Journal of Biological Research-Thessaloniki **15**: 3 – 16, 2011
J. Biol. Res.-Thessalon. is available online at http://www.jbr.gr
Indexed in: WoS (Web of Science, ISI Thomson), SCOPUS, CAS (Chemical Abstracts Service) and DOAJ (Directory of Open Access Journals)

An evolutionary theory based on a protein-mRNA co-synthesis hypothesis

Yannis GOUNARIS*

Department of Agriculture, *University of Thessaly*, *Fytokou Street*, *38446 New Ionia*, *Greece*

Received: 10 March 2010 Accepted after revision: 7 September 2010

A new hypothesis is proposed, that early life possessed a mechanism for the simultaneous synthesis of a polypeptide and its coding mRNA. Early tRNAs and mRNAs are considered to have been pairs of simple complementary nucleotide triplets, able to catalyze peptide bond formation and joining of the codon triplets. An *in silico* build stereochemical model for such a process is presented. A model for the evolution of modern tRNAs and ribosomal translation from such a primitive apparatus is also presented. Modern tRNAs are viewed as consisting of two halves, one of which evolved from ancient elongating mRNA and the other from a primitive also elongating tRNA. The probable origin of the ACC triplet at the amino-acid-accepting arm of the tRNAs is discussed. A theoretical scheme for the generation of mutations in response to the presence of chemical or physical factors and to their effect on newly synthesized polypeptides, based on the polypeptide-mRNA co-synthesis hypothesis, is presented. Messenger RNA nicking during each step of peptide bond formation and *trans*-translation of a different mRNA or codon triplet, in response to ribosome stalling, are the essential elements of the proposed theory for directed mutability.

Key words: anticodon, mutation, reverse translation, ribosome, tRNA.

INTRODUCTION

Since its appearance 4.5 billion years ago, life evolved into billions of species. According to the Lamarckian hypothesis on the appearance of the various life forms, it is the environment that induces the morphological, biochemical and genetic changes, by an unknown yet mechanism. This theory (that the genetic changes are induced in a directed way by environmental factors) has not yet been supported by clear and unambiguous experimental evidence. The theory of Darwinian evolution states that environment selects for individuals best adapted to it and favors their propagation, without however explaining how these individuals arise in the first place. It is, in general, believed that mutations arise randomly but I have in vain searched Darwin's writings to find a phrase stating it clearly. Darwin gave a theoretical explanation on what happens *after* the appearance of the genetic change. Like Lamarck, he never gave an explanation on *how* these changes appear, which is the point comprising the heart of the evolutionary process. The concept of discrete genetic units (genes), located in the nucleus and determining the individual characteristics of organisms appeared much latter (de Vries, 1910), as a modification of Darwin's 'gemmule' theory (Darwin, 1868) and the concept of sudden genetic changes (mutations) of these hereditary determinants was developed in 1901 (de Vries, 1901). Even so, a molecular-level explanation of the mutation mechanism would be impossible at that time, since present day advances at the molecular biology did not exist. The scientific dispute over the exact mechanism creating the evolutionary useful genetic changes remains even today and focuses on the question if these changes are random or not or if both kinds occur and both used by environmental selective pressure.

* Corresponding author: tel.: +30 24210 93278, fax: +30 24210 93278, e-mail: igoun@agr.uth.gr

ANTICODON- OR CODON-AMINO ACID STRUCTURAL RELATIONS

The anticodon on the tRNAs is at a distance of about 70 Å from the terminal 3′-end adenine, where the

corresponding amino acid esterifies. This is used in support of the opinion that there is no stereochemical relation between the amino acid and its (anti)codon(s). In accordance to the principle of randomness in evolution, Crick (1968) supported the view that the nowadays observed correspondence between an amino acid and its codons constitutes a 'frozen accident' in evolution. By this it is meant that once an accidental, not stereochemically justified association of a tRNA with a particular amino acid is established, this is promulgated and adopted unaltered by the cell for all future translational processes.

Quite the opposite, a year before, Woese (1967) argued in favor of a stereochemical matching between amino acids and the base triplets of their corresponding codons. Stereochemical relationships between the codon (Freeland & Hurst, 1998; Chiuzano *et al*., 2000; Karasev & Stefanov, 2001; Biro *et al*., 2003) or anticodon (Jungck, 1978; Konecny *et al*., 1993) base sequence and the amino acid side chain have been reported. Also, in solution, some amino acids interact specifically with their codons (Yarus, 2000). A tabulation of the amino acids and their codons in a form of periodic table, published by Biro *et al*. (2003), sums up these structural interrelationships. The most obvious observations in these studies are that the second position of the (anti)codon is the most strongly related to the amino acid structure (Table 1). Pyrimidines (uracil or cytosine) in the second position of the codon are negatively associated with charged amino acids. Uracils in the second position of the codon (adenine in the anticodon) code only for hydrophobic amino acids and cytosines (guanine in the anticodon) code for either hydrophobic or uncharged hydrophilic (polar) amino acids. The charged amino acids and the stop signals are all coded by codons having purines in the second position.

In addition to the above observations, a simple inspection of a genetic code table shows that the third base (3′-end) of the codon, which is the first 5′-end base of the anticodon, is the least important in determining the coded amino acid. In half of the 16 possible base combinations of the first two positions of the codon, the third base makes no difference on the nature of the coded amino acid. In the other half of the cases, a different amino acid is coded or a stop signal is created mainly when the third base changes from pyrimidine to purine. With the only exception of the leucine (UUG/A)-phenylalanine (UUU/C) couple, amino acids with longer side chains are coded or stop signals are created when the third base of the codon changes from pyridine to purine (Table 2). Measuring the distances along the direction of the bond of the side-chain to the alpha-carbon to a plane perpendicular to the axis of this bond and contacting the most distal atom of the side-chain, even the couple Leu/Phe follows the above rule. This distance was measured to be only 3.21 Å in Phe instead of the 3.90 Å in Leu (Fig. 1). It is therefore concluded that the third base of the codon or the first base of the anticodon are in a stereochemical relation to the length of the amino acid side-chain along the direction of its bond to the alpha carbon. It is difficult to explain how a longer base in the third position of the codon could stereochemically relate to a longer amino acid and also to a stop signal. In translation, when a stop codon is reached by the ribosome, it is a very small molecule, water, that takes the position of the amino acid

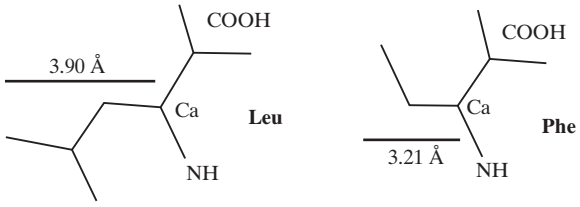

FIG. 1. Comparison of the Leu and Phe side-chain lengths along a direction parallel to their bond to the α-carbon. The Pymol molecular simulation program was used for the structures and for the measurements.

TABLE 1. Stereochemical relationships between the base in the second position of the (anti)codon and the amino acid side chain

Base in second position of the codon	Base in second position of the anticodon	Coded amino acid
U	A	hydrophobic
C	G	hydrophobic or uncharged hydrophilic (polar)
A	U	charged or polar or stop codon
G	C	charged or polar or glycine or stop codon

TABLE 2. Relationship between the third base of the codon and the length of the amino acid side chain. Distances from the α-carbon to the most distal carbon, oxygen or nitrogen atom of the side chain were measured using the Pymol molecular simulation program

Case	Codon	Anticodon	Amino acid	Side chain length (Å)
1	GAU/C	CUA/G	aspartic acid	3.57
	GAA/G	CUU/C	glutamic acid	4.89
2	CAU/C	GUA/G	histidine	4.62
	CAA/G	GUU/C	glutamine	4.96
3	UUU/C	AAA/G	phenylalanine	4.99
	UUA/G	AAU/C	leucine	3.90
4	UAU/C	AUA/G	tyrosine	6.31
	UAA/G	AUU/C	stop	
5	UGU/C	ACA/G	cysteine	2.75
	UGA	ACU	stop/tryptophan	
	UGG	ACC	tryptophan	5.93
6	AUU/C	UAA/G	isoleucine	3.91
	AUA	UAU	isoleucine/methionine	
	AUG	UAC	methionine	5.25
7	AGU/C	UCA/G	serine	2.46
	AGA/G	UCU/C	arginine/stop	7.23
8	AAU/C	UUA/G	asparagine	3.75
	AAA/G	UUU/C	lysine	6.29

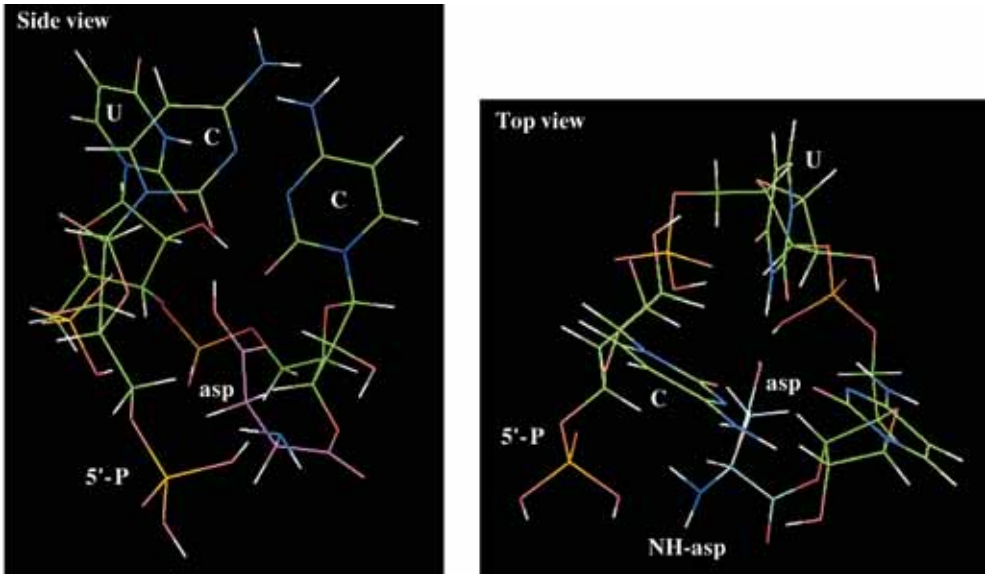

FIG. 2. A line representation for the aspartic acid anticodon (5′-p-CUC-3′-OH) modeled as a right-hand helix around the aspartic acid molecule. The aspartic acid is positioned inside the helical cavity, esterified to the 3′-end of the anticodon and the axis of its side-chain is parallel to the axis of the helix. The α-amino group of the amino acid and the 5′-phosphate of the anticodon are positioned to the outside of the helical cavity.

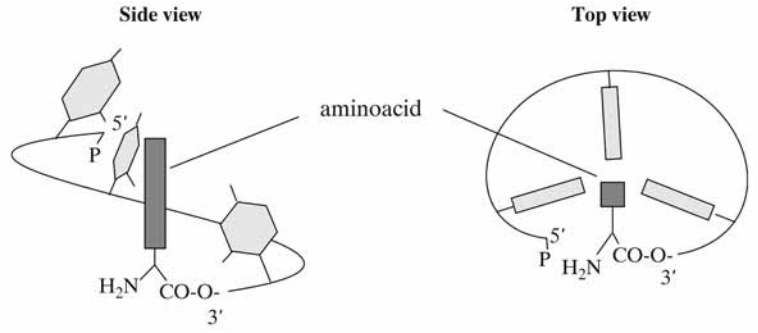

FIG. 3. A graphical representation of an anticodon triplet helically wrapped around its amino acid. The three bases are bent towards the center of the cavity, repulsed by the surrounding hydrophilic aqueous medium. A small side-opening in the almost complete helical step allows the amino acid amino group to face the exterior.

in the peptidyl transferase center of the ribosome and the synthesized polypeptide is hydrolyzed from its carrier tRNA.

At this point, I would like to propose a possible stereochemical interaction of anticodons to their related amino acids, which is compatible with permissible bond angle strain using the molecular simulation program Pymol and which is in agreement with the nowadays available observations on amino acid-(anti)codon relationships. According to this proposed scheme, the anticodon triplet surrounds the amino acid, forming an almost complete right-turn helical step around it, with the amino acid carboxyl group being esterified to the 3′-OH of the 3′-end ribose of the anticodon, as is normally done with the 3′-adenosine of the tRNAs (Fig. 2). The three bases are rotated to a conformation *anti* relatively to their riboses. That is, the base hydrogen-forming groups point towards the center of the helix. This brings them to bend towards the inside of the helix, forming a hydrophobic cavity (Fig. 3). The amino acid is positioned inside this cavity with the axis of the side-chain coinciding or being parallel to the axis of the helix, whereas its α-amino group protrudes to the outside. The central part of the wall of this cavity is primarily lined by groups of the second base of the anticodon. In contrast, little is lined by groups of the third base of the anticodon, because this is positioned higher up towards the roof of the cavity due to the helical conformation of the triplet. Purines are larger molecules than pyrimidines and create a smaller cavity excluding or severely limiting the access of water to its interior. Amino acids with hydrophilic or polar side-chains would be expected to be hydrated and therefore to be excluded from small cavities created by purines, especially when the purines are in the second position of the anticodon, that lines the major part of the cavity wall. Also, large cavities are expected to accommodate more water and are best suited as translational stop signals.

In primitive life, it is likely that tRNAs were much smaller and different in structure. Perhaps they were simple tri-ribonucleotides. Three nucleotides, joined in a series, are the minimum required to formulate a cyclic or helical structure, a sort of primitive pincher holding the amino acid. That would explain why the contemporary codons and anticodons consist of three bases and not one or two or four or more. Nature's economy kept this sufficient minimum of three bases throughout the entire evolution. It can be said that it is an event 'frozen' in evolution but not an 'accident'.

A HYPOTHESIS FOR A PRIMITIVE TRANSLATION MECHANISM CATALYZING THE SIMULTANEOUS SYNTHESIS OF PROTEINS AND OF THEIR CODING mRNA

Assuming that primitive tRNAs were consisted of the anticodon triplet only and that the complementary triplet codons were the primitive mRNAs, a model for the evolution of the contemporary mRNAs is presented in this section. A codon (mRNA) anticodon (tRNA) annealing model with the two triplets in helical conformation is presented in Figure 4. This arrangement was drawn with the help of the Pymol molecular simulation program, using purely bond rotations and avoiding overlapping of atoms when they were represented as spheres. Therefore, the model does not involve any significant bond stretching or bending and is sterically and thermodynamically possible. The codons are tri-phosphorylated at their 5′-ends. The

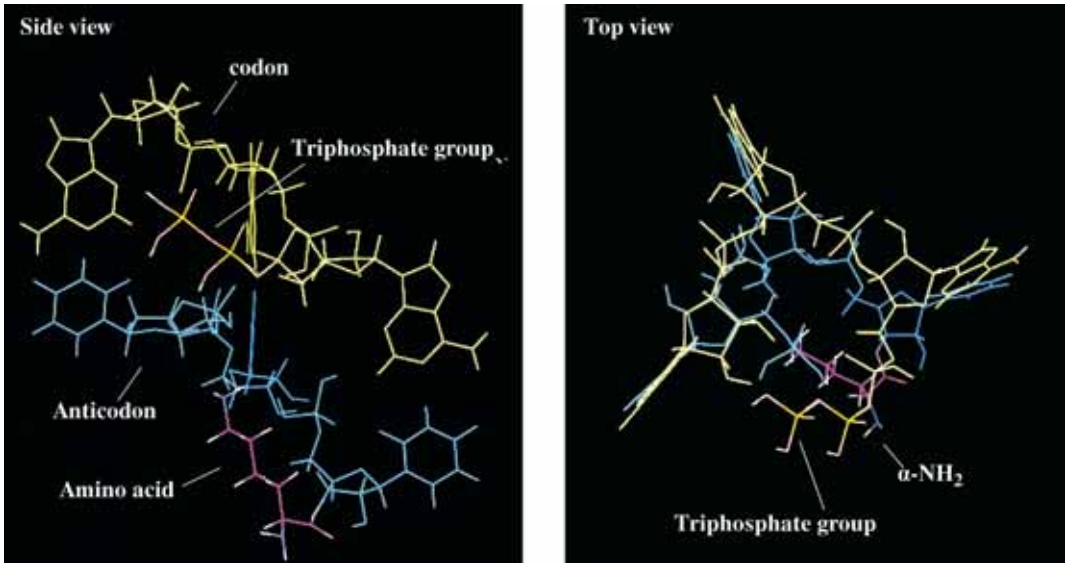

FIG. 4. A model for the annealing of an anticodon triplet to its corresponding codon triplet. The model has been build using the Pymol molecular simulation program. The codon (yellow) is triphosphorylated at its 5′-end and the anticodon (blue) bears its corresponding amino acid (magenta) esterified to its 3′-end. The triphosphate and the α-amino groups protrude to the helix exterior through the helix side crevice.

three bases of the anticodon are now more upright relatively to their position in the non-annealed form and the same holds for the codon triplet. Both triplets form an almost complete helical step around the amino acid. The codon triphosphate group at its 5′-end is protruding to the aqueous medium through the helix side crevice and is bent upward. The amino acid is esterified to the 3′-OH of the anticodon, the axis of its side-chain is parallel to the axis of the helix and its amino α-NH_2 also exits through the side opening of the helix.

In a very primitive translation apparatus, two such codon (mRNA)-anticodon (tRNA) complexes, each carrying its own amino acid, could align their sides in

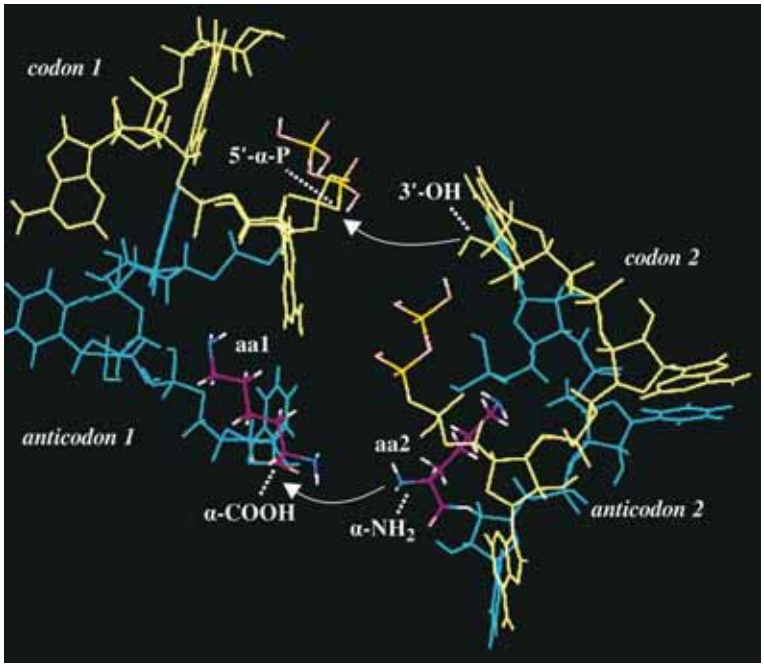

FIG. 5. Simultaneous formation of a peptide bond and a phosphodiester bridge in two aligned triphosphorylated codon (yellow) – anticodon (blue) complexes in a hypothetical primitive translation apparatus. The formed dipeptide remains attached to the incoming anticodon (tRNA) 2 whereas the 3′-end of codon 2 is joined to the 5′-end of codon 1, forming a mRNA larger by 3 nucleotides.

a way bringing the α-amino group of one amino acid close to the esterified carboxyl group of the other amino acid (Fig. 5). This would also bring the triphosphate group of one of the codons close to the 3′-OH of the other. In such an alignment, formation of a peptide bond between the two amino acids is possible by nucleophilic attack of the amino group of one amino acid to the carboxyl group of the other. Simultaneously, the 3′-OH of one of the codons can attack (as a nucleophile) the α-phosphate group of the triphos-

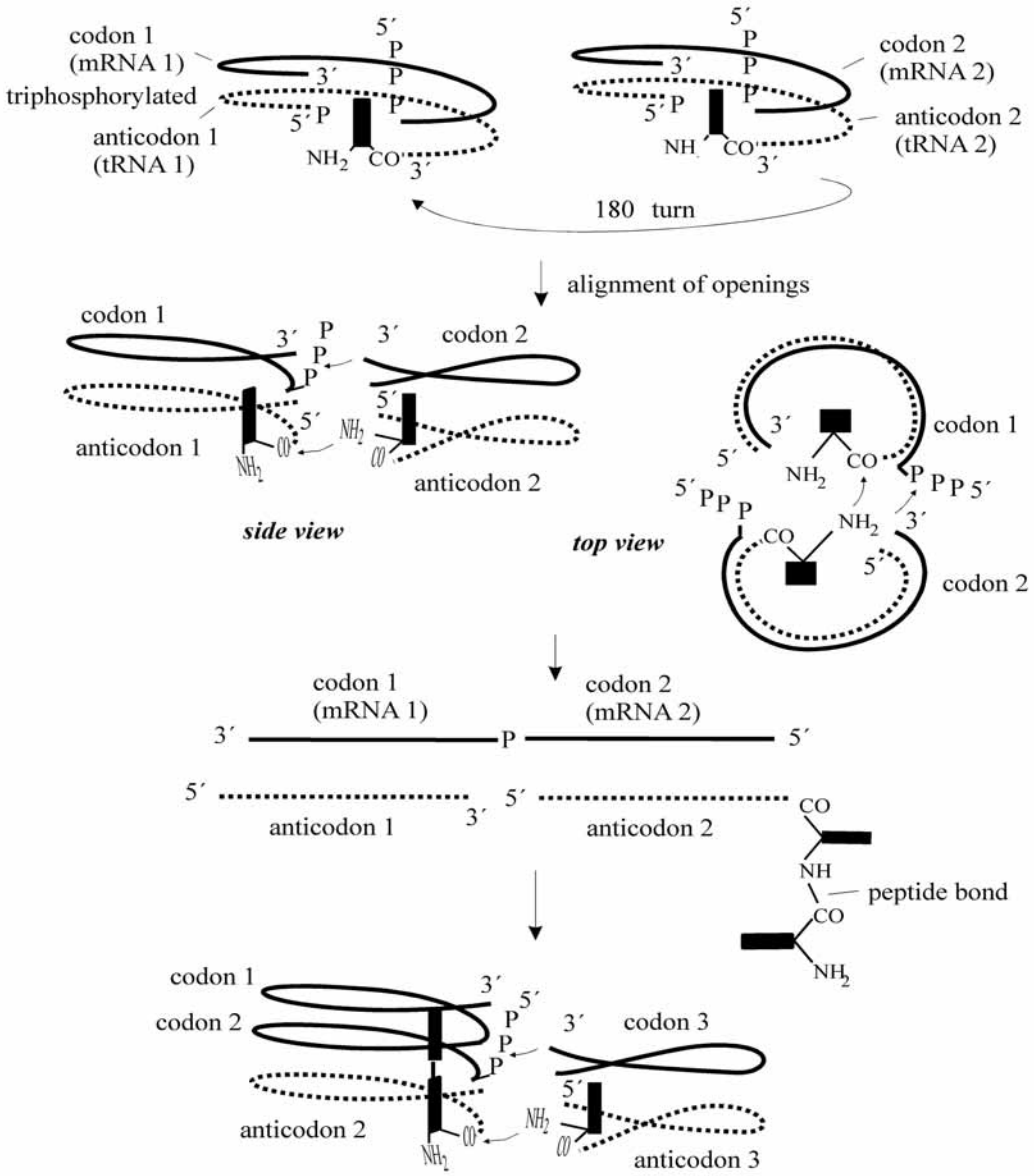

FIG. 6. A graphical representation of the proposed primitive mechanism of simultaneous protein and mRNA synthesis, from which contemporary ribosomes and translation might have evolved. According to this model, early tRNAs were consisted of anticodon triplets bearing their amino acid at their 3′-end and primitive mRNA consisted of the corresponding codon triplets, triphosphorylated at their 5′-end. Two codon-anticodon complexes are aligned with the side openings of their helices opposite to each other. This brings the α-amino group of one amino acid close to the carboxyl group of the other and the 3′-OH of one of the codons close to the triphosphorylated 5′-end of the other. Formation of the peptide bond occurs between the two amino acids as well as formation of a phosphodiester bridge between the two codons. New rounds of this process with new codon-anticodon complexes elongates simultaneously the polypeptide, by adding amino acids to its carboxyl end and its coding mRNA by adding codon triplets to its 3′-OH end. The elongating polypeptide is surrounded helically by the elongating mRNA, in a way that three bases are found around each amino acid. This means that the helix step is around 3.6 Å , same as the distance between successive peptide bonds.

phate chain at the 5′-end of the other codon, effecting the joining of the two codons (mRNAs) via a phosphodiester bridge to form a 6-nucleotide mRNA. The dipeptide formed remains attached to the incoming tRNA, as is done in contemporary translation and is used in a new round of simultaneous peptide bond-phosphodiester bridge formation with a new incoming codon-anticodon complex. Gradual elongation of the polypeptide chain is accompanied by simultaneous elongation of the coding mRNA, with new codon triplets being added to its 3′-end. A more detailed graphical representation of the proposed model for this primitive translation-mRNA formation procedure is given in Figure 6.

THE EVOLUTION OF MODERN tRNAs

A first look at the secondary structure of tRNAs (Fig. 7A) reveals a similarity of the central part of its clover leaf structure to a Holiday cross (Holiday junction). Such crosses are locations where complementary regions of two different double stranded nucleic acids or two complementary regions of the same double stranded nucleic acid hybridize to form a four-strand cross. Nucleic acid nicking and ligation occurs in such regions resulting in recombination. When the rehybridizing regions are part of the same double-stranded nucleic acid molecule they often are encountered as palindromic sequences and allow transformation of the cross back into the extended linear double-stranded nucleic acid format and the reverse. Also, many tRNAs have an intron starting 1-2 bases downstream of the anticodon and extending to the variable loop. Processing to a mature tRNA involves removal of the intron by nicking at its ends, which could be considered as related to nicks occurring during recombination in Holiday crosses. A second possibility of tRNA nicking at the point close to the anticodon is that case of transfer-messenger RNAs (tmRNAs). In tmRNAs, a mRNA is inserted in the tRNA, starting a base downstream of the anticodon (Keiler *et al*., 1996; Zwieb *et al*., 2001; Ivanova *et al.*, 2007), at a position where the introns also are found. This inserted mRNA is translated as the stalled ribosomes change the readable transcript (trans-translation) as a way to overcome the translational blockage.

The lengths of the amino acid accepting arm and of the anticodon arm of the contemporary tRNAs are such that, provided their base sequence was complementary in the beginning of evolution, the tRNA clover leaf structure (Holiday cross-like structure) can transform into the equivalent straight double strand format (Fig. 7B). Considering the nicking events, contemporary tRNAs can be considered as composed of two halves. One of the halves could have been evolved from the elongating mRNA (codon) of the hypothesis I presented (Fig. 6) and the other from its complementary tRNA (anticodon). A new incoming codon (XXX)-anticodon (ACC) pair, carrying a new amino acid, could be annealed with the two halves of the hypothetical ancient ‘tRNA’-‘mRNA’ pair. Formation of phosphodiester bridges would join the codon XXX to the 5′-end of the elongating ‘mRNA’. The codon, which nowadays is always ACC, was attaching to the 3′-end of the ‘mRNA’ but also perhaps to the 5′-end of the ‘tRNA’, resulting in the extension at both molecules. At the beginning, a different codon-anticodon pair of complementary triplets was used for each round of peptide bond formation. Evolutionary modifications of this early scheme selected the ACC as the default triplet whereas the triplet XXX was different in each round of peptide elongation. Formation of phosphodiester bonds at both sides of the codon and anticodon is a possibility, and would have produced a circular single stranded RNA, with base complementarity along its entire length, as is the case with contemporary viroids. The existence of such a structure should however be transitionary, as it would have to be nicked again for a new round of peptide elongation. The present day attachment of the ACC triplet to the 3′-end of modern tRNAs could be an evolutionary relic of the formation of a closed circular single-stranded ‘tRNA’-‘mRNA’ molecule. The trinucleotide ACC, attached to every contemporary tRNA during its processing, is the anticodon for glycine, the simplest and perhaps the most ancient amino acid. As the ribosomes were evolutionary developing towards using already existing mRNAs and abandoned their mRNA synthesizing activity, there was no more need for a different triplet in each round of peptide bond formation and ACC was adopted as the default triplet, as an evolutionary remnant of no specific coding role.

The idea that tRNAs are composed of two halves is not new. In a series of articles, Di Giulio (1992, 1995, 1999, 2004, 2006a, 2006b, 2008a, 2008b, 2009) provided very convincing proof that tRNAs are made of two similar halves, a 3′-half and a 5′-half, transcribed from two different copies of a duplicated gene, that at ancestral times were positioned apart. The intervening sequence between the two copies of the gene corresponds to the introns encountered even to-

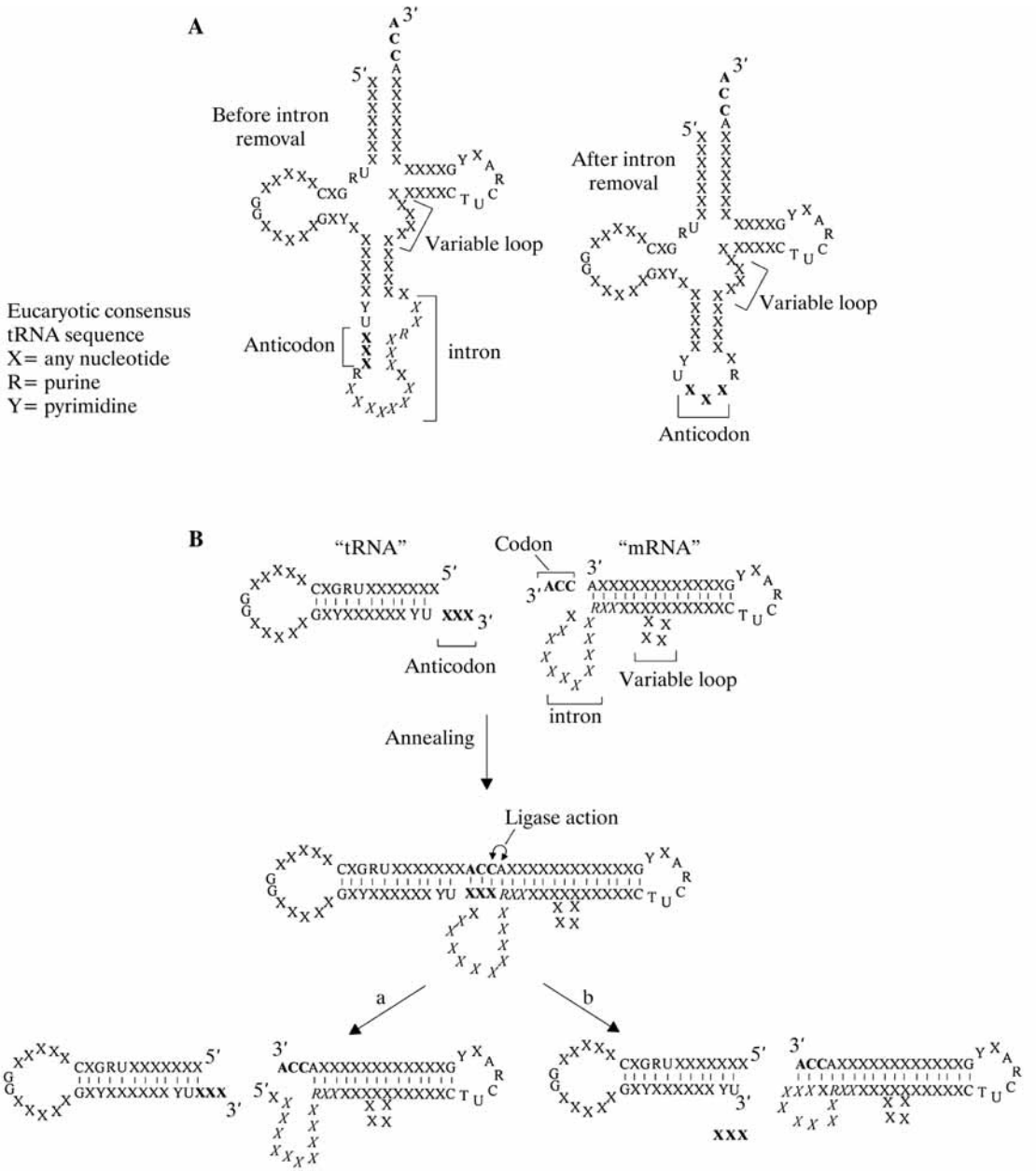

FIG. 7. Probable evolution of modern tRNAs from primitive tRNA-mRNA couples co-synthesizing proteins and mRNAs. A: Consensus sequence of contemporary eukaryotic tRNA and their unprocessed precursors having intron regions. The structure reminisces a Holiday cross where nucleic acid nicks and recombination takes place. Such conformations often result from palindromic sequences that can rearrange into a linear double-stranded chain, as shown in plane B. B: A straight chain format of a primitive tRNA, shown as two halves, one of which evolved from ancient 'tRNA' and the other from an elongating ancient 'mRNA', according to the hypothesis proposed in this article. A new incoming codon/anticodon pair (XXX-ACC) hybridizes and anneals to the two halves. The contemporary codons were actually anticodons in the primitive translation apparatus. Formation of a phosphodiester bridges, would join the codon XXX to the 5′-end of the elongating 'mRNA'. The codon, which nowadays is always ACC, was attaching to the 3′-end of the 'mRNA' but also perhaps to the 5′-end of the 'tRNA', resulting in the extension at both molecules. At the beginning, a different codon-anticodon pair of complementary triplets was used for each round of peptide bond formation. Evolutionary modifications of this early scheme selected the ACC anticodon for glycine as the default triplet whereas the codon XXX was different in each round of peptide elongation. Formation of phosphodiester bonds at both sides of the codon and anticodon is a possibility, and would have produced a circular single stranded RNA, with base complementarity along its entire length, as is the case with contemporary viroids. The existence of such a structure should however be transitionary, as it would have to be nicked again for a new round of peptide elongation. The present day attachment of the ACC triplet to the 3′-end of modern tRNAs could be an evolutionary trace of formation of a closed circular single-stranded 'tRNA'-'mRNA' molecule.

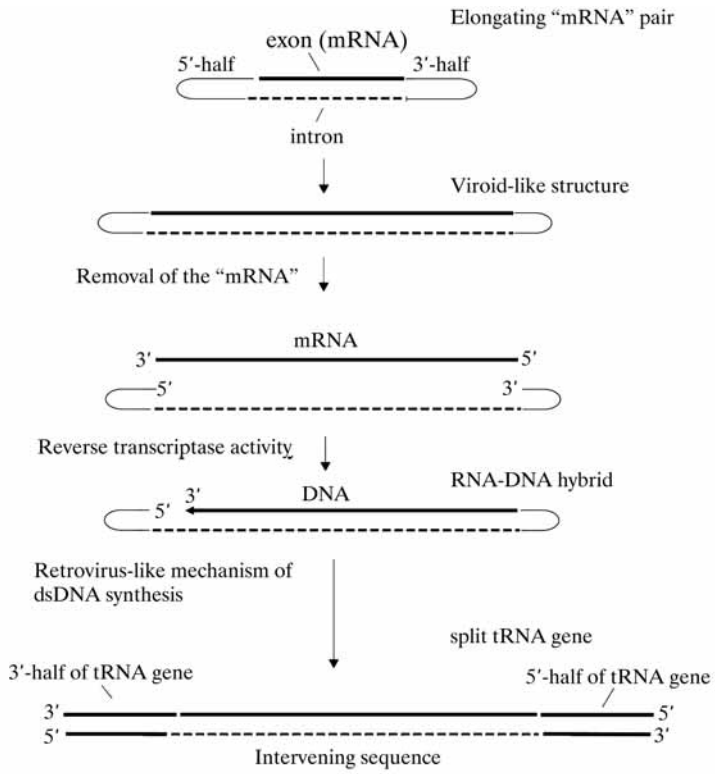

FIG. 8. A hypothesis for the evolution of split tRNA genes from a primitive elongating ssRNA, as proposed in this article. An mRNA is synthesized together with the polypeptide it codes for. At some time, the mRNA is excised. Then, a dsDNA synthesis mechanism, similar to that used by contemporary retroelements, produces a split tRNA gene having the two tRNA halves at its ends separated by an intervening sequence. Gradual elimination of the intervening sequence, produced the contemporary tRNA genes, having no intervening sequence and promoter regions in both halves.

day in some tRNAs. From this ancestral form evolved the common contemporary tRNA genes that are have the two halves in succession without any intervening part. Di Giulio's model supposes a gene duplication event. With what I propose in this section, an alternative mechanism of creating a split tRNA gene is apparent. A diagrammatic representation of this alternative evolutionary route is given in Figure 8.

According to this scheme, the continuously elongating 'mRNA'-'tRNA' pair assumes a circular ssRNA format related and probable ancestral to contemporary viroids. After removal of the mRNA part, the employment of reverse transcriptase and of a retroelement-like mechanism for dsDNA synthesis, would lead to the formation of a split tRNA gene having an intervening sequence between its 3′ and 5′ halves. It is interesting to note that, according to this proposed model, any promoter sequences residing at the two ends of the viroid-like ssDNA (which become the two halves of the tRNA gene), will appear separately inside the two halves tRNA gene. Indeed, the promoters of tRNA genes are split in two parts (A and B regions), located inside on two halves of the gene and promoters exist inside the long terminal repeats (LTR) of retroviruses, retrotransposons and retroelements in general.

THE EVOLUTION OF THE CONTEMPORARY TRANSLATION APPARATUS

We could assume that the 5′-ends of the codon and perhaps also of the anticodon triplets of the primitive protein-mRNA co-synthesizing system proposed here were already triphosphorylated. This could be related to the way these triplets were synthesized from nucleotide triphosphates. However, in later tRNAs and protein synthesizing apparatuses that evolved from them, nicking of the tRNA strand would result in 'tRNA' and 'mRNA' halves having monophosphorylated free 5′-ends, unable to ligate to the 3′-ends of the new coming ACC and XXX triplets.

The present day ribosome action could have evolved from a primitive protein-mRNA co-synthesizing complex as shown in Figure 9. The ACC triplet remained simply as an amino acid-carrying end structure. Its coding role has been replaced by an already existing pre-synthesized mRNA. We know that RNA ligases use an ATP to form an AppN group via a 5′-pp-5′ bond at the 5′-end of the nucleic acid they are about to ligate (Sugino *et al*., 1978; Ho *et al*., 2004). It can be hypothesized that contemporary ribosomes do the same using GTP. In each round of peptide bond formation, nicking of the mRNA just upstream of the

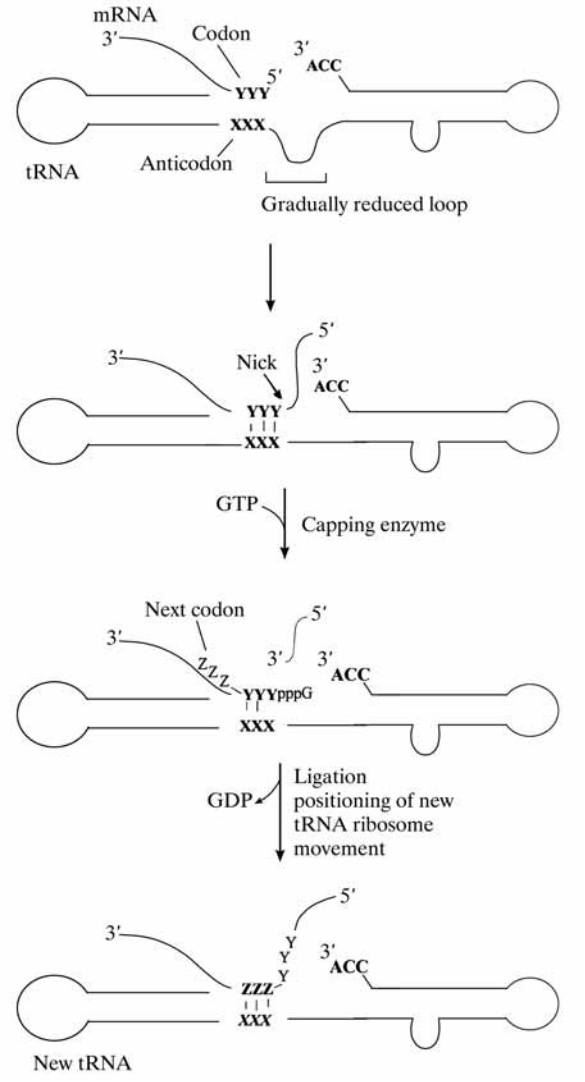

FIG. 9. The emergence of the contemporary translation mechanism form the primitive protein-mRNA co-synthesizing system proposed in this article. The ACC 3′-triplet of the modern tRNAs has only amino-acid carrying role. Its coding role has been replaced by the codons of an incoming mRNA. The GTP necessary for polypeptide chain elongation could be used to form a guanylