyclical of tRNA sampling during decoding. Most of the gatekeepers used as test conditions are associated with the A-site; moreover, at most one of these can be present on the same ribosome. Besides the ground configuration of an mRNA codon but no tRNA in the POST-translocation ribosome, these are aminoacylated-tRNA (aa), peptidyl-tRNA (pp), tm-RNA (tm), RF1/RF2, RRF, and IF1. Two gatekeepers are associated with the P-site, viz. deacylated tRNA (deP) and IF3 (IF3P), and a third gatekeeper is associated with the nascent polypeptide exit tunnel (SP).

The Gibbs energy available for the ribosome elongation cycle is at least -30 kcal / mol from the hydrolysis of two (or more) GTPs and exergonic peptidyl transfer (Table 16-3). This compares to only -12 kcal / mol for the conjectural duplisome driven by the opening and closing dRNA thermal motors and isoergonic nucleotidyl transfer. The Gibbs energy of folding for RNA and protein domains are comparable, but folding of RNA domains is immediately offset by the cost of unfolding the RNA template, while the Gibbs energy of protein folding ultimately derives from cellular metabolism that supplies the amino acid pools.

	duplisome	ΔG kcal / mol	ribosome	ΔG kcal / mol
Initiation			IF2	-12
Elongation		- 12 / step		- 30 / step
Read	dRNA opening	12	EF-Tu	-12
Add	nucleotidyl transfer	~ 0	peptidyl transfer	-6
Move	dRNA closing	-24	EF-G	-12
Termination			RF3	-12
Recycling			EF-G	-12
Folding	RNA domain	-10	protein domain	-10

TABLE 16-3. ENERGETICS OF POLYNUCLEOTIDE DUPLICATION & POLYPEPTIDE TRANSLATION

17. Protein-supported RNA replication

The greater speed and fidelity of RNA-directed RNA polymerase (RdRP) enzymes forced the retirement of ribozymatic copying in favor of enzymatic replication. The duplisome hypothesis suggests this selective sweep was driven mostly by speed, while greater accuracy of all-protein replisomes emerged later. Thus, given merely adequate fidelity, polymerase enzymes that used NDPs or NTPs to extend primers would easily outpace duplisomes that used dRNAs to add one duplicon per day. One continuity from the duplisome to RdRP, and later DdRP, enzymes is that polynucleotides are made from 5' to 3'. Thus, despite differences in mechanism, step size, and especially, elongation rate, folding of nascent polyribonucleotides proceeds in the same direction, and perhaps via similar folding intermediates.

Attesting the importance of NTP substrates, as well as the virtuosity of early proteins, template-directed polymerase enzymes arose independently several times in protein life (Koonin et al 2020a). Thus, three lineally unrelated, catalytic core domains are found in polymerases today called Polβ-type, RRM-type, and DPBB-type. Any, or all of these folds may trace to RdRPs used in RNA replication, recombination, or repair in protein life. Today the RRM-Palm domain is found in replicative RdRPs of RNA viruses, while DPBB domains are found in regulatory RdRPs of eukaryotic RNA interference. Whatever the primitive RdRP polymerase enzymes, the new means of ribogenome reproduction now beat out the duplisome.

One immediate downside to enzymatic replication was that RdRPs produce long duplexes that require strand separation, and invest every other round of copying in making reverse complements (minus strands) that neither fold as ribozymes nor function as mRNAs. Similar limitations of spontaneous RNA copying are conjectured to have given a leg up to the duplisome in an early RNA world (Campbell 1991; Noller 2012; Zhou et al 2021; Ding et al 2023). Seen as passive intermediates of RNA replication, reverse complements and long RNA duplexes presented an immediate challenge for polynucleotide copying. Thus, there was strong selection for RdRP-based replisomes to accrete additional activities, *viz*. helicase, sliding clamp, and proof-reading, to overcome obstacles to processivity and accuracy of RNA copying.

Whereas thermodynamic and kinetic obstacles to enzymatic RNA replication were immediate, more profoundly, its duplex intermediates and antisense products afforded new levels of regulation of genome reproduction and gene expression, as well as new strategies and tactics of genetic mixis and genome defense. Hitherto, in duplisome-mediated RNA copying, regulation concerned whether to initiate a polynucleotide, and whether and where to pause or terminate elongation, and whether to thread a second template to create a novel ribogene. The various full-length and truncated products might themselves be catalytically inactive, fold as active ribozymes and riboswitches, or now, serve as mRNAs. In ribosome-mediated protein translation, regulation was enriched to codon-directed initiation and termination, conditional pausing, and co-translational secretion (section 16).

In enzymatic RNA replication, there were now two RNA strands, plus and minus, that could be replicated fully or partly, folded, sequestered, or degraded, separately of the other stand. Antisense RNAs arising as replication intermediates afforded new forms of *sequence-directed regulation*, both *cis*-regulation of the sense RNA, and *trans*-regulation of otherwise unrelated RNAs that share the cognate recognition sequence. Networks of nuclease enzymes and argonaute proteins arose that prepare and target small interfering RNAs (siRNAs) to reversibly alter activity of ribozymes and riboswitches, or prevent translation of mRNA, as well as direct their irreversible endonucleolytic cleavage (Fire & Mello Nobel; Ambros & Ruvkum Nobel; Al-Hashimi et al 2024).

With the takeover of RNA replication by polymerase enzymes, ribozyme P had no more dRNAs to load, and was relegated to clearing away the leaders of pre-tRNAs. We suggested above that this latter activity, which depends on the S-domain of RNase P RNA, arose to prevent adventitious charging of 5’ end of tRNAs with oligomers that interfered with tRNA functions in polynucleotide termination, and later, polypeptide elongation. Today RNase P clears away transcribed pre-tRNA leaders, but even this vestigial role is marginal as some leaderless tRNAs are primary transcripts, indicated by 5’ triphosphate in the mature tRNA (Gupta 1984). First reported for $tRNA^{fMet}$ of the archaeon *Haloferax volcanii*, leaderless transcripts are rare, and no organism has eliminated all need for RNase P in tRNA maturation. Presumably, it has been easier to

retain this activity than position all tRNA transcription starts so precisely. In the kingdoms of cellular life, ribozyme P has accreted various protein partners, and moonlights on novel RNA substrates (ref). In eukarya, the catalytic RNA duplicated and diverged as RNase P and RNase MRP, and was replaced in the mitochondrion by an all-protein RNase P (ref).

18. The DNA world

Required for folding and activity of ribozymes, the 2’ OH group of ribose sugars is a liability for general and durable information storage for three reasons: First, O2’ catalyzes strand scission via its 2’, 3’ cyclic phosphate intermediate. Second, this hydroxyl group restricts the sugar pucker, limiting RNA duplexes to the A-form. Finally, its proclivity to form hydrogen bonds with the RNA backbone and nucleobases, as well as water and metal cations, allow RNA sequences to explore complex folds. Eschewing the 2’ OH, polynucleotides based on deoxyribose are more covalently stable and conformationally uniform polymers. Thus, the DNA backbone is stable to spontaneous strand scission, and for want of better alternatives, high complexity DNA sequences settle into long regular B-form duplexes with their reverse complement. Of these two chief advantages of deoxyribogenes, the greater backbone stability was immediate, while error-free repair pathways emerged over time to exploit the informational redundancy of dsDNA.

The genome handover to DNA required not one, but two polymerase enzymes: the DdRP *transcriptase* (DPBB fold), and the RdDP *reverse transcriptase* (RRM fold). Both enzymes were no doubt exapted from an existing RdRP of the corresponding fold. Together these polymerases copied RNA into dsDNA for longterm storage and VGT, and copied DNA into noncoding RNAs and mRNAs for gene expression. One parsimonious suggestion is that the transcriptase arose from the replicative RdRP through a facile change in template preference from RNA to DNA (Koonin et al 2020a). If so, the breakout of DNA life required only the reverse transcriptase, plus a robust source of dNDPs/dNTPs.

In cellular metabolism dNTPs are made enzymatically from NDPs by *ribonucleotide reductase,* which converts them to the corresponding dNDPs, followed by *nucleoside diphosphate kinase* which promotes dNDPs to dNTPs at the expense of ATP. This observation that deoxyribonucleotides are made from their ribonucleotides, and not from deoxyribose directly, was one of the original arguments for the RNA world (Long et al 2022). Mixed polymers of ribonucleotides and deoxyribonucleotides had some, but not all advantages of pure deoxyribogenes. An example of orthogonal evolution, it

was crucial for expression of active ribozymes that the transcriptase exclude dNDP/dNTP substrates, but less important for gene replication that the reverse transcriptase exclude NDP/NTP substrates.

In a best of both worlds for chromosome life, DNA strands hold together well enough for genome storage, yet separate easily enough for genome copying, and gene expression. Redundant sequence information on complementary strands is near at hand, not just to template new copies, but for error-free repair of damage confined to one existing strand. One further advantage of duplex DNA as the genetic polymer is that the rates of spontaneous hydrolysis of nucleobases, notably depurination of A and G, and deamination of C, are much lower in paired than unpaired regions. Even so, mutations from depurination were reduced further by the error-free enzymatic pathway of *abasic site repair* (ref).

Rather than discovering a chemically more stable nucleobase to replace cytosine, evolution replaced uracil, its hydrolysis product, by thymine (5-methyl uracil) in DNA (ref). The key idea for detection and repair is that uracil should not normally be present in DNA, unlike RNA where it is a standard nucleobase. Shunted to dTTP, the pool of dUTP is kept quite low, so that the DdDP enzyme incorporates thymine (T) when copying adenine (A). There are two residual sources of uracil in DNA today: (1) incorporation of trace dUTP opposite A by the polymerase during replication, resulting in a U:A pair; and (2) spontaneous deamination of cytosine in duplex DNA after replication, creating a U.G mispair. Both are efficiently detected and corrected through error-free repair pathways (ref). The former is subject to uracil-DNA glycosylase for *base excision repair*, while the latter is subject to *mismatch repair* as well. The default pathway for undetected U.G mispairs is error-prone repair to U:A and C:G in the next replication round.

The replacement of uracil by thymine in DNA required three new enzymes to shunt dUTP to dTTP (ref). First, dUTP, derived from UDP by the actions of ribonucleotide reductase and nucleoside diphosphate kinase, is efficiently demoted to dUMP by *deoxyuridine triphosphatase* (dUTPase). Spontaneous or enzymatic deamination of dCMP are additional minor sources of dUMP. Next dUMP is converted to dTMP by *thymidylate synthetase*, and then promoted to dTDP by *thymidylate phosphate kinase*,

and finally, to dTTP by the common *nucleoside diphosphate kinase*. Controlled at the level of substrate pools, DdDPs do not distinguish between the major pool of dTTP, and traces of dUTP kept low by dUTPase. *Is this generally true for DNA polymerases?*

Two features of nucleotide metabolism and DNA replication suggest there was likely a transitional stage in the evolution of deoxyribogenomes where mixed pools of dUTP and dTTP were used by the DdDP according to their availability. First, there is no *dUDP phosphatase* enzyme to convert the dUDP made by ribonucleotide reductase directly to dUMP. Rather dUDP it is promoted to dUTP, and then hastily demoted to dUMP as an afterthought. This suggests that dUTP was once used for DNA replication and this dUTPase was a later development. Second, any possibility of error-free repair of uracil DNA requires extremely low levels of uracil in normal DNA. This suggests that low and variable ratios of thymine to uracil nucleobases in DNA conferred some immediate advantage, perhaps in replicase recognition or nuclease resistance, while the error-free repair pathways for uracil-DNA emerged later when nearly uracil-free DNA was the norm. Perhaps relevant, DdRPs can bypass uracil in RNA transcription. *What was the original purpose of spiking deoxyribogenes with thymine?*

In the final curtain call of the RNA world, DdDPs all but eliminated the need for RNA intermediates in DNA replication. Although some components of the DNA replisome are conserved in all cellular life, *viz*. sliding clamp, clamp loader ATPase, and ssDNA-binding protein, other components, *viz*. helicase, primase, and the replicative DdDP itself, are not homologous among bacteria, archaea and eukarya (Leipe et al 1999). One attractive proposal is that the replicative DdDP of LUCA arose from the universal DdRP transcriptase with its DPBB fold (Koonin et al 2020a). That DPBB fold (PolD) DNA polymerase, which remains the main replicative DdDP in archaea, has been replaced by RRM fold (PolB) or Polβ fold (PolC) DNA polymerases in eukarya and bacteria, respectively. Whatever this history of genome replication and repair enzymes in cellular and viral life, these fundamental processes were fluid, not fixed, with occasional switching of template preference from RNA to DNA, and back, or nucleotide preference from NDP/NTP to dNDP/dNTP, and back. Interleaving these lineal changes in proteins through VGT, polymerases and ancillary components underwent repeated non-orthologous gene displacements through HGT (Koonin et al 1996).

Alongside chemical kinetics and thermodynamics, new information sciences provided a conceptual framework for molecular biologists exploring biopolymers and their processes of replication, transcription, and translation. Claude Shannon formulated memory stores to WRITE and later READ messages, or communication channels to SEND and remotely RECEIVE messages as physical media configured for alternative messages or states (Shannon & Weaver 1949). He showed how the current state of any memory store or communication channel could be abstracted as a word, or sequence taken from some fixed alphabet of two or more letters. Useful storage media afford: (1) uniform storage of all well-formatted messages; (2) message stability in passive storage and active use (repeatable or NON-DESTRUCTIVE READ); (3) mechanisms of copying (WRITING) from one store to another of the same medium/format; (4) mechanisms of updating or editing (OVERWRITING) in the same medium/format; (5) mechanisms of translating into a different medium/format; and (6) adequate separation of well-formatted messages to allow ERROR DETECTION-AND-CORRECTION.

Explaining its instant popularity, Shannon's simple and general theory of communication abstracted the quantity of information from the particular quality, or meaning of any message (Watanabe 1969). Just as thermodynamics considers the relative occupancies of states, not the reaction paths that give processes their meaning, information theory considers the relative frequencies p_i of messages, not the encoding and decoding functions that give any message m_i its meaning. All noticed the close relation of information $I = -\Sigma p_i ln p_i$ to the concept of configurational entropy from statistical mechanics. In simple cases, encoding and decoding are 1-to-1 functions, and no distinction is made between the message as written and stored, say as a tape or polymer, and the message as read and used (Watson & Crick 1954b). More remarkable decodings are needed to interpret computable functions encoded as say Gödel numbers or Turing tables, metabolic reactions encoded as say gene sequences, memories and behavior encoded as say cerebral circuits, or abstracted as linguistic discourse (Gödel 1931; Turing 1936). Like thermodynamics, information theory was inadequate for understanding the organization and function of regulatory mechanisms, so biologists and cognitive scientists continued their quests for a *lingua franca*.

Happily, the language of discrete processes of actions and tests describes plans of motivated behavior with explicit goals, as readily as procedures and mechanisms with extrinsic purpose known only to their watchmaker (section 19).

DNA chromosomes created an entirely new level of regulation for genome replication, gene expression, and indeed, gene invention. Early molecular biologists demonstrated that duplex DNA, working with RNAs and enzymes for transcription and translation, elegantly satisfy most of the criteria for an ideal information storage medium from computer science. Whereas ribosomes and tRNAs were central to translating mRNAs into proteins, the real smarts of genome replication and gene transcription appeared to be deoxyribogenes in collaboration with DNA binding proteins and polymerase enzymes. Thus, in their influential analysis of gene regulation in the bacterial *lac* operon, François Jacob and Jacques Monod established the paradigm of sequence-specific DNA binding proteins and gene-proximal regulatory sequences providing positive and negative control of mRNA transcription (Jacob & Monod 1961).

Since the heyday of the central dogma, the hegemony of DNA and proteins has been challenged by discoveries of RNA functions in conservative and creative processes of heredity, from gene expression to gene invention. Noticing the vast amount of repetitive and poorly transcribed DNA in eukaryotic chromosomes, Roy Britten and colleagues conjectured that non-coding RNAs act to regulate transcription of themselves, or other genes, but never leave the nucleus (Britten & Kohne 1968; Britten & Davidson 1969). Whereas bacterial operons matched protein keys to DNA locks to regulate transcription, it seemed that gene regulation in eukaryotes might match antisense RNAs to DNA genes, or their nascent RNA transcripts. However, the evidence for such processes was speculative at best, and their mechanisms entirely unknown. Whatever their role, RNAs were still the handmaiden of their deoxyribogene: “We do not, in this model, wish to specify a mode of action for the receptor gene—that is, the nature of the molecular events occurring between the DNA, histones, polymerases, and so forth, present in the receptor complex. This model is concerned primarily with interrelations among the DNA sequences present in the genome” (Britten & Davidson 1969 p 350).

No sooner had the stable rRNAs and tRNAs, and more mercurial mRNAs, taken their places in protein translation, than palimpsests and imprimaturs of RNA life were found to pervade all aspects of genome replication and gene expression. Many roles were clearest in eukarya because of the strict spatial and temporal segregation of DNA replication and RNA transcription in the nucleus, from protein translation in the cytoplasm (Koonin). Thus, guide-RNAs catalyze rRNA maturation in the nucleolus, while tRNA maturation commences with RNase P cleavage in the nucleoplasm. Meanwhile, guide-RNAs of splicosomes, and nuclear introns themselves, are remnants of self-splicing group II introns from RNA life that had snuck into chromosomes via reverse transcription. As undesirable and parasitic as these seem, the process of intron splicing afforded new opportunities for post-transcriptional regulation of gene expression in eukarya.

Beyond catalyzing the maturation of rRNAs, tRNAs, and mRNAs, other roles of RNAs in genome replication, gene expression, and gene invention were soon discovered. In all cellular kingdoms, replication of DNA chromosomes requires a DdRP to make the short RNA primers that initiate DNA replication. These *primase* enzymes of DNA replisomes were likely exapted from RNA replisomes. In eukarya, replication of the ends of nuclear chromosomes require telomerase, an RdDP enzyme and guide RNA that maintains these telomeres against erosion in lagging strand replication.

In eukarya, cis-acting RNAs guide the creation and maintenance of repressive chromatin through argonaute-mediated matching [RNA/RNA, RNA/DNA]. The small functional centromere with inner kinetochore proteins is transcribed and this centromeric RNA represses nearly identical centromeric repeats that flank it. A similar mechanism uses XIST RNA made from the X-inactivation center (XIC) to guide *cis*-inactivation of genes along the X-chromosome, and *trans*-inactivation of the XIC on the active X chromosome. Whereas centromeres have characteristic functions in chromosome segregation, *enhancers*, the hallmark regulatory element of eukaryote gene transcription, are deoxyribogenes that express lncRNAs to regulate themselves and topologically associated genes through argonaute-mediated transcriptional silencing. Thus, RNPs of argonaute with guide RNAs that target mRNAs in the cytoplasm were exapted in chromosome life for regulation of DNA replication and

transcription by chromosome-associated RNAs in the nucleus. The key unit of this chromatin-based regulation is the DNA loop defined as the region between two active CRCF binding elements oriented head-to-head to delimit loop extrusion by cohesion motors.

Working through their RNA/RNA or RNA/DNA duplexes formed with support of argonautes or other duplex-matching proteins, the handmaiden RNAs of the nucleus were now recognized as “smart” chromatin marks that program and maintain patterns of deoxyribogene replication and expression across mitotic cell divisions (Miga & Alexandrov 2021; Stefanov & Nowacki 2024). This RNA-guided regulation of DNA replication and transcription allowed deoxyribogenes to feedback onto themselves and forward onto cognate targets in stable states of chromatin that could perdure through one or more cell divisions. Thus, patterns of gene expression, and cell fate decisions could be transmitted to daughter cells as the landscape of euchromatin and heterochromatin. The same pattern could be inherited by both daughters in simple proliferation, or by one daughter only, in the stem-leaf pattern of one daughter resembling the mother, while the other is reprogrammed to a novel leaf fate.

The functions of RNA in DNA replication and transcription seemed to fall under the rubric of conservative search, *viz*., retracing and exploiting past successes, not creative search, *viz*. exploring the epigenome, or even the genome. Whether controlled mixis and in-house discovery, or promiscuous HGT, genome tinkering is a double-edged sword, promising new affordances, and threatening hostile takeover by selfish replicators. Thus, the creative processes of heredity are contrasted as *gene invention*, *viz*. curiosity about novel genes that might be opportunities, and its obverse, *immunity*, *viz*. caution about foreign genes that might be Trojan horses. One emergent principle of chromosome life is that the DNA duplexes are normal for deoxyribogenes, and single-strand RNAs are normal for gene products, but RNA duplexes are often indicative of selfish or foreign elements. Thus, argonaute-based mechanisms exapted to repress genes present in suspiciously many copies, transcribed in both orientations, or unvetted sources … (ref). [piwi] Meanwhile, adenosine deaminase (ADAR) to edit A-to-I in RNA duplexes … prevent export from the nucleus (ref).

Falling under the rubric of exploration, any process of creative heredity, whether spontaneous mutation or planned (catalyzed) variation, ultimately changes the length and sequence of the genome. In chromosome life, single-nucleotide substitutions and small indels occur spontaneously at high-rates, and are corrected by error-free DNA repair pathways. Various forms of homologous, or site-specific recombination allow cautious forms of sampling the pangenome in prokarya, or comparing in-house alternatives of alleles and haplotypes in eukarya. Unequal crossing over can increase or increase the length of tandem repeats.

More dramatic genome variations entail chromosome rearrangement, gene combination and duplication, as well as movements of mobile genetic elements (MGEs). Conceptually, the simplest process of introducing a new gene into the chromosome, whether imported from without, or provided in-house, is to create a double-strand break (DSB) and insert the foreign/novel DNA in between. The endonuclease enzyme of DNA transposons, or transposase, catalyzes essentially this cut-and-paste mechanism with minimal sequence requirement for their chromosome insertion sites. There are many variations on the process wherein copying and insertion are embedded within a larger process of DNA replication, possibly with RNA transcription and reverse transcription. These include the generation of tandem repeats through rolling-circle replication, and insertion of LINE retrotransposons through target-site primed reverse transcription.

Across all kingdoms of cellular life, RNAs transcribed from chromosomal genes, viruses, and transposons, are important intermediates in creative heredity, including gene invention and mobile forms of self/nonself recognition. Not only can RNAs provide the sequence to be inserted through reverse transcription, but they are implicated in targeting the DNA sites for insertion. It bears repeating that in the transition from protein life to DNA life, reverse transcription had an obligate role in the replication of deoxyribogenes before DdDP enzymes, and perhaps in assembling the first DNA chromosomes after DdDPs.

In eukaryotes, there is close relation between genomic variation and epigenomic repression. Geneticists early noticed that heterochromatin has higher mutation rates than euchromatin and changes more rapidly in evolution (Roberts & Gordenin 2014).

Before the genome sequencing projects, heterochromatin was known as a junkyard of repetitive and selfish DNA. Technical difficulties in sequencing and assembling these regions left these simple ideas about heterochromatin stand until recently. Indeed the conservative processes of heredity and Mendelian exchange are compartmented to euchromatin, *viz*. early DNA replication in the cell division cycle, error-free repair of DSBs in S/G2 by homologous recombination if needed, and meiotic crossovers between homologs vetted by the synaptonemal complex. Conversely, the creative processes of gene invention and HGT are compartmented to heterochromatin, *viz*. late DNA replication, RNA catalyzed DSBs, and novel combination of deoxyribogenes from NHEJ or other error-prone DSB repair.

Molecular geneticists discovered that the repressive heterochromatin found subtelomeric, pericentromeric, or sporadic positions along chromosomes, was unstable both in its epigenomic state (facultatively repressive or active chromatin) and its genomic state (error-prone replication).

HR repair in S/G2

NHEJ at all stages of cell division cycle

AEJ alternative end joining

SSA single-strand annealing

In both unicellular and multicellular eukarya, most of the chromatin landscape is reset with each generation through the events of meiosis and zygosis. Unexpectedly, at some sites the epigenomic state can be transmitted across one or more generations. This discovery of trangenerational inheritance of epigenomic states gave new life to the idea of Lamarkian inheritance, and the Baldwin effect, driven out of the biology of multicellular plants and animals by August Weissman's theory of segregation of the germline and soma.

How is epigenomic regulation coupled to planned variation of the genome?. Examining the draft sequence from the *Caenorhabditis elegans* genome consortium, we found remarkable evidence for a protracted and concerted process of gene discovery in facultative heterochromatin (Hutter et al 2000). In these dynamic regions, dubbed local gene clusters in nematodes, or segmental duplications in mammals, candidate genes or pseudogenes are combined, duplicated, and deleted through RNA-

mediated processes of transcription and reverse transcription. In the short run, novel genes at the euchromatin-heterochromatin boundaries are sometimes expressed and contribute to fitness through their RNA products, including translated mRNAs, perhaps for several generations. But meanwhile, they are threatened with repression, and possible deletion, catalyzed by RNAs from homologous gene candidates in either orientation within the same cluster or at a local gene cluster on another chromosome entirely. Importantly, typical clusters contain admixtures of two or more gene families because the mechanism of RNA-mediated copying favors nearby genes, but is not strict. This concerted evolution in gene clusters underlies the phenomena of copy number variation.

To survive in the long run, one unique gene must emerge, and all homologous clusters must shrink, fixing this gene within euchromatin, subject only to the traditional allelic tourneys of population genetics (Hutter et al 2000). Looking for analogies to the concerted evolution of genes in local gene clusters, and their rapid-prototyping in this facultative heterochromatin, there is a fair analogy to the tenure mechanism on social organizations that separates the initial hiring decision from the extended vetting process for retention. In these extended trials, the probationary employee performs a range of likely tasks, and is perhaps compared to other interns, before a final selection. In section 19 we suggest the process of gene tenure from heterochromatin to euchromatin is an epigenomic version of habit formation, or downward consolidation through repeated practice in multilevel search. Thus, to become habitual, a successful gene is consolidated from creative inheritance and optional expression in facultative heterochromatin to conservative inheritance and reliable expression in constitutive euchromatin. In these tourneys, the dynamic population of paralogs of the successful gene found nearby in “local gene clusters” or “segmental duplications” are lost.

[LOCAL DUMP]

[combines recent discoveries and older search results]

At first, domesticated retroelements in telomere maintenance were considered exotic, perhaps vestigial, but in fact the back and forth of transcription and reverse transcription plays a principal role in the creative search for invention of possibly useful

genes in eukarya (Hutter et al 2000; Miga & Alexandrov 2021; Al-Hashimi et al 2024; Stefenov & Nowacki 2024).

Hints of more from Similarly, reverse transcription, part of the life cycle of mobile retroelements, is used to replicate and repair the ends of nuclear chromosomes in eukarya, or the vestigial centromeres flanking …..

Today dsRNA is the central element in sequence-directed gene regulation and genome defense ….(Al-Hashimi et al 2024).

In this process of gene discovery and invention. remarkable unforeseen and unforeseeable opportunities, breakouts … immediate advantages and emergent opportunities … [exploration]…. without invoking teleos, explore directions believed promising based on recent experience, abductive bias …. the mechanisms of recombination underlying gene duplication, regulation of euchromatin and heterochromatin, to maximize R & D without compromising the faithful transmission of well-tested or tenured genes .. mixis, horizontal, risk selfish and useless at cost of discovery …. recombination in regular mixis, recombination in radical exploration

The concept of core and pangenomes works well for ribogene communities in the RNA world. The regular organization of genes into chromosomes … new opportunities for gene discovery … and opportunities for gaming the system with MGEs and sessile cliques. Whereas one gene in MGE, linked cliques or disbursed cliques.

To effectively exploit past experience, adaptive systems must either keep records of past searches in an immediately useful form, viz. executable routines of present behavior, or have the time and means to compile, or translate them, from higher-level formats of search maps, to lower-level formats of search plans.

First, what were the maps inherited from past experiences of the species? Second, how were these compiled as executable plans to mediate and regulate the varied processes of cell biology, development and behavior, as well as evolutionary search itself?

duplication and specialization, recombination to join together

Ga	Era	RNA	polypeptides	DNA
4.5	prebiotic	random	random	none
	early RNA life	spontaneous copying	random	none
	late RNA life	duplisome	random	none
	early protein life	duplisome →	ribosome	none
	late protein life	RdRP enzyme	ribosome	none
	early DNA life	DdRP enzyme	ribosome	RdDP enzyme
3.7	late DNA life	DdRP enzyme	ribosome	DdDP enzyme

TABLE 18-1. THE FIRST BILLION YEARS OF POLYMER LIFE ON EARTH

19. *Plus ultra*

The saga of polymer life through LUCA spans the first billion years of life on Earth (Table 18-1), while eukaryogenesis, leading to sexual reproduction, spans the second billion years (Maynard Smith & Szathmary 1995; Martin & Koonin 2006). New levels of evolutionary search, no less remarkable, have emerged in the greater age since then. Focusing on just our species, our society, and our sciences, we hope to reconstruct the major stages of metazoan evolution from segregation of germ cells and multicellularity of the soma, through the neural circuits of behaving, learning, and reasoning, and now, the uniquely human faculty for sharing fragments of our cognitive map with linguistic conspecifics (Pinker 2015, 2022; Grillner 2021; Lamanna et al 2023). All acknowledge that our cerebrum and its specialized functions evolved from simpler stages of the vertebrate brain, though some still carve out an exception, or impose a firewall for language (Chomsky 1995; Berwick & Chomsky 2016).

Proposals that a common logic of discovery underlies biological and social organization, including our arts and sciences, have appeared many times, and been rejected as often (Spencer; Veblen 1899; Hofstadter 1944; Campbell 1990; Holland 1992; Cziko 1995). Historians and political philosophers long pondered the regulatory principles of social organization, but comparable principles of biological organization only become known with Mendel. Today we know the principles of biological and social organization are encoded and preserved as nucleic acid polymers and natural language texts, respectively. There is no agreement the processes that create and use such different forms of knowledge can be usefully explained within one framework, much less what that formulation might look like. Like replicator dynamics for populations of biopolymers and cells, *meme theory* describes some population phenomena in sociology without framing any deeper principles (Dawkins 1976). More promisingly, Donald Campbell's model of multilevel search, dubbed *evolutionary epistemology,* explains abduction as a creative process of variation and selection in which higher-level ends naturally beget vicarious, or lower-level ends, as their means (Campbell 1974).

Lumping easy and hard targets together, critics of the relation between learning and evolution argue that the apparent similarities between different forms and levels of organization, are superficial, while their overt differences are profound. In his critique of meme theory and evolutionary epistemology, philosopher Paul Thagard made abduction the central pillar of cognitive science and scientific discovery, even as he doubted any mechanism and role of abduction in biological evolution (Thagard 1980). This error was only possible because he, and many others, conflate the simple theories of population genetics and replicator dynamics[32] with the actual processes, and remarkable products of evolution.

Responding to Isaac Newton's union of geometry and dynamics in *Philosophiæ Naturalis Principia Mathematica*, Giambattista Vico doubted that true laws of nature could be captured in this, or any artificial language (Newton 1687; Vico 1710). In his aphorism *verum esse ipsum factum*, Vico claimed that humans understand manufactured artifacts more intuitively than discovered nature. Indeed philosophers first looked to contemporary human arts and machines for concepts and principles to understand the regulatory mechanisms encountered in nature. In psychology, for example, the working metaphors of mind progressed over time from clockworks to Jacquard looms and telegraph networks to computers (Boden 2006). But from planetary motion to atomic structure, Vico's aphorism proved wrong, *viz*. our most fundamental theories invented to explain natural phenomena have preceded, not followed significant applications to say artificial satellites, pharmaceutics, or genome engineering.

---

[32] Replicator dynamics formulates pure r-selection. In the search theory of knowledge, neither reproduction of individuals, nor increase of their numbers *tout court* are essential for evolution, nor sound measures of progress or fitness. Only the first copy (creation) and the last copy (extinction) of genomes or texts are critical, not head-counts in between. For variation and selection, it suffices that two versions are compared, and the better one retained. Even this can be a comparison between the current system and an earlier restore point, so that the evolving lineage has just one working representative at any time. Beyond planned variation, or the creative role of heredity, the conservative role of heredity is to ensure a back-up copy is available whenever the working copy is casually destroyed. Under ordinary conditions this is akin the faithful copying and safe-storage of irreplaceable business records, something to do, but not the make-or-break of a successful business. Under pure r-selection, destruction of individuals, and extinction of lineages *in toto* become the means of selection, but they are not the focus of variation and selection, nor necessary to discovery in general.

If Vico was mistaken to dismiss formal languages as epistemological interfaces for the natural sciences (*Naturwissenschaften*), he seemed correct in concluding these languages were ill-suited to understanding human society (Vico 1725). Contesting the hegemony of natural sciences, new social sciences (*Geisteswissenschaften*) argued that texts were irreducible to Cartesian geometry or Newtonian dynamics, but divided over whether formal languages or quantitative models had particular uses. Postmodern philosophers doubted the interpretive disciplines should be called sciences at all, eschewing formal languages for poetic forms of natural language (Geertz 1973; Rorty 1979). Meanwhile, the "hardboiled" or "scientific" wings of economics, organization science, and linguistics continued to formulate fragments of social sciences as automata, games, or quantitative models.

In the long back-and-forth of thought between biology and political economy, scientists framed sundry phenomena of multilevel organization and coopetition in terms of various agents, constitutions, contracts, ecosystems, firms, laws, markets, mutualism, sectors and species. Despite many insights, on say microeconomic theory of markets, contractual organization of the firm, or embedding games within evolutionarily stable strategies (ESS), and no small number of Nobel Prizes in economics, most models of the statics and dynamics of social organization lost their naive charm in attempts to generalize them to more phenomena (Smith; Marshall; Walras; Coarse; Arrow; Simon; Maynard-Smith).

As contentious as comparing biological and social organization, there is quiet angst in the natural sciences, if not loud disagreement, about how to formulate the Darwinian concept of purpose beyond its original gloss as survival and reproduction. Looking to dynamical concepts from the physical sciences for similarities and differences of living systems from their non-living surroundings, many pinned their hopes for an essence of life variously on evolving fixed-points and eigenvectors, dissipative interfaces between system and surround, or Spencer's catch-all of self-organized complexity (Kauffman 1993). It is fair to say that none of these frameworks inspired from physics is a better *lingua franca* for evolving organization and function than those inspired by biological and social sciences. Whatever their origin, all of these theories are either broad

descriptive languages with many parameters and few constraints, or narrow predictive models that seem difficult to extend to other problems.

In the theory of knowledge as maze search, a robust framework from mathematical logic and computer science, Darwinian survival and reproduction are cashed out as mapping unfamiliar paths as means of preserving maps of familiar ones. For many, the idea of search without any definite object seems paradoxical. However, the logic of active exploration, and basic problems of search complexity, are no different for autonomous, or living systems that construct maps, marked with plans of present behavior, for the purpose of map-preservation, and non-autonomous adaptive systems that construct them to attain or maintain some externally mandated goal.

Not to bury the lede, biological species are no longer the sole example of living systems that explore immediate affordances, and invent intermediate ones, for no higher purpose than continuity of these maps and plans. The emergence of natural language augmented nonverbal demonstration with spoken instruction, and later, written instruction. Relying on texts, not just individual memories, for their charters and working plans, writing allowed social organizations to outlive their human agents. From churches to nation states to transnational corporations, perduring social organizations have durable principles and more ephemeral procedures, *genomes and epigenomes in all but name*, that greatly exceed the individual powers and memories of human agents contracted to advance their interests (Simon 1947; Weber 1965; Galbraith 1967). These organizations explore and exploit their affordances on the timescale of human history, not human lifespans, *for no higher purpose than preserving and propagating their institutional memories*. Thus, highly autonomous, multigenerational social organizations are living individuals in the textual world in the same dynamical sense that ribogene communities were living individuals in the RNA world.[33]

There are social analogs of basic phenomena of multilevel coopetition discovered in biology. In a social analog of *genomic conflict*, human agents can violate enlistment oaths, or breech employment contracts, to game any social organization (Veblen 1899). In a social analog of *germ and soma*, lower-level organizations, like schools and

[33] To be sure, neither social organizations nor polymer communities are conscious, that peculiar phenomenon and useful function of the vertebrate cerebrum, that excites neuroscientists and confounds philosophers (Hofstadter 2007; Dennett 2017; Pinker 2022; Koch 2024).

businesses, have lesser autonomy, lesser perdurance, and greater reliance on the knowledge and purpose of their human agents, as well as greater acceptance of vicarious ends and constrained means imposed from higher-level governance (Campbell). In a social analog of horizontal exchange of knowledge between organizations or lineages, some cultural institutions are generally useful means, abstracted from particular ends, that can be acquired and used by any sufficiently capable social organization. Thus, sundry organizations apply general theories of basic sciences to their specific plans of say treating disease, building cars, or waging war (Roger Bacon). Finally, in a social analog of the mobility of self/nonself recognition, religious creeds and scriptures define orthodoxy/heterodoxy for churches that otherwise share no organizational history or synod.

The project of life science is to understand the organization of living systems, that is, their relations of mechanism and function. Analyzing processes into states and their transitions, biological mechanisms are commonly cashed out as say biopolymer chemistry, intracellular regulation, and intercellular communication. Identifying the purpose of these mechanisms, biological functions are cashed out as say homeostasis, immunity, motivated behavior, survival and reproduction, and ultimately, evolution. Cellular processes, i.e., metabolism and polymer synthesis, are mechanistically simple, but their contributions to survival and reproduction, and ultimately evolutionary search, are indirect and intertwined. Conversely, our cerebral processes, i.e., planning, acting, and learning, are mechanistically complex, but their contributions to achieving immediate and even long-range goals are direct and well-sorted.

Francis Bacon noticed that crucial tests of rival explanations, made by planned experiment, are necessary for scientific discovery. In his parable of ants, spiders, and bees from *Novum Organon*, Bacon explained this back-and-forth relation between theory and observation: “The men of experiment are like the ant, they only collect and use; the reasoners resemble spiders, who make cobwebs out of their own substance. But the bee takes a middle course: it gathers its material from the flowers of the garden and of the field, but transforms and digests it by a power of its own” (Bacon 1620). Obviously, this entails an experimental interface of actuators and sensors to manipulate and observe any process. Less obviously, it entails active learning through planned

experiment (Popper 1959; Angluin 1987). Comparing two explanations of overlapping phenomena, these alternatives are made strict rivals by framing both in a common experimental interface, and identifying some crucial test where they make opposite predictions, or the weaker theory fails to predict.

After the revolution of physical sciences in the 17th and 18th centuries, and the consilience of physics, chemistry, and biology in the 19th and 20th centuries, there was no place left for vitalism of energy, or dualism of matter in natural science. Still something important was missing. Beginning in mid 20th century, a new consilience, *the theory of discrete processes*, has unified descriptions of dynamical behavior from chemistry and biology, to electrical engineering, computation and language. Whereas continuous processes were the glory of science since Newton, intelligent behavior is better explained with logic and algebra, unadorned with probability and calculus. Discrete processes go by as many names as there are disciplines, but their generic description as sequences of tests (sensing states of the surroundings) and actions (changing states of the surroundings) was abstracted from reaction pathways and catalysis in chemistry, as well as communication and control in electrical engineering (Kleene; Moore 1956; Petri 1962; Milner).

The basic language of discrete processes is learned in a single lecture hour, but understanding the algorithms and complexity of searching simple mazes, much less ones with distributed states, nested levels, or concurrent processes, has spanned a scientific century, and unified a great deal of computer science, engineering, logic and linguistics. Here we use *maze* as a useful metaphor, and convenient shorthand, for discrete processes, such as finite automata, with an interface of sensors to perform Boolean tests on states, and actuators to cause transitions between states (Turing 1936; Kleene 1952; Moore 1956; Harel et al 2001). In this metaphor, states and partial states are places in the maze, while transitions are passages (generally one-way) from one place to another. A maze search algorithm marks familiar paths to exploit and curious ones to explore on the same working *map*. Using this map to travel the maze, a typical travel *plan* includes tests to select alternative routes along the way, as well as less explored places where familiar paths break down, with only some hints of opportune paths to explore further, or suspected hazards to avoid if possible.

In early AI and cognitive science, intelligent behavior was modeled as different activities, *viz*. communication, perception, concept formation, learning, memory, movement, planning, and reasoning (Nilsson 1980). Confronting this fragmentation, Allen Newell and Herbert Simon framed the central tasks of intelligent systems as search problems, and identified some common search strategies, *viz*. heuristics, satisficing, multilevel decomposition, and small world networks (Simon 1969; Newell & Simon 1972). Later researchers repeated the call for unified theories of cognition, or study of intelligent agents as a whole (Newell 1990; Russell & Norvig 1995). But what they mostly meant were uniform implementations of miscellaneous activities, say as production systems of rewrite rules, not universal principles, much less actual algorithms of goal-directed search.

Even as the profane wing of AI and cognitive science championed by Newell and Simon explored a trove of search hacks, its sacred wing championed by John McCarthy continued the original quest of Thomas Hobbes, Gottfried Leibniz, George Boole, and later logicians for the laws of thought (Hobbes 1651; Boole 1854; Strickland & Lewis 2022). Concept formation, pattern recognition, and similar models of machine learning encountered the same paradoxes in inductive generalization from examples and counterexamples that had confounded earlier philosophers and logicians (Mills 1843; Carnap 1966). Meanwhile, despite provably sound and complete rules of deductive inference, theorem-provers, goal-planners, and similar models of automated reasoning struggled to go beyond the simpler problems.

From polymer life to human cognition, natural language and artificial intelligence, the common obstacle to a logic of discovery was the conventional separation of learning and reasoning (James 1890). More and more, researchers called for a logic of informed guesswork that unified induction and deduction, which C. S. Peirce called *abduction*, applicable to problems of creative reasoning from gene invention and concept formation to scientific hypotheses and their experimental tests (Peirce 1878). Remarkably, computer science, not philosophy of science, nor natural sciences such as biology or psychology, was first to formulate discovery as learning through successive rounds of planned experiment. Demonstrating a tractable solution to Moore's problem of exploring finite automata, Dana Angluin proved that maze learning

is aimless without planned trials, and planning is intractable without active experimentation (Moore 1956; Conway 1971; Angluin 1987).

In the theory of active learning and reasoning as *maze search*, adaptive systems map their surroundings, marking their current plans on these maps to exploit resources and explore opportunities, improving these maps and updating their plans as they follow marked paths through the maze. Maze search is at once a general theory of knowledge and a useful one. It is general in so much as it encompasses diverse theories of adaptation in natural and artificial systems, and can readily formulate the sundry search hacks found in biology, organization science, and AI. It is useful because it explains the emergence of new search levels in a hierarchical logic of discovery.

Proffering the false choice of *realism* versus *nominalism*, early philosophers problematized the epistemological interface between subject and object. Any interface that mediates our search, they argued, imposes conventions and constrains actions, distorting or obscuring the essential features of reality, and limiting the accessible states. Nineteenth century scientists skirted about the function of search interfaces, *viz*. acquisition of knowledge by the system of its surroundings for some purpose. Ignoring purpose and agency of the system, the interface was reduced to its Newtonian description as reciprocal interactions of system and surrounding through exchange of matter and energy. Twentieth century scientists tried to reintroduce the elided knowledge and purpose through the backdoor of thermodynamics and information theory. Here adaptive systems have *dissipative interfaces* that preserve and increase their organizational complexity at the expense of their surroundings (Schrodinger 1944; Shannon & Weaver 1949; Prigogine 1978; Kaufmann 1993). Satisfied that statistical measures of complexity avoided the pratfalls of dualism, few noticed they had discarded the nature and emergence of meaning and purpose, the real baby of evolution, along with the bath waters.

It is not, of course, that our surroundings are mazes *in essence*, only that adaptive systems naturally interpret, or *map* these affordances in terms of their process interface of tests and actions. As Jacques Herbrand first demonstrated for the universe of first-order models, this interface, here a given set of predicate and function symbols, frames the search space of possible worlds, allowing us to discover, or otherwise define the

actual one (Herbrand 1931). Unfortunately, philosophers, logicians, linguists, and early AI researchers, overlooked the emergence of search interfaces, and assumed there was just one interface, given and fixed *before an agent began to explore*. This error fueled the hoary debates over realism and nominalism, and now armchair conceits of understanding the brain through mental introspection, or understanding the hardware and software behind our word processor from experience with the user interface of keyboard and console.

Adaptive systems have two different logical forms of learned purpose: PRE goals and LOWER-LEVEL goals. Showing how FINAL goals beget earlier and earlier ones, Aristotle explained unilevel planning as means-ends analysis, or *reasoning backward from ends to means* (Russell & Norvig 1995). Showing how HIGHER-LEVEL processes beget lower and lower ones, others explained multilevel planning as *reasoning downward from specification to implementation* (Campbell 1974; ref). Whereas PRE goals for reasoning backward emerge in both unilevel and multilevel maze search, LOWER-LEVEL goals for reasoning downward emerge in multilevel search from invention of new affordances, or process interfaces.

Darwinism identified the ultimate end of living species as preservation and propagation of their evolving genome, that is, map *continuity tout court*. Having struggled to banish any other *teleos* from their science, biologists were slow to recognize learned purpose. Thus, the early theory of reinforcement learning in S-R psychology suffered the same absence of planned variation and learned goals as the contemporary theory of natural selection (Thorndike 1911). To banish explicit PRE goals and planning from animal behavior, behaviorists modeled innate instincts and acquired habits as reflex chains of stimulus-response or S-R links, each response creating an immediate stimulus for the next one. In Figure 19-1 we show a simple reflex as a process with one Boolean test (A) for its adequate stimulus, and one action (P) for its elicited response. In this, and other process diagrams below, we show **tests** in red, and **actions** in green (Petri 1962). For convenience, we overload the symbols for tests, using say the same letter A for the Boolean variable, and its values, A or a. Thus, as values, upper and lower case letters show the Boolean holds or does not hold, without using an explicit negation operator (De Morgan). As explained below, we also overload

the symbols for actions, using say the uppercase letter P for performing that action, and the lowercase letter p for skipping it.

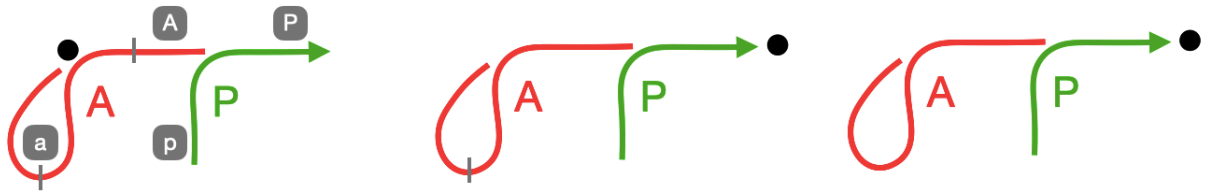


FIGURE 19-1 SIMPLE REFLEX ARC

On the left in Figure 19-1, the control token (filled circle) marks the PRE-state with the process monitoring its stimulus. If test A holds immediately, so the guard (single-bar) is satisfied, then action P is performed, and the process terminates with the token in the POST-state (middle Figure 19-1). If test A fails, it is tried again, and again, *ad nauseum*. The process language does not worry about when, only that the test will be repeated, nor does it assert that the test will ever be satisfied for the process to terminate. But once it is satisfied, action P will be performed and the process will terminate (left Figure 19-1). Terminating runs of this simple process have the form (a)*AP as Kleene regular expressions. Condition A is tested in the PRE-state, but not retested in the POST-state, that is, it may stay true, or may turn false when action P is performed. Meanwhile, other significant features may become true, stay true, become false, or stay false, but they are not explicitly tested in this process.

In *Cybernetics, or Control and Communication in the Animal and the Machine*, Norbert Wiener generalized feedback, homeostasis, and other regulatory concepts, from animal behavior and physiology, to human-made machines (Wiener 1948). In a thermostat, for sample, one and the same test is the stimulus to action, and the goal of behavior. In *Plans and the Structure of Behavior*, George Miller and colleagues, introduced an elementary *while*-program called a TOTE unit, mnemonic for TEST-OPERATE-TEST-EXIT (Miler et al 1960). Here the same PRE-test that is the stimulus to action (OPERATE), is the POST-test allowing the process to terminate (EXIT). Terminating runs have the form (AP)*a, that is, the process repeats action P so long as test A holds, and terminates once test A fails. In particular, it exits without any action if test A fails on

the first try (middle diagram in Figure 19-2). In a sense, the purpose of the TOTE unit is to turn A off or leave it off. Whether it never, sometimes, or always succeeds and exits, and whether action P somehow contributes to success, depend of course, on the relation of action P to condition A in the actual maze or affordance.

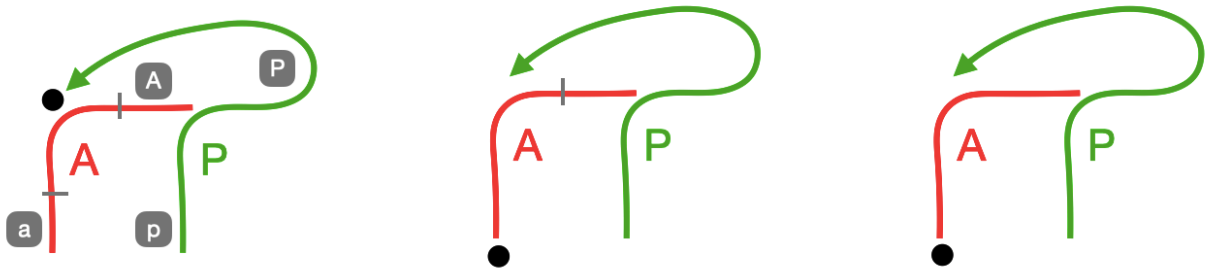


FIGURE 19-2 SIMPLE *while*-PROGRAM OR TOTE UNIT

A thermostat-controlled furnace, or any other homeostat, can be described as a modified TOTE unit (Figure 19-3). Whereas the TOTE unit terminates once test A fails, the simple homeostat runs forever without halting. Test A monitors whether the room is too cold. If so, furnace P runs, and the control token returns to the start. If the room is not too cold, the furnace run is skipped, and the control token returns to the start. Thus, runs of the system, *viz*. thermostat-controlled furnace, coupled to its surrounding, *viz*. heated room in winter, have the form …(AP)*(ap)*(AP)*(ap)*… where products in the Kleene stars denote zero or more cycles of the simple test-action reflexes (AP) or (ap), respectively.

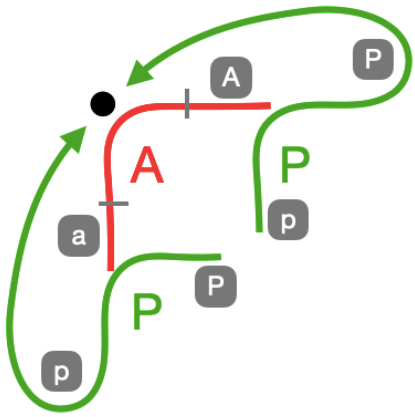


FIGURE 19-3. SIMPLE HOMEOSTAT

The process interfaces that adaptive systems use to map their surroundings, and plan their own behavior, continually evolve through synthesis of yet higher-level tests and actions, and analysis of "elementary" tests and actions into yet lower-level, or finer grain observations and manipulations. In this emergent hierarchy of interfaces and their search maps, plans of one level can *call upon* plans at levels below, or *return to* plans at levels above. Thus, multilevel maze search combines reasoning backward from POST goals to PRE goals, and reasoning downward from HIGHER-LEVEL to LOWER-LEVEL plans.

Seeking a common language for HIGHER-LEVEL specifications of function and LOWER-LEVEL implementations of the corresponding procedure, Tony Hoare introduced a more general process than the S-R arc or TOTE unit (Hoare 1969). In what became known as Hoare triples, two different Boolean tests are used, test A for the PRE-state, and test B for the POST-state. The Hoare triple A{P}B asserts that if test A holds in the PRE-state then test B holds in any POST-state following from action P. For Aristotle's means-ends analysis in the unilevel maze, this cashes out: to attain goal B, first establish its PRE-goal A, and then take action P. In his logic of programming, Hoare showed how to specify functions using Hoare triples, and then derive procedures that meet these specifications. Although his rules of program derivation are sound and complete, like automata learning without planned experiment, they have intractable complexity. That is, no theory of program derivation as pure downward reasoning from HIGHER-LEVEL specification in relational form to LOWER-LEVEL procedure in functional (executable) form was ever going to work beyond toy problems.

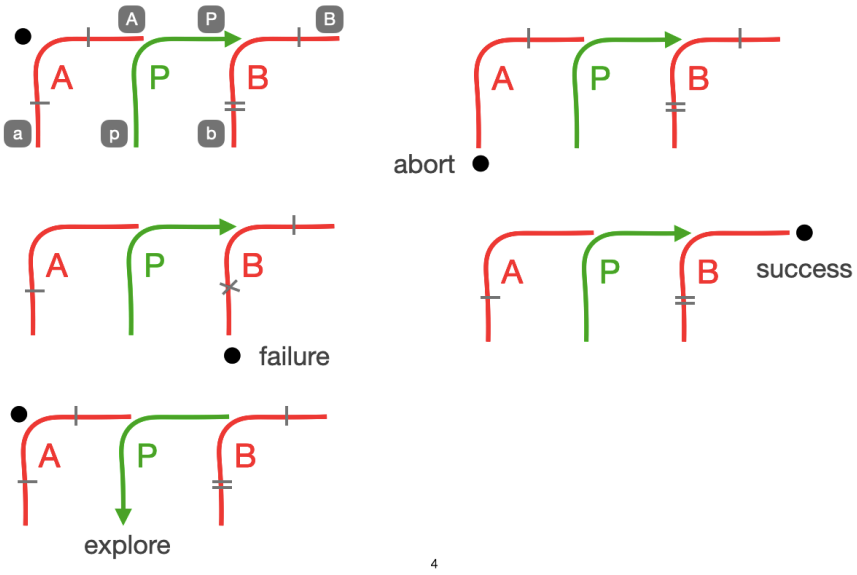


FIGURE 19-4 HOARE TRIPLES

Although our surroundings are not mazes *simpliciter*, we explore and exploit them as such, that is, our maps and plans are processes of tests and actions. Despite their simplicity, Hoare triples, or the equivalent process diagrams, are adequate and perspicacious for formulating multilevel maze search (Hedgecock, Proença & Mastwal, in preparation). Unlike S-R arcs and TOTE units that have one exit, or homeostats that never terminate, Hoare triples have three exits. First, when PRE-condition A fails, the process *aborts* with no action is taken (Figure 19-4). Second, when A holds, and action P is performed, test B may hold, and the process ends in *success*. Third, when A holds, and action P is performed, test B may fail, and the process ends in *failure*. We use the double-bars (=) on a test to indicate an impossible value, that is, one never before witnessed, and the crossed-bars (X) to indicate the value has now been witnessed, marking a path to update now, and perhaps explore later.

Process diagrams depict the maps and plans of the system, that is, their working model of surroundings as affordances to explore and exploit. The system gets the values of tests from the process interface, but those values are put to the interface by the surroundings. These handshake interactions called *synchronization* of system and surrounding are indicated by single-bars (—) on tests. The system also puts actions to the process interface that can change the state of the surroundings. Thus, states and their changes are known partly, but directly, through tests, and more fully, but indirectly,

through comparing the recent sequence of tests and actions to the known paths in our working map.

Sequences of attended tests and intended actions are planned by the system. One and the same map can be marked with several plans to reach different goals, or avoid changing hazards. Boolean tests are choice-points in the process diagram, but actions are not. Curiously, it is convenient to overload the process symbols for actions, using the same symbol in upper and lower case to indicate performing (P) or skipping (p) that same action, respectively (Kozen). In our process diagrams we use leftward arcs to perform and downward arcs, respectively, to skip the labelled action. For purpose of manipulating our surroundings, skipped actions have no effects (beyond the time for the control token to move past), but for active learning of mazes, they allow perspicacious records of past runs for planning future experiments. In Figure 19-4 we show a simple plan that skips action P to explore further tests and actions.

Back to animal life, today biologists ascribe purpose to sundry mechanisms from homeostats that maintain internal conditions or metabolite pools, to immune responses that detect and destroy foreign cells and polymers, cell cycle checkpoints that delay the cell cycle at various stages to allow for repair of DNA damage, or shunt cells to apoptosis, and myriad other forms of cell and polymer quality control. Sometimes these functions are identified first (as from a loss-of-function trait in Mendelian genetics) and their mechanism sought. Other times mechanisms are encountered first (as in RNAi) before their functions are understood. We give just one example of multilevel regulation of animal behavior: For temperature regulation and energy homeostasis, the hypothalamus specifies the goals or set-points to innate brainstem mechanisms, and relays copies of these HIGH-LEVEL triples of tests and actions to the limbic cerebrum, the highest level of our cognitive map. There these triples direct our lifelong individual learning of LOWER-LEVEL cerebral plans as supplemental means of temperature regulation and energy homeostasis, *e.g.*, finding likely sources of food, water, and nesting materials, warm and cool places in the environment, etc. But even our innate regulatory processes, *viz*. triples of hypothalamic tests and brainstem actions, are vicarious means to the animal ends of survival and reproduction acquired through natural selection of the species.

In exploring an unfamiliar maze, we follow tentative plans with intermediate tests for alternative paths, and occasional interrupts for unexpected failures, or unpredicted successes, that call for reinforcement learning and replanning. William James noticed that with repeated practice our behavior becomes rapid and predictable, or one fixed sequence of actions from start to finish (James 1890). In terms of maze search, our plans have narrowed to a single path on the map for a routine run through the maze. This was most obvious with overtraining on a predictable task in an impoverished environment (Thorndike 1913; Watson 1919; Skinner 1938). Many noticed these well-practiced plans, or acquired habits of individuals had an uncanny similarity to innate instincts of the species, and posited an analogy, if not actual connection between habit formation and instinct evolution (Lamarck, Baldwin). The appealing idea that learned habits could become hereditable instincts was abandoned after Weissmann introduced the firewall between germline and soma including the nervous system.

Habit formation in animal learning, and the underlying memory consolidation from limbic to neocortical areas, reflects the logical redistribution, or *downward consolidation*, of plans from higher to lower levels in the cognitive map. Downward consolidation is the logical imprimatur of repeated practice of a multilevel plan in similar surroundings. Examples of consolidation range from reorganization of mining operations from exploration (mineral prospecting) to exploitation (mineral extraction), to derivation/compilation of programs from specification to procedure, and gene tenure from heterochromatin to euchromatin (section 18). Unlike unilevel search with one fixed process interface with the maze (affordance) to map, in multilevel search, there is an evolving hierarchy of affordances to explore. Each affordance is at once a maze to be mapped by search levels above, and a map recording past searches of maze levels below. In downward consolidation, needed tests are moved toward the initial states, and needless ones are pruned entirely. In a consolidated plan for a familiar maze, any selected sequence of lower-level actions, once begun, has few if any interruptions. One useful metaphor likens consolidated plans to the trajectories of ballistic missiles, aimed accurately from initial tests, and then needing no midcourse corrections, and more tentative plans of less familiar mazes to cruise missiles requiring intermediate tests, or

feedback through GPS updates, and selection of alternative paths for midcourse correction.

Philosophers have characterized downward consolidation in multilevel maze search as a transformation of knowledge from declarative to procedural forms. Biologists have studied the origin and nature of these search levels themselves. Like bread in a Dagwood sandwich, new search levels can be placed at top or bottom of the map hierarchy, but more generally they can insert anywhere like a slice of cheese or lunchmeat. Unlike the sandwich metaphor, any hierarchy of nested maps need only be partially ordered. The ease and challenges of inserting new affordances is seen in the continual invention and rapid prototyping of new protein-coding genes anywhere in HIGHER-LEVEL regulatory mechanisms, or LOWER-LEVEL metabolic pathways. Likewise, in the vertebrate cerebrum, new cortical areas are inserted or removed, and old ones subdivided or merged, at all levels from limbic to executive/perceptual to motor/sensory.[34] Most remarkable of all, affordances are mobile within the multilevel map, so that new levels build upward and downward from earlier levels, but they can also detach to reattach elsewhere.

Back to RNA life, tests and actions were cashed out as riboswitches and ribozymes, respectively, responding to physical or chemical conditions, and acting on RNAs or other chemical affordances. For cellular life, biochemistry reserved the suffix *-ase* for the actions of enzymes and ribozymes that make or break covalent bonds, but cell biology needed a more general concept of regulation. Computer science demonstrated the usefulness of abstracting processes, *viz*. switching circuits and sequential circuits with feedback and memory, from kinetic and energetic details of their implementation (Shannon 1936; Turing 1936). Significant events just had to be "fast enough", "irreversible enough", or "deterministic enough" for the process at hand. Thus, the kinetic concept of catalyst became the concept of reusable instruction for conditional transition between definite states. In the deterministic case, an irreversible transition occurs just in case the PRE-state satisfies the conditions of the current

[34] Interestingly, a progression of insertions occurs during the course of development as new intermediate areas come online, as judged by myelination of their axonal outputs, and BOLD measures of their activity. To be sure, the order of myelination in development (ontogeny) has no known relation to the order of appearance of cerebral areas in evolution (phylogeny).

instruction. For process logic and algebra, it is convenient to factor conditional actions, aka reflexes, into a test for input-from the surroundings, composed with an action for output-to the surroundings (Harel et al 2001). Such tests and actions together define the basic process interface.

The unification of Mendelian genetics as the level of function, and biochemistry/cell biology as the level of process, produced the new science of molecular biology (Mayr 1961; Judson 1979; Watson et al 2016). Its success whets our appetite for a similarly profound unification of function and process in cognitive science. We have just begun to understand the multilevel organization of the vertebrate cerebrum, and how its search functions, *viz*. planning, acting, and learning, are realized neuronal circuits. Beyond the cerebrum *in toto*, our unique language faculty that underlies human society and science, and how it emerged within the vertebrate cerebrum, are objects of the greatest interest. Our working hypothesis is that we abstract (*parse)* mid-level paths through the cognitive map, what Thomas Hobbes dubbed conscious trains of thoughts, as tacit or spoken discourse (Hobbes 1651; Hedgecock, Proença & Mastwal, in preparation). From these sentences, the speaker, or any competent listener (aka linguistic conspecific), reinstantiates (*bisimulates*) those paths elsewhere within the cognitive map.

Natural language provides a means of replicating and unifying fragments of the cognitive map both within and between individuals. This explains two hoary chestnuts of mental philosophy, how beasts have conscious thoughts, yet no language, and why discourse is vicarious to our conscious thought. No one doubts the emergent advantage of language for human communication and society, but was there an immediate advantage of “inner speech”? Perhaps Hobbes was close to the mark in thinking that planning from given ends to likely means is the common mode of reasoning in man and beast, while planning from given means to likely ends entails creative combination of map fragments mediated through natural language: “The train of regulated thoughts is of two kinds: one, when of an effect imagined we seek the causes or means that produce it; and this is common to man and beast. The other is, when imagining anything whatsoever, we seek all the possible effects that can by it be

produced; that is to say, we imagine what we can do with it when we have it" (Hobbes 1651, Part I Chapter III).

Implementing processes *in silico*, we foresee handovers to artificial life and intelligence as profound and abrupt as the breakout of protein translation, or takeover of DNA chromosomes. Meanwhile, genome engineering, a new level of planned variation, has begun to combine ancient inventions of polymer life and modern discoveries of biological science. From gene therapy to optogenetic regulation of behavior, this new search level dramatizes the potentially seamless mixis of our sciences and our species. For a precedent of our human brain and social institutions feeding back on their own development and instruction, we can turn to the profound evolutionary feedback of coded proteins on the processes of protein synthesis, genome reproduction, and gene expression. Finally, in what surely will be the greatest mixis of knowledge since the origin of life on Earth, contacting species of extraterrestrial origin, and contracting their/our means and our/their ends, shall combine learning and discovery across the multiplicity of worlds foreseen by Nicolas de Cusa and Giodano Bruno.

All of this makes abundantly clear that the problem of life and evolution is one and the same as the problem of knowledge and discovery itself. No doubt, simple models of *blind variation and selection* that cannot explain *the arrival of the fittest* in biological evolution, cannot explain the invention of scientific hypotheses and their crucial experiments. In fact, biological evolution, the human brain, and scientific discovery all work remarkably better than our working explanations of these phenomena. Nearly a century ago, on the tercentenary of Francis Bacon, C. D. Broad contrasted the many successes of inductive sciences and the hoary paradoxes of inductive reasoning as the *glory of science*, and the *scandal of philosophy*, respectively.[35] A more pithy statement of the gap between these explanada and our present explanans, attributed to physicist Richard Feynman, is "Philosophy of science is as useful to scientists as ornithology is to birds."

---

[35] "May we venture to hope that when Bacon's next centenary is celebrated the great work which he set going will be completed; and that Inductive Reasoning, which has long been the glory of Science, will have ceased to be the scandal of Philosophy?" (Broad 1926 p67).

Evolutionary transitions, no less than scientific breakthroughs, require effective plans of exploration (planned variation) that realize the multilevel logic of search. Our descriptive theories of search heuristics, abductive bias, or genomic and epigenomic variation intimate the phenomena, but do not formulate their search principles. With another centenary upon us, we must kick this can down the road a bit more, awaiting a clear and simple formulation of the logic of multilevel maze search, and its algorithmic realizations, that explains the arrival of the fittest and their natural selection, as readily as the construction of scientific hypotheses and their experimental tests. Polymer life, eukaryotic (sexual) reproduction, the vertebrate brain, and our language faculty have thrown down the gauntlet, showing that maze search works remarkably well, revealing its processes, and hinting at its principles.

Acknowledgements

This paper grew out of an annual lecture on the RNA world to our human genetics class at Johns Hopkins University; we wish to thank over a decade of students for letting us explore these ideas. We also wish to acknowledge our teachers and scientific mentors: Maurice Bessman, Marshall Blaufarb, Glenn Flowers, Phil Hartman, Ed Lewis, Hill, Ray Owen, Dick Russell, Aaron Shatkin, John Sulston and John White. Finally, we thank our colleagues Matthew Eroglu, Bertrand Garcia-Moreno, Virginia Hedgecock, Blake Hill, Oliver Hobert, Harald Hutter, Edward Kipreos, Robert Leheny, Carolyn Norris, Mark Van Doren, Bruce Vogel, and Haiqing Zhao for encouragement to tackle something so inherently speculative, as well as outside our comfort zone. Most of all, this paper is dedicated to our families, Emma Louise and Nathan David, who give our lives meaning.

References

Abramov & Mojzsis (2009).

Agmon (2022).

Agmon I, Bashan A, Zarivach R & Yonath A (2005). Symmetry at the active site of the ribosome: structural and functional implications. *Biol Chem* 386, 833-844.

Agmon I, Davidovich C, Bashan A & Yonath A (2009). Identification of the prebiotic translation apparatus within the contemporary ribosome. Nat Prec …

Agris PF, Eruysal ER, Narendran A, Vare VYP, Vangaveti S & Ranganathan SV (2018). Celebrating wobble decoding: half a century and still much is new. *RNA Biology* 15, 537-553.

Ahl et al (2015).

Ainfinsen (1973).

Ainfinsen et al (1961).

Aleksashin NA, Leppik M, Hockenberry AJ, Klepacki D, Vazquez-Laslop N, Jewett MC et al (2019). Assembly and functionality of the ribosome with tethered subunits. Nat Commun 10, 930-.

Alkatib S, Scharff LB, Rogalski M, Fleischmann TT, Matthes A, Seeger S, Schottler MA, Ruf S & Bock R (2012). The contributions of wobbling & superwobbling to the reading of the genetic code. *PLOS* 8, 1-16.

Alva V, Soding J & Lupas AN (2015). A vocabulary of ancient peptides at the origin of folded proteins. *eLife* 4, e09410.

Amort et al (2007).

Anastasi C, Crowe MA, Powner MW & Sutherland JD (2006). Direct assembly of nucleoside precursors from two- and three-carbon units. *Angew Chem Int Ed* 45, 6176-6179.

Andachi Y, Yamao F, Muto A & Osawa S (1989). Codon recognition patterns as deduced from sequences of the complete set of transfer RNA species in *Mycoplasma capricolum*. Resemblance to mitochondria. *J Mol Biol* 209, 37-54.

Atkins JF & Bjork GR (2009). A gripping tale of ribosomal frameshifting: extragenic suppressors of frameshift mutations spotlight P-site realignment. *Microbiol Mol Biol Rev* 73, 178-210.

Attwater J, Raguram A, Morgunov AS, Gianni E & Holliger P (2018). Ribozyme-catalysed RNA synthesis using triplet building blocks. *eLife* 7, 1-25.

Attwater J, Wochner A & Holliger P (2013). In-ice evolution of RNA polymerase ribozyme activity. *Nat Chem* 5, 1011-1018.

Attwater J, Wochner A, Pinheiro VB, Coulson A & Holliger P (2010). Ice as a protocellular medium fro RNA replication. *Nat Commun* 1, 76-.

Attwater J, Wochner A, Pinheiro V, Coulson A & Holliger P (2010). Ice as a protocellular medium for RNA replication. Nature Comm ##, 1-9.

Balasanyants SM, Aleksandrova EV & Polikanov YS (2021). The role of release factors in the hydrolysis of ester bond in peptidyl-tRNA. *Biochemistry* 86, 1122-1127.

Baltimore D (1970). RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* 226, 1209-1211.

Ban N, Nissen P, Hansen J, Moore PB & Steitz TA (2000). The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science* 289, 905-920.

Bao et al (2022).

Bare GA & Joyce GF (2021). Cross-chiral, RNA-catalyzed exponential amplification of RNA. J Am Chem Soc 143, 19160-19166.

Bartel DP & Szostak JW (1993). Isolation of new ribozymes from a large pool of random sequences. *Science* 261, 1411-1418.

Beck & Moll (2018).

Becker S, Thoma I, Deutsch A, Gehrke T, Mayer P, Zipse H & Carell T (2016). A high-yielding, strictly regioselective prebiotic purine nucleoside formation pathway. *Science* 352, 833-836.

Becker S, Feldmann J, Wiedmann S, Okamura H, Schneider C, Iwan K, Crisp A, Rossa M, Amatov T & Carell T (2019). Unified prebiotic syntheses of pyrimidine and purine RNA building blocks driven by wet-dry cycles. *Science* 366, 76-82.

Beckert B, Kedroy A, Sohmen D, Kempf G, Wild K, Sinning I, Stahlberg H, Wilson DN & Beckmann R (2015). Translational arrest by a prokaryotic signal recognition particle is mediated by RNA interactions. *Nature Structural Mol Biol* 22, 767-773.

Been MD & Cech TR (1988). RNA as an RNA polymerase: net elongation of an RNA primer catalyzed by the Tetrahymena ribozyme. *Science* 239, 1412-1416.

Benner SA, Kim HJ & Carrigan MA (2012). Asphalt, water, and the prebiotic synthesis of ribose, ribonucleosides, and RNA. *Accounts Chem Res* 45, 2025-2034.

Benner SA, Kim HJ & Biondi E (2018). Mineral-organic interactions in prebiotic synthesis. The discontinuous synthesis model for the formation of RNA in naturally complex geological environments. In *Prebiotic Chemistry & Chemical Evolution of Nucleic Acids* (ed Menor-Salvan C) Springer International.

Bernal (1951).

Bernhardt HS (2012). The RNA world hypothesis: the worst theory of the early evolution of life (except for all the others. *Biol Direct* 7, 23-.

Bernstein M (2006). Prebiotic material from on and off the early Earth. Phil Trans R Soc B Biol Sci 361, 1689-1702.

Bjork GR, Wikstrom PM & Bystrom AS (1989). Prevention of translational frameshifting by the modified nucleoside 1-methylguanosine. Science 244, 986-989.

Bhattacharyya S & Varshney U (2016). Evolution of initiator tRNAs and selection of methionine as the initiating amino acid. *RNA Biology* 13, 810-819.

Bohnsack & Bohnsack (2019).

Bokov K & Steinberg SV (2009). A hierarchical model for evolution of 23S ribosomal RNA. *Nature* 457, 977-980.

Bonitz SG, Berlani R, Coruzzi G, Li M, Macino G, Nobrega RG, Nobrega MP, Thalenfeld BE & Tzagoloff A (1980). Codon recognition rules in yeast mitochondria. *Proc Natl Acad Sci USA* 77, 3167-3170.

Bose T, Fridkin G, Davidovich C, Krupkin M, Dinger N, Falkovich AH, Peleg Y, Agmon I, Bashan A & Yonath A (2022). Origin of life: protoribosome forms peptide bonds and links RNA and protein dominated worlds. *Nucleic Acids Res* 50, 1815-1828.

Bousset L, Mary C, Brooks MA et al (2014). Crystal structure of a signal recognition particle Alu domain in the elongation arrest conformation. RNA 20, 1955-1962.

Brännvall M, Kikovska E, Wu S & Kirsebom LA (2007). Evidence for induced fit in bacterial RNase P RNA-mediated cleavage. *J Mol Biol* 372, 1149-1164.

Breslow (1959).

Bretscher (1968).

Broad CD (1926). *The Philosophy of Francis Bacon*. Cambridge University Press.

Brunelle JL, Shaw JJ, Youngman EM & Green R (2008). Peptide release on the ribosome depends critically on the 2'-OH of the peptidyl-tRNA substrate. *RNA* 14, 1526-1531.

Burma et al (1961).

Burt & Trivers (2006).

Buskirk & Green (2017).

Butlerow (1861).

Cafferty et al (2018).

Cairns-Smith AG (1965). The origin of life and the nature of the primitive gene. *J Theoret Biol* 10, 53-88.

Cairns-Smith AG & Davies CJ (1977). The design of novel replicating polymers. In *Encyclopaedia of Ignorance* (editors Duncan R & Weston-Smith M). Pergamon Press, Oxford. pp 391-403.

Cairns-Smith AG (1982). *Genetic Takeover: and the Mineral Origins of Life*. Cambridge University Press, Cambridge UK.

Cairns-Smith AG & Hartman H (1986). *Clay Minerals and the Origin of Life*. Cambridge University Press, New York.

Campbell (1974).

Campbell JH (1991). An RNA replisome as the ancestor of the ribosome. *J Mol Evol* 32, 3-5.

Campbell TD, Febrian R, McCarthy JT, Kleinschmidt HE, Forsythe JG & Bracher PJ (2019). Prebiotic condensation through wet-dry cycling regulated by deliquescence. Nature Comm 10, 1-7.

Carbone et al (2021).

Carlson BA, Kwon SY, Chamorro M, Oroszlan S, Hatfied DL & Lee BJ (1999). Transfer RNA modification status influences retroviral ribosomal frameshifting. Virology 255, 2-8.

Carpenter FH (1960). The free energy change in hydrolytic reactions: the non-ionized compound convention. *J Am Chem Soc* 82, 1111-1122.

Caskey CT, Beaudet AL, Scolnick EM & Rosman M (1971). Hydrolysis of fMet-tRNA by peptidyl transferase. *Proc Natl Acad Sci USA* 68, 3163-3167.

Cech TR (1986). A model for the RNA-catalyzed replication of RNA. Proc Natl Acad Sci 83, 4360-4363.

Cech TR (2009). Crawling out of the RNA world. *Cell* 136, 599-602.

Chamberlin & Berg (1962).

Chan CW, Chetnani B & Mondragon A (2013). Structure and function of the T-loop structural motif in noncoding RNAs. *WIREs RNA* 4, 507-522.

Cech et al (1981).

Chillon I & Marcia M (2021). Self-splicing group II introns. In Ribozymes … 143-167.

Chyba C & Sagan C (1992). Endogenous production, exogenous delivery and impact-shock synthesis of organic molecules: an inventory for the origins of life. *Nature* 355, 125-132.

Chyba CF, Thomas PJ, Brookshaw L & Sagan C (1990). Cometary delivery of organic molecules to the early Earth. *Science* 249, 366-373.

Cobb AK & Pudritz RE (2014). Nature's starships. I. Observed abundances and relative frequencies of amino acids in meteorites. *Astrophys J* 783, 140-.

Cohen et al (2022).

Collins & Greider (1993).

Cojocaru R & Unrau PJ (2021). Processive RNA polymerization & promotor recognition in an RNA World. *Science* 371, 1225-1232.

Cojocaru R & Unrau PJ (2021). Phosphoryl transfer ribozymes. Ribozymes (ed Muller S, Masquida B & Winkler W). Wiley-VCH. pp331-357

Cojocaru R & Unrau PJ (2021). Phosphoryl transfer ribozymes. In Ribozymes … 331-358.

Coughlin DJ, Pleiss JA, Walker SC, Whitworth GB & Engelke DR (2008). Genome-wide search for yeast RNase P substrates reveals role in maturation of intron-encoded box C/D small nucleolar RNAs. *Proc Natl Acad Sci USA* 105, 12218-12223.

Crick FHC (1955/1958). On degenerate templates and the adaptor hypothesis. *Symp Soc Exp Biol* 12, 138-163.

Crick FHC (1956). On protein synthesis. *Symp Soc Exp Biol* 12, 138-163.

Crick FHC (1966). Codon-anticodon pairing : the wobble hypothesis. *J Mol Biol* 19, 548-555.

Crick FHC (1968). The origin of the genetic code. *J Mol Biol* 38, 367-379.

Crick FHC (1970). Central dogma of molecular biology. *Nature* 227, 561-563.

Crick FHC, Barnett L, Brenner S and Watts-Tobin RJ (1961). General nature of the genetic code for proteins. *Nature* 192, 1227-1232.

Crick FHC, Griffith JS & Orgel LE (1957). Codes without commas. Proc Natl Acad Sci USA 43, 416-421.

Crick & Orgel (1973).

Crick et al (1976).

Cronin (1989).

Cronin JR & Pizzarello S (1997). Enantiomeric excesses in meteoritic amino acids. *Science* 275, 951-955.

Damer B & Deamer D (2020). The hot spring hypothesis for an origin of life. Astrobiology 20, 429-452.

Dance (2023).

Dao Duc et al (2019).

Darwin C (1859). Origin of Species

Dawkins R (1976). *The Selfish Gene*. Oxford University Press.

Deamer D & Weber AL (2010). Bioenergetics and life's origins. Cold Spring Harbor Perspect Biol 2, ###.

Dedkova LM & Hecht SM (2019). Expanding the scope of protein synthesis using modified ribosomes. *J Am Chem Soc* 141, 6430-6447.

Demeshkina N, Jenner L, Yusupova G & Yusupov M (2010). Interactions of the ribosome with mRNA & tRNA. Current Opinion Struc Biol 20, 325-332.

Demeshkina N, Jenner L, Westhof E, Yusupov M & Yusupova G (2012). A new understanding of the decoding principle on the ribosome. *Nature* 484, 256-259.

Deng et al (2022).

Dever & Green (2012).

Dick TP & Schamel WWA (1995). Molecular evolution of transfer RNA from two precursor hairpins: implications for the origin of protein synthesis. J Mol Evol 41, 1-9.

Dickerson et al (1982).

Diener (1971).

Di Giulio (1992). On the origin of the transfer RNA molecule. *J theor Biol* 159, 199-214.

Di Giulio M (2004). The origin of the tRNA molecule: implications for the origin of protein synthesis. *J theor Biol* 226, 89-93.

Dill KA & Chan HS (1997). From Levinthal to pathways to funnels. Nat Struct Mol Biol 4, 10-19.

Djumagulov M, Demeshkina N, Jenner L, Rozov A, Yusupov M & Yusupova G (2021). Accuracy mechanism of eukaryotic ribosome translocation. Nature 600, 543-546.

Dobson CM (2004). Principles of protein folding, misfolding and aggregation. *Semin Cell Dev Biol* 15, 3-16.

Docter BE, Horowitz S, Gray MJ, Jakob U & Bardwell JC (2016). Do nucleic acids moonlight as molecular chaperones? *Nucleic Acids Res* 44, 4835-4845.

Doerfel LK et al (2013). EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. *Science* 339, 85-88.

Doolittle WF & Sapienza C (1980). Selfish genes, the phenotype paradigm and genome evolution. Nature 284, 601-603.

Doudna JA & Lorsch LR (2005). Ribozyme catalysis: not different, just worse. Nat Struct Mol Biol, 395-

Doudna JA & Szostak JW (1989). RNA-catalysed synthesis of complementary-strand RNA. *Nature* 339, 519-522.

D'Urso G, Guyomar C, Chat S, Giudice E & Gillet R (2022). Insights into the ribosomal trans-translation rescue system : lessons from recent structural studies. FEBS J ##, 1-12.

Eakin (1963).

von Ehrenstein, Weisblum B & Benzer S (1963). The function of sRNA as amino acid adaptor in the synthesis of hemoglobin. *Proc Natl Acad Sci USA* 49, 669-675.

Eigen M (1971). Self-organization of matter and the evolution of biological macromolecules. *Naturwissenschaften* 58, 465-523.

Eigen M & Schuster P (1979). The hypercycle: a principle of natural self-organization. Part A: emergence of the hypercycle. *Naturwissenschaften* 64, 541-565.

Eigen M & Winkler-Oswatitsch R (1981a). Transfer-RNA: the early adaptor. Naturwissenschaften 68, 217-228.

Eigen M & Winkler-Oswatitsch R (1981b). Transfer-RNA, an early gene? Naturwissenschaften 68, 282-292.

Eigner J, Boedtker H & Michaels G (1961). The thermal degradation of nucleic acids. *Biochim Biophys Acta* 51, 165-168.

Ekland, Szostak & Bartel (1995).

Ekland EH & Bartel DP (1996). RNA-catalysed RNA polymerization using nucleoside triphosphates. *Nature* 382, 373-376.

Erastova et al (2017).

Eschenmoser A (1999). Chemical etiology of nucleic acid structure. *Science* 284, 2118-2124.

Eschenmoser A (2004). The TNA-family of nucleic acid systems: properties and prospects. *Orig Life Evol Biosph* 34, 277-306.

Eschenmoser (2005).

Fahnestock S & Rich A (1971). Ribosome-catalyzed polyester formation. *Science* 173, 340-343.

Ferris JP (2005). Mineral catalysis and prebiotic synthesis: Montmorillonite-catalyzed formation of RNA. Elements 1, 145-149.

Ferris JP (2006).

Fialho DM, Roche TP & Hud NV (2020). Prebiotic syntheses of noncanonical nucleosides and nucleotides. *Chem Rev* 120, 4806-4830.

Fine & Pearlman (2023).

Fire A & Mello C

Flores et al (2004).

Forsythe et al (2015).

Fraenkel-Conrat (1956).

Frank & Gonzales (2010)

Freier SM, Kierzek R, Jaeger JA, Sugimoto N, Caruthers MH, Neilson T & Turner DH (1986). Improved free-energy parameters for predicitions of RNA duplex stability. Proc Natl Acad Sci USA 83, 9373-9377.

Fry M (2022). Crick's adaptor hypothesis and the discovery of transfer RNA: experiment surpassing theoretical prediction. Philos Theor Pract Biol 14, 11-31.

Fu Z, Indrisiunaite G, Kaledhonkar S, Shah B, Sun M, Chen B, Grassucci RA, Ehrenberg M & Frank J (2019). The structural basis for release-factor activation during translation termination revealed by time-resolved cryogenic electron microscopy. Nature Comm 10, 1-7.

Fujishima K, Sugahara, Tomita M & Kanai A (2008). Sequence evidence in the Rachael genomes that tRNAs emerged through the combination of ancestral genes. *PLoS ON*E 3, e1622.

Fujishima K, Sugahara J, Kikuta K, Hirano R, Sato A, Tomita M & Kanai A (2009). Tri-split tRNA is a transfer RNA made from 3 transcripts that provides insight into the evolution of fragmented tRNAs in archaea. *Proc Natl Acad Sci USA* 106, 2683-2687.

Fukuda et al (2020).

Fukushima K & Esaki H (2021). Theoretical study of the mechanism of ribosomal peptide bond formation using the ONIOM method. Chem Pharm Bull 69, 734-740.

Furth et al (1962).

Gade et al (2020).

Gamov G (1954). Possible relation between deoxyribonucleic acid and protein structures. Nature 173, 318.

Garside EL, Kent OA & MacMillan AM (2021). The spliceosome: an RNA-protein ribozyme derived from ancient mobile genetic elements. In Ribozymes … 169-191.

Geiduschek et al (1961).

Gibard C, Bhowmik S, Karki M, Kim EK & Krishnamurthy R (). Phosphorylation, oligomerization and self-assembly in water under potential prebiotic conditions. *Nature Chem* 10, 212-217.

Gilbert W (1986). Origin of life: The RNA world. *Nature* 319, 618.

Gierer & Schramm (1956).

Giurgiu C, Li L, O'Flaherty DK, Tam CP & Szostak JW (2017). A mechanistic explanation for the regioselectivity of nonenzymatic RNA primer extension. *JACS* 139, 16741-16747.

Goldman AD & Kacar B (2021). Cofactors are remnants of life's origin and early evolution. *J Mol Evol* 89, 127-133.

Gordon KHJ (1995). Were RNA replication & translation directly coupled in the RNA (+protein?) world? *J Theor Biol* 173, 179-193.

Gordon & Mikailowky (2021).

Gößringer M, Schencking I & Hartmann RK (2021). The RNase P ribozyme. In *Ribozymes* 1, 227-279.

Gottesman & Mustev (2019).

Goyal et al (2017)

Green R & Szostak JW (1992). Selection of a ribozyme that functions as a superior template in a self-copying reaction. *Science* 258, 1910-1915.

Greider & Blackburn (1989).

Griffith & Vaida (2012).

Grosjean H & Westhof E (2016). An integrated, structure- and energy-based view of the genetic code. *Nucleic Acids Res* 44, 8020-8040.

Gruber AR, Lorenz R, Bernhart SH, Neubock R & Hofacker IL (2008). The Vienna RNA websuite. *Nucleic Acids Res* 36, W70-W74.

Grunberg-Manago M, Ortiz P & Ochoa S (1956). Enzymic synthesis of polynucleotides. I. polynucleotide phosphorylase of Azotobacter vinelandii. *Biochim Biophys Acta* 20, 269-285.

Grundy & Henkin (2006).

Guerrier-Takada C, Gardiner K, Marsh T, Pace N & Altman S (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* 35, 849-857.

Gunisova et al (2017)

Gupta R (1984). *Halobacterium volcanii* tRNAs. *J Biol Chem* 259, 9461-9471.

Han et al (2023)

Hanczyc et al (2003).

Hao et al (2022).

Haruna I, Nozu K, Ohtaka Y & Spiegelman S (1963). An RNA “Replicase” induced by and selective for a viral RNA: Isolation and properties. *Proc Natl Acad Sci USA* 50, 905-911.

Haugen P, Simon DM & Bhattacharya D (2005). The natural history of group I introns. *Trends Genet* 21, 111-119.

Higgs PG (2009). A four-column theory for the origin of the genetic code: tracing the evolutionary pathways that gave rise to an optimized code. *Biol Direct* 4, 16.

Higgs PG & Pudritz RE (2009). A thermodynamic basis for prebiotic amino acid synthesis and the nature of the first genetic code. *Astrobiology* 9, 483-490.

Hoagland M (1959). Biochemical activities of nucleic acids: the present status of the adaptor hypothesis. Brookhaven Symp Biol 12 (November), 40-46.

Holden et al (2022).

Holley RW, Apgar J, Everett GA, Madison JT, Marquisee M, Merrill SH, Penswick JR & Zamir A (1965). Structure of a ribonucleic acid. *Science* 147, 1462-1465.

Hopfield JJ (1974). Kinetic proofreading: a new mechanism for reducing errors in biosynthetic proccesses requiring high specificity. Proc Natl Acad Sci 71, 4135-4139.

Hsiao C, Mohan S, Kalahar BK & Williams LD (2009). Peeling the onion: ribosomes are ancient molecular fossils. *Mol Biol Evol* 26, 2415-2425.

Huang & Karplus (2019).

Huang S, Aleksashin NA, Loveland AB, Klepacki D, Reier K, Kefi A, Szal T, Remme J, Jaeger L, Vazquez-Laslop N, Korostelev AA & Mankin AS (2020). Ribosome engineering reveals the importance of 5S rRNA autonomy for ribosome assembly. *Nature Communication* 11, 1-13.

Huber et al (2023)

Hud NV, Cafferty BJ, Krishnamurthy R & Williams LD (2013). The origins of RNA and “my grandfather’s axe.” *Chem Biol* 20, 466-474.

Hutter H et al (2000).

Ianeselli A, Atienza M, Kudella PW, Gerland U, Mast CB & Braun D (2022). Water cycles in a Hadean CO2 atmosphere drive the evolution of long DNA. Nature Physics 18, 579-585.

Ianeselli A, Tetiker D, Stein J, Kuhnlein A, Mast C B, Braum D & Tang TYD (2022). Non-equilibrium conditions inside rock pores drive fission, maintenance & selection of coacervate protocells. Nature Chemistry 14, 32-39.

Ianeselli A, Salditt A, Mast C, Ercolano B, Kufner CL, Scheu B & Braun D (2023). Physical non-equlibria for prebiotic nuclei acid chemistry. *Nature Rev Physics*

Inokuchi et al (2023)

Inoue T & Orgel LE (1983). A nonenzymatic RNA polymerase model, *Science* 219, 859-862.

Ishida S, Terasaka N, Katoh T & Suga H (2020). An aminoacylation ribozyme evolved from a natural tRNA-sensing T-box riboswitch. *Nature Chemical Biology* 16, 702-709.

Jacome R, Campillo-Balderas JA, Becerra A & Lazcano A (2022). Structural analysis of monomeric RNA-dependent polymerases revisited. J Mol Evol 90, 283-295.

Jadhav VR & Yarus M (2002). Coenzymes as coribozymes. *Biochimie* 84, 877-888.

Janosi et al (1996)

Javaux EJ (2019). Challenges in evidencing the earliest traces of life. *Nature* 572, 451-460.

Jeffares DC, Poole AM & Penny D (1998). Relics from the RNA world. J Mol Evol 46, 18-36.

Jencks WP (1975). Binding energy, specificity, and enzymic catalysis: the circe effect. Adv Enzymol Relat Areas Mol Biol 43, 219-410.

Jin et al (2018).

Johnston WK, Unrau PJ, Lawrence MS, Glasner ME & Bartel DP (2001). RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. *Science* 292, 1319-1325.

Joyce GF & Orgel LE (1993). Prospects for understanding the origin of the RNA world.

Joyce GF, Visser GM, Van Boeckel CAA, Van Boom JH, Orgel LE & Van Westrenen J (1984). Chiral selection in poly (C)-directed synthesis of oligo (G). Nature 310, 602-604.

Kacian DL, Mills DR, Kramer FR & Spiegelman S (1972). A replicating RNA molecule suitable for a detailed analysis of extracellular evolution and replication. *Proc Natl Acad Sci USA* 69, 3038-3042.

Kamerlin SCL, Sharma PK, Prasad RB & Warshel A (2013). Why nature really chose phosphate. *Quart Rev Biophysics* 46, 1-132.

Kanai A (2015). Disrupted tRNA genes and tRNA fragments : a perspective on tRNA gene evolution. *Life* 5, 321-331.

Kanavarioti A, Monnard PA & Deamer DW (2001). Eutectic phase in ice facilitate nonenzymatic nucleic acid synthesis. *Astrobiol* 1, 271-281.

Kasting & Brown (1998).

Kavita & Breaker (2022).

Kawabata M, Kawashima K, Mutauro-Aoki H, Ando T, Umehara T & Tamura K (2022). Peptide bond formation between aminoacyl-minihelices by a scaffold derived from the peptidyl transferase center. *Life* 12, 573-583.

Kazemi M, Socan J, Himo F & Aqvist J (1918). Mechanistic alternatives for peptide bond formation on the ribosome. Nucleic Acids Res 46, 5345-5354.

Keiler KC, Walker PR & Sauer RT (1996). Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. Science 271, 990-993.

Kerkhofs K, Garg J, Fafard-Couture, Elela SA, Scott MS, Pearlman RE & Bayfied MA (2022). Altered tRNA processing is linked to a distinct and unusual La protein in Tetrahymena thermophila. Nature Comm 13, 1-17.

Khaitovich P & Mankin AS (1999). Effects of antibiotics on large ribosomal subunit assembly reveals possible function of 5 S rRNA. *J Mol Biol* 291, 1025-1034.

Khade PK, Shi X & Joseph S (2013). Steric complementarity in the decoding center is important for tRNA selection by the ribosome. J Mol Biol 425, 3778-3789.

Kim SH, Sussman JL, Suddath Fl, Quigley GJ, McPherson A, Wang, AHJ, Seeman NC & Rich A (1974). The general structure of transfer RNA molecules. *Proc Nat Acad Sci USA* 71, 4970-4974.

Kim SC, O'Flaherty DK, Giurgiu C, Zhou L & Szostak JW (2021). The emergence of RNA from the heterogeneous products of prebiotic nucleotide synthesis. *J Am Chem Soc* 143, 3267-3279.

Kirsebom LA & Svard SG (1994). Base pairing between Escherichia RNase P RNA and its substrate. *EMBO J* 13, 4870-4876.

Kirsebom LA & Trobro S (2009). RNase P RNA-mediated cleavage. *Life* 61, 189-200.

Kirschning A (2021). Coenzymes and their role in the evolution of life. *Angew Chem Int Ed* 60, 6242-6269.

Kisonaite M, Wild K, Lapouge K, Ruppert T & Sinning I (2022). High-resolution structures of a thermophilic eukaryotic 80S ribosome reveal atomistic details of translocation. Nature Comm 13, 1-12.

Kitadai & Muruyama (2018).

Kleene (1952).

Knight RD, Freeland SJ & Landweber LF (2004). Adaptive evolution of the genetic code. In *The Genetic Code and the Origins of Life* (ed de Pouplana LR). Kluwer Academic.

Kofman C, Lee J & Jewett MC (2021). Engineering molecular translation systems. Cell Systems 12, 593-607.

Koga T & Naraoka H (2017). A new family of extraterrestrial amino acids in the Murchison meterorite. *Scientific Reports* 7, 636-643.

Koonin EV (2011). *The Logic of Chance: The Nature and Origin of Biological Evolution*. FT Press. Upper Saddle River NJ.

Koonin EV (2017). Frozen accident pushing 50: stereochemistry, expansion, and chance in the evolution of the genetic code. Life 7, 1-13.

Koonin EV, Krupovic M, Ishino S & Ishino Y (2020a). The replication machinery of LUCA: common origin of DNA replication and transcription. *BMC Biology* 18, 1-8.

Koonin EV, Makarova KS, Wolf YI & Krupovic M (2020b). Evolutionary entanglement of mobile genetic elements and host defence systems: guns for hire. *Nature Rev* 21, 119-131.

Koonin EV, Mushegian AR & Bork P (1996). Non-orthologous gene displacement. *Trends Genetics* 12, 334-336.

Kornberg A (1969). Active center of DNA polymerase. *Science* 163, 1410-1418.

Korostelev (2021).

Kortmann J & Narberhaus F (2012). Bacterial RNA thermometers: molecular zippers & switches. Nature Rev Microbiol 10, 255-265.

Kovacs NA, Petrov AS, Lanier KA & Willams LD (2017). Frozen in time: the history of proteins. *Mol Biol Evol* 34, 1252-1260.

Kruger K, Grabowski PJ, Zaug AJ, Sands J, Gottschling DE & Cech TR (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell* 31, 147-157.

Krupkin M, Matzov D, Tang H, Metz M, Kalaora R, Belousoff MJ, Zimmerman E, Bashan A & Yonath A (2011). A vestige of a prebiotic bonding machine is functioning within the contemporary ribosome. *Phil Trans R Soc* B 366, 2972-2978.

Kuhlenkoetter S, Wintermeyer W & Rodnina MV (2011). Different substrate-dependent transition states in the active site of the ribosome. Nature 476, 351-355.

Kurland (1992).

Kushner (2018).

Kvenvolden et al (1990).

Lahry K, Datta M & Varshney U (2024). Genetic analysis of translation initiation in bacteria: an initiator tRNA-centric view. *Molecular Microbiology* 00, 1-17.

Lan et al (2018).

Lazcano (2016).

Lee et al (1997).

Lehman N & Jukes TH (1988). Genetic code development by stop codon takeover. *J theor Biol* 135, 203-214.

Lehmann J, Jossinet F & Gautheret D (2013). A universal RNA structural motif docking the elbow of tRNA in the ribosome, RNAse P and T-box leaders. *Nucleic Acids Res* 41, 5495-5502.

Leidig C, Thoms M, Holdermann I, Bradatsch B, Berninghausen O, Bange G, Sinning I, Hurt E & Beckmann R (2014). 60S ribosome biogenesis requires rotation of the 5S ribonucleoprotein particle. Nature Comm 5, 3491.

Leininger SE, Narayan K, Deutsch C & O'Brien EP (2019). Mechanochemistry in translation. *Biochemistry* 58, 4657-4666.

Leipe DD, Aravind L & Koonin EV (1999). Did DNA replication evolve twice independently? Nucleic Acids Res 27, 3389-3401.

Leipe DD, Wolf YI, Koonin EV & Aravind L (2002). Classification and evolution of P-loop GTPases and related ATPases. *J Mol Biol* 317, 41-72.

Leipe DD, Koonin EV & Aravind L (2003). Evolution and classification of P-loop kinases and related proteins. *J Mol Biol* 333, 781-815.

Leiva & Katz (2022).

Leopold et al (1992).

Le Vay & Mutschler (2019).

Levinthal C (1969). How to fold graciously.

Levy M & Ellington AD (2001). The descent of polymerization. *Nat Struct Mol Biol* 8, 580-582.

Li Y & Breaker R (1999). Kinetics of RNA degradation by specific base catalysis of transesterification involving the 2'-hydroxyl group. *J Am Chem Soc* 121, 5364-5372.

Licznar P, Mejlhede N, Prere MF, Wills N, Gesteland RF, Atkins JF & Fayet O (2002). Programmed translational -1 frameshifting on hexanucleotide motifs and the wobble properties of tRNAs. *EMBO J* 22, 4770-4778.

Lipmann F (1971). Attempts to map a process evolution of peptide biosynthesis. *Science* 173, 875-884.

Liu Z, Wu LF, Xu J, Bonfio C, Russell DA & Sutherland JD (2020). Harnessing chemical energy for the activation and joining of prebiotic building blocks. *Nat Chem* 12, 1023-1028.

Lohse & Szostak (1996).

Long et al (2022).

Longo LM, Jablonska J, Vyas P, Kanade M, Kolodny R, Ben-Tal N & Tawfik DS (2020). On the emergence of P-loop NTPase and Rossmann enzymes from a beta-alpha-beta ancestral fragment. *eLife* 9, 1-16.

Lohrmann R & Orgel L (1973). Prebiotic activation processes. *Nature* 244, 418-420.

Lorenz MR et al (2017). Proton gradients and pH oscillations emerge from heat flow at the microscale. Nat Commun 8, 1897.

Lupas AN & Alva V (2017). Ribosomal proteins as documents of the transition from unstructured (poly)peptides to folded proteins. *J Structural Biol* 198, 74-81.

Maini R, Dedkova LM, Paul R, Madathil MM, Chowdhury SR, Chen S & Hecht SM (2015). Ribosome-mediated incorporation of dipeptides and dipeptide analogues into proteins in vitro. *J Am Chem Soc* 137, 11206-11209.

Maracci C & Rodnina MV (2016). Translational GTPases. *Biopolymers* 105, 463-475.

Martin (1998).

Martínez-Giménez JA & Tabarés-Seisdedos R (2021). Possible ancestral functions of the genetic and RNA operatonal precodes and the origin of the genetic system. *Origins Life Evol Biospheres* 51, 167-183.

Martins et al (2007).

Matzov D, Taoka M, Nobe Y, Yamauchi Y, Halfon Y, Asis N, Zimermann E, Rozenberg H, Bashan A, Bhushan S, Isobe T, Gray MW, Yonath A & Shalev-Benami M (2020). Cryo-EM structure of the highly atypical cytoplasmic ribosome of Euglena gracilis. *Nucleic Acids Res* 48, 11750-11761.

Maynard Smith J (1983). Models of evolution. *Proc Roy Soc Lond* 219, 315-325.

Maynard Smith J & Szathmary E (1995). *The Major Transitions in Evolution*. Freeman Press: Oxford.

McClain WH, Guerrier-Takada C & Altman S (1987). Model substrates for an RNA enzyme. *Science* 238, 527-530.

McRae EKS, Wan CJK, Kristoffersen EL, Hansen K, Gianni E, Gallego I, Curran JF, Attwater J, Holliger P & Andersen ES (2022). Cryo-EM structure and functional landscape of an RNA polymerase ribozyme. bioRxiv

Meng K, Chung CZ, Soll D & Krahn N (2022). Unconventional genetic code systems in archaea. *Frontiers Microbiol* ##, 1-18.

Milicevic et al (2023).

Miller SL (1953). A production of amino acids under possible primitive earth conditions. *Science* 117, 528-529.

Miller & Urey (1959).

Mills DR, Peterson RL & Spiegelman S (1967). An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proc Natl Acad Sci USA* 58, 217-224.

Mitchell et al (1997).

Miyazaki Y & Korenaga J (2022). A wet heterogeneous mantle creates a habitable world in the Hadean. Nature 603, 86-90.

Mizuuchi R & Ichihashi N (2021). Primitive compartmentalization for the sustainable replication of genetic molecules. *Life* 11, 1-17.

Moore (1956).

Moore PB & Steitz TA (2002). The invovlement of RNA in ribosome function. Nature 418, 229-235.

Moore SD & Sauer RT (2007). The tmRNA system for translational surveillance and ribosome rescue. Annu Rev Biochem 76, 101-124.

Mueller et al (1990).

Muller F, Escobar L .... et al 2022. A prebiotically plausible scenario of an RNA-peptide world. Nature 605, 279-284.

Murakami H, Ohta A, Ashigai H & Suga H (2006). A highly flexible tRNA acylation method for non-natural polypeptide synthesis. *Nature Methods* 3, 357-359.

Mutscher H, Wochner A & Holliger P (2015). Freeze-thaw cycles as drivers of complex ribozyme assembly. *Nature Chem* 7, 502-508.

Nagaswamy U & Fox GF (2003). RNA ligation and the origin of tRNA. Origins Life Evol Biosphere 33, 199-209.

Nam et al (2018).

Natsidis P, Schiffer PH, Salvador-Martinez I & Telford MJ (2019). Computational discovery of hidden breaks in 28S ribosomal RNAs across eukaryotes and consequences for RNA Integrity Numbers. *Nature Scientific Reports* 9, 19477.

Nelsesteun GL (1980). Origin of life: consideration of alternatives to proteins and nucleic acids. J Mol Evol 15, 59-72.

Nelson JW & Breaker RR (2017). The lost language of the RNA World. *Sci Signal* 10, 1-10.

Nesbitt et al (1999).

Nicholls JWF, Chin JP, Williams TA, Lenton TM, O'Flaherty V & McGrath JW (2023). On the potential roles of phosphorus in the early evolution of energy metabolism. *Frontiers Microbiol* 14, 1239189.

Nielsen (1993).

Nielsen PE, Egholm M, Berg RH & Buchardt O (1991). Sequence-selective recognition of DNA by strand displacement with a thymine-substitute polyamide. *Science* 254, 1497-1500.

Nilsson OB, Hedman R, Marino J, Wickles S, Bischoff L, Johansson M, Muller-Lucks A, Trovato F, Puglisi JD, O'Brien EP, Beckmann & von Heijne G (2015). Cotranslational protein folding inside the ribosome exit tunnel. Cell Reports 12, 1533-1540.

Ninio J (1975). Kinetic amplification of enzyme discrimination. *Biochimie* 57, 587-595.

Nirenberg M (2004). Historical: deciphering the genetic code—A personal account. *Trends Biochem Sci* 29, 46-54.

Nissen P, Hansen J, Ban N, Moore PB & Steitz TA (2000). The structural basis of ribosome activity in peptide bond synthesis. Science 289, 920-930.

Nissen P, Ippolito JA, Ban N, Moore PB & Steitz TA (2001). RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. Proc Natl Acad Sci USA 98, 4899-4903.

Noller HF (2004). The driving force for molecular evolution of translation. *RNA* 10, 1833-1837.

Noller HF (2010/2012). Evolution of protein synthesis from an RNA world. *Cold Spring Harbor Perspect Biol* 4, 1-14.

Noller HF, Hoffarth V & Zimniak L (1992). Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* 256, 1416-1419.

Nutman AP, Bennett VC, Friend CRL, Van Kranendork MJ & Chivas AR (2016). Rapid emergence of life shown by discovery of 3,700-million-year-old microbial structures. *Nature* 537, 535-538.

Oba Y, Takano Y, Furukawa Y, Koga T, Glavin DP, Dworkin JP & Naraoka H (2022). Identifying the wide diversity of extraterrestrial purine and pyrimidine nucleobases in carbonaceous meteorites. *Nature Comm* 13, 1-10.

Oberbeck et al (1991).

Ogle JM, Brodersen DE, Clemons WM, Tarry MJ, Carter AP & Ramakrishnan V (2001). Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science* 292, 897-902.

Ogle JM, Murphy FV, Tarry MJ & Ramakrishnan V (2002). Selection of tRNA by the ribosome requires a transition from an open to a closed form. *Cell* 111, 721-732.

Olenginski et al (2024).

Oparin (1938)

Orellana O, Cooley L & Soll D (1986). The additional guanylate at the 5' terminus of Escherichia coli $tRNA^{His}$ is the result of unusual processing by RNase P. *Mol Cell Biol* 6, 525-529.

Orgel LE (1968). Evolution of the genetic apparatus. *J Mol Biol* 38, 381-393.

Orgel (1986).

Orgel LE (2003).

Orgel LE (2004).

Orgel LE & Crick FHC (1980). Selfish DNA : the ultimate parasite. *Nature* 284, 604-607.

Orgel LE & Sulston JE (1971).

Oro J (1961). Mechanism of synthesis of adenine from hydrogen cyanide under plausible primitive earth conditions. *Nature* 191, 1193-1194.

Oro & Kiball (1960).

Pace NR & Marsh TL (1985). RNA catalysis and the origin of life. *Origins Life* 16, 97-116.

Paisley TE & Van Tuyle GC (1994). The processing of wild type and mutant forms of rat nuclear pre-$tRNA^{Lys}$ by homologous RNase P. *Nucl Acids Res* 22, 3347-3353.

Pasek MA (2020). Thermodynamics of prebiotic phosphorylation. *Chem Rev* 120, 4690-4706.

Pavlova et al (2019).

Pearce BKD, Tupper AS, Pudritz RE & Higgs PG (2018). Constraining the time interval for the origin of life on earth. *Astrobiol* 18, 343-364.

Peng BZ, Bock LV, Belardinelli R, Peske F, Grubmuller H & Rodnina MV (2019). Active role of elongation factor G in maintaining the mRNA reading frame during translation. *Sci Adv* 5, 1-8.

Persson T, Cuzic S & Hartmann RK (2003). Catalysis by RNase P RNA: unique features and unprecedented active site plasticity. *J Biol Chem* 278, 43394-43401.

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.

Pettersson BMF, Ardell DH & Kirsebom LA (2005). The length of the 5' leader of Escherichia coli tRNA precursors influences bacterial growth. J Mol Biol 351, 9-15.

Phan HD, Lai LB, Zahurancik WJ & Gopalan V (2021). The many faces of RNA-based RNase P, an RNA-world relic. *Trends Biochem Sci* 46, 976-991.

Pinna S, Kunz C, Halpern A, Harrison SA, Jordan SF, Ward J, Werner F & Lane N (2022). A prebiotic basis for ATP as the universal energy currency. PLOS Biology 20, 1-25.

Pizzarello (2006).

Polacek N, Gomez MJ, Ito K, Xiong L, Nakamura Y & Mankin A (2003). The critical role of the universally conserved A2602 of 23S ribosomal RNA in the release of the nascent peptide during translation termination. *Mol Cell* 11, 103-112.

Polacek N & Mankin AS (2005). The ribosomal peptidyl transferase center: structure, function, evolution, inhibition. *Critical Reviews Biochem Mol Biol* 40, 285-311.

Polikanov YS, Steitz TA & Innis CA (2014). A proton wire to couple aminoacyl-tRNA accommodation and peptide-bond formation of the ribosome. *Nature Structural Mol Biol* 21, 787-793.

Poole AM, Jeffares DC & Penny D (1998). The path from the RNA world. *J Mol Evol* 46, 1-17.

Popper (1959).

Poudyal et al (2018).

Poulis P (2022). Ribosome dynamics during spontaneous frameshifting. Doctoral Dissertation, Univ Gottingen.

Poulis P, Patel A, Rodnina MV & Adio S (2022). Altered tRNA dynamics during translocation on slippery mRNA as determinant of spontaneous ribosome frameshifting. Nature Comm 13, 4231-####.

Powner MW, Gerland B & Sutherland JD (2009). Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* 459, 239-242.

Prabhakar et al (2017).

Prigogine I (1978). Time, structure, and fluctuations. *Science* 201, 777-785.

Rae CD, Gordiyenko Y & Ramakrishnan V (2019). How a circularized tmRNA moves through the ribosome. Science 363, 740-744.

Ramachandran GN, Ramakrishnan C & Sasisekharan V (1963). Stereochemistry of polypeptide chain configurations. *J Mol Biol* 7, 95-99.

Randau L & Soll D (2008). Transfer RNA genes in pieces. *EMBO Rep* 9, 623-628.

Reimann & Zubay (1999).

Reinhold-Hurek B & Shub DA (1992). Self-spicing introns in tRNA genes of widely divergent bacteria. *Nature* 357, 173-176.

Rich A (1962). On the problems of evolution and biochemical information transfer. In *Horizons in Biochemistry* (ed Kasha M & Pullman B). New York: Academic Press. 103-126.

Riddle DL & Carbon J (1973). Frameshift suppression: a nucleotide addition in the anticodon of a glycine transfer RNA. Nature New Biol 242, 230-234.

Ring D, Wolman Y, Friedmann N & Miller SN (1972). Prebiotic synthesis of hydrophobic and protein amino acids. *Proc Natl Acad Sci* 69, 765-768.

Robertson & Joyce (2012).

Robertus JD, Ladner JE, Finch JT, Rhodes D, Brown RS, Clark BFC & Klug A (1974). Structure of yeast phenylalanine tRNA at 3 A resolution. *Nature* 250, 546-551.

Rodriguez-Garcia et al (2015).

Rogalski M, Karcher D & Bock R (2008). Superwobbling facilitates translation with reduced tRNA sets. *Nature Structural Mol Biol* 15, 192-198.

Roovers M, Droogmans L & Grosjean (2021). Post-transcriptional modifications of conserved nucleotides in the T-loop of tRNA: a tale of functional convergent evolution. *Genes* 12, 1-19.

Ross DS & Deamer D (2016). Dry/wet cycling and the thermodynamics and kinetics of prebiotic polymer synthesis. Life 6, 28-.

Ross DS & Deamer D (2019). Prebiotic oligomer assembly: What was the energy source? Astrobiology 19, 517-521.

Ross DS & Deamer D (2022). Duplex formation and the origins of homochirality. Astrobiology 22, 192-196.

Ross DS & Deamer D (2023). Template-directed replication and chiral resolution during wet-dry cycling in hydrothermal pools. Life 13, 1749-1758.

Rozov A, Demeshkina N, Khusainov I, Westhof E, Yusupov M & Yusupova G (2016). Novel base-pairing interactions at the tRNA wobble position crucial for accurate reading of the genetic code. *Nature Comm* 7, 1-10.

Rubio Gomez MA & Ibba M (2020). Aminoacyl-tRNA synthetases. *RNA* 26, 910-936.

Saenger (1984).

Saito et al (2020).

Salibi E, Peter B, Schwille P & Mutschler H (2022). Exchange, catalysis & amplification of encapsulated RNA driven by periodic temperature changes. Research Square 1-21.

SantaLucia J (1998). A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proc Nat Acad Sci USA 95, 1460-1465.

Sasselov et al (2020).

Schilling-Bartetzko et al (1992).

Schimmel P & Ribas de Pouplana L (1995). Transfer RNA: from minihelix to genetic code. Cell 81, 983-986.

Schmeing TM & Ramakrishnan V (2009). What recent ribosome structures have revealed about the mechanism of translation. Nature Reviews 461, 1234-1242.

Schurman JG (1887). The Ethical Import of Darwinism. New York. Charles Scribner's Sons.

Schwartz AW & Orgel LE (1985). Template-directed synthesis of novel, nucleic acid-like structures. *Science* 228, 585-587.

Selmer M, Dunham CM, Murphy FV, Weixlbaumer A, Petry S, Kelley AC, Weir JR & Ramakrishnan V (2006). Structure of the 70S ribosome complexed with mRNA and tRNA. Science 313, 1935-1942.

Shannon CE (1936).

Shannon CE & Weaver W (1949). *The Mathematical Theory of Communication*. University of Illinois Press. Urbana.

Sharma U, Conine CC, Shea JM, Boskovic A, Derr AG, Bing XY, Balleannee C, Kucukural A, Serra RW, Sun F, et al. (2016). Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. *Science* 351, 391-396.

Sharp PA (1985). On the origin of RNA splicing and introns. *Cell* 42, 397-400.

Shin et al (2023).

Shine & Dalgarno (1974).

Shirokikh NE (2023). Pathways of early evolution from the perspectives of a riboreplisome - the ultimate RNA machine of life. (www.preprints.org)

Simmel FC, Yurke B & Singh HR (####). Principles and applications of nucleic acid strand displacement reactions. Chemical Reviews

Spirin (1968).

Spirin (2009).

Stevens (1961).

Stueken EE, Anderson RE, Bowman JS, Brazelton WJ, Colangelo-Lillis J, Goldman AD, Som SM & Baross JA (2013). Did life originate from a global chemical reactor? *Geobiology* 11, 101-126.

Suddala et al (2023).

Sulston J, Lohrmann R, Orgel LE & Miles HT (1968a?b). Nonenzymatic synthesis of oligoadenylates on a polyuridylic acid template. *Proc Nat Acad Sci USA* 59, 726-733.

Sun L, Campbell FE, Zahler NH & Harris ME (2006). Evidence that substrate-specific effects of C5 protein lead to uniformity in binding and catalysis by RNase P. EMBO J 25, 3998-4007.

Sutherland (2016).

Sutherland JD (2017). Studies on the origin of life - the end of the beginning. *Nature Rev Chem* 1, 1-7.

Syroegin EA, Aleksandrova EV & Polikanov YS (2023). Insights into the ribosome function from the structure of non-arrested ribosome-nascent chain complexes. *Nature Chem* 15, 143-153.

Szostak JW (2012). The eightfold path to non-enzymatic RNA replication. *J Syst Chem* 3, 1-14.

Szostak (2017).

Szathmary E (1993). Coding coenzyme handles: A hypothesis for the origin of the genetic code. *Proc Natl Acad Sci USA* 90, 9916-9920.

Szathmary E (1999). The origin of the genetic code: amino acids as cofactors in an RNA world. Trends Genet 15, 223-229.

Takeuchi N, Kaneko K & Koonin EV (2014). Horizontal gene transfer can rescue prokaryotes from Muller's rachet: benefit of DNA from dead cells and population subdivision. *G3 (Bethesda)* 4, 325-330.

Tallsjo A & Kirsebom LA (1993). Product release is a rate-limiting step during cleavage by the catalytic RNA subunit of *Escherichia coli* RNase P. *Nuc Acids Res* 21, 51-57.

Tamura K & Schimmel P (2003). Peptide synthesis with a template-like RNA guide and aminoacyl phosphate adaptors. *Proc Natl Acad Sci* 100, 8666-8669.

Tamura K & Schimmel P (2006). Chiral-selective aminoacylation of an RNA minihelix: mechanistic features and chiral suppression. *Proc Natl Acad Sci USA* 103, 13750-13752.

Tamura K (2015). Origins and early evolution of the tRNA molecule. *Life* 5, 1687-1699.

Tanaka T & Kikuchi Y (2001). Origin of the cloverleaf shape of transfer RNA - the double-hairpin model : implication for the role of tRNA intron and the long extra loop. *Viva Origino* 29, 134-142.

Temin HM & Mizutani S (1970). RNA-dependent DNA polymerase in virions of rous sarcoma virus. *Nature* 226, 1211-1213.

Thagard P (1980). Against evolutionary epistemology. *Proc Biennial Meeting Phil Sci Assc* pp187-196.

Tinoco I & Bustamante C (1999). How RNA folds. *J Mol Biol* 293, 271-281.

Vaidya N, Manapat ML, Chen IA, Xulvi-Brunet R, Hayden EJ & Lehman N (2012). Spontaneous network formation among cooperative RNA replicators. *Nature* 491, 72-77.

Vetsigian K, Woese C & Goldenfeld N (2006). Collective evolution and the genetic code. *Proc Natl Acad Sci USA* 103, 10696-10701.

Vetter IR & Wittinghofer A (1999). Nucleoside triphosphate-binding proteins: different scaffolds to achieve phosphoryl transfer. *Quart Rev Biophysics* 32, 1-56.

Vicens Q & Kieft JS (2022). Thoughts on how to think (and talk) about RNA structure. *Proc Natl Acad Sci USA* 119, 1-9.

Viegas SC, Matos RG & Arraiano CM (2020). The bacterial counterparts of th eukaryotic exosome: an evolutionary perspective. In *The Eukaryotic RNA Exosome* (eds LaCava J & Vanacova S). Springer. pp37-46

Visser CM (1984). Evolution of biocatalysts 1. Possible pre-genetic-code RNA catalysts which are their own replicase. *Orig Life* 14, 291-300.

Voorhees RM & Hedge RS (2015). Structure of the scanning and engaged states of the mammalian SRP-ribosome complex. Elife 4, e07975.

Wachtershauser G (1988). Before enzymes and templates: theory of surface metabolism. *Microbiol Rev* 52, 452-484.

Wachowius F & Holliger P (2021). RNA replication and the RNA polymerase ribozyme. In Ribozymes (edited Muller S, Masquida B & Winkler W) Wiley-VCH.

Wang Q & Su H (2022). A tale of water molecules in the ribosomal peptidyl transferase reaction. Biochemistry 61, 2241-2247.

Watanabe S (1969). *Knowing and Guessing: A Formal and Quantitative Study*. New York: John Wiley & Sons.

Watson JD & Crick FHC (1953a). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* 171, 964-967.

Watson JD & Crick FHC (1953b). Genetical implications of the structure of DNA. *Nature* ##, ###.

Watson JD et al (2016).

Weber AL (1989). Model of early self-replication based on covalent complementarity for a copolymer of glycerate-3-phosphate and glycerol-3-phosphate. *Origins Life* 19, 179-186.

Weber & Orgel (1980).

Wei et al (2020).

Weimann BJ, Lohrmann R, Orgel LE, Schneider-Bernloehr H & Sulston JE (1968). Template-directed synthesis with adenosine-5'-phosphorimidazolide. *Science* 161, 387-.

Weiner AM & Maizels N (1987). tRNA-like structures tag the 3' ends of genomic RNA molecules for replication : Implications for the origin of protein synthesis. *Proc Acad Sci USA* 84, 7383-7387.

Weinger JS, Parnell KM, Dorner S, Green R & Strobel SA (2004). Substrate-assisted catalysis of peptide bond formation by the ribosome. Nature Struc Mol Biol 11, 1101-1106.

Weiss R & Cherry J (1993). Speculations on the origin of ribosomal translocation. *Cold Spring Harbor Monograph Series* 24, 71-89.

Welting TJM, Kikkert BJ, Van Venrroij WJ & Pruijn GJM (2006). Differential association of protein subunits with human RNase MRP and RNase P complexes. *RNA* 12, 1373-1382.

Werren (2011).

Westheimer FH (1987). Why nature chose phosphate. *Science* 235, 1173-1178.

White HB (1976). Coenzymes as fossils of an earlier metabolic state. *J Mol Evol* 7, 101-104.

Widmann J, Di Giulio M, Yarus M & Knight R (2005). tRNA creation by hairpin duplication. J Mol Evol 61, 524-530.

Wilmer (2022).

Wilson & Lilly (2021).

Wilson & Szostak (1999).

Woese C (1967). *The Origins of the Genetic Code*. New York: Harper & Row.

Wolfenden R, Lu X & Young G (1998). Spontaneous hydrolysis of glycosides. *J Am Chem Soc* 120, 6814-1815.

Wolfenden R (2011). Benchmark reaction rates, the stability of biological molecules in water, and the evolution of catalytic power in enzymes. *Annu Rev Biochem* 80, 645-667.

Wolman et al (1972).

Wong JT (1975). A co-evolution theory of the genetic code. *Proc Natl Acad Sci USA* 72, 1909-1912.

Wong JT, Ng SK, Mat WK, Hu T and Xue H (2016). Coevolution theory of the genetic code at age forty: pathway to translation and synthetic life. *Life* 6, 12.

Wu T & Orgel LE (1992). Nonenzymatic template-directed synthesis on hairpin oligonucleotides. J Am Chem Soc 114, 5496-5501.

Wu LF & Sutherland JD (2019). Provisioning the origin and early evolution of life. *Emerging Topics in Life Sciences* 3, 459-468.

Wu LF, Su M, Liu Z, Bjork SJ & Sutherland JD (2021). Interstrand aminoacyl transfer in a tRNA acceptor stem-overhang mimic. *J Am Chem Soc* 143, 11836-11842.

Yadav et al (2020).

Yakhnin AV (2007). A model for the origin of protein synthesis as coreplicational scanning of nascent RNA. *Orig Life Evol Biosph* 37, 523-536.

Yamamoto et al (2016).

Youngman EM, Brunelle JL, Kochaniak AB & Green R (2004). The active site of the ribosome is composed of two layers of conserved nucleotides with distinct roles in peptide bond formation and peptide release. *Cell* 117, 589-599.

Zaher HS, Shaw JJ, Strobel SA & Green R (2011). The 2'-OH group of the peptidyl-tRNA stabilizes an active conformation of the ribosomal PTC. *EMBO J* 30, 2445-2453.

Zamecnik PC (1960). Historical and current aspects of the problem of protein synthesis. *Harvey Lectures* 1958-1959, 250-281. New York: Academic Press.

Zaug AJ & Cech TR (1985). Oligomerization of intervening sequence RNA molecules in the absence of proteins. *Science* 229, 1060-1061.

Zaug AJ & Cech TR (1986). The Tetrahymena intervening sequence ribonucleic acid enzyme is a phosphotransferase and an acid phosphatase. Biochemistry 25, 4478-4482.

Zaug AJ & Cech TR (1986). The intervening sequence RNA of *Tetrahymena* is an enzyme. *Science* 231, 470-475.

Zhang B & Cech TR (1997). Peptide bond formation by *in vitro* selected ribozymes. *Nature* 390, 96-100.

Zhang J & Ferre-DAmare AR (2016a?b). Trying on tRNA for size: RNase P and the T-box riboswitch as molecular rulers. *Biomolecules* 6, 1-14.

Zhang SJ, Duzdevich D, Ding D & Szostak JW (2022). Freeze-thaw cycles enable a prebiotically plausible & continuous pathway from nucleotide activation to nonenzymatic RNA copying. *Proc Natl Acad Sci USA* 119, 1-7.

Zhou L, Kim SC, Ho KH, O'Flaherty DK, Giurgiu C, Wright TH & Szostak JW (2019a). Non-enzymatic primer extension with strand displacement. *eLife* 8, 1-14.

Zhou J, Lancaster L, Donohue JP & Noller HF (2019b). Spontaneous ribosomal translocation of mRNA and tRNAs into a chimeric hybrid state. Proc Natl Acad Sci USA 116, 7813-7818.

Zhu J, Huang W, Zhao J, Huynh L, Taylor DJ & Harris ME (2022). Structural and mechanistic basis for recognition of alternative tRNA precursor substrates by bacterial ribonuclease P. *Nature Comm* 13, 1-13.

Ziv G, Haran G & Thirumalai D (2005). Ribosome exit tunnel can entropically stabilize alpha-helices. Proc Nalt Acad Sci USA 102, 18956-18961.

Zu et al (2022).

Zuo Z, Peng D, Yin X, Zhou X, Cheng H & Zhou R (2013). Genome-wide analysis reveals origin of transfer RNA genes from tRNA halves. *Mol Biol Evol* 30, 2087-2098.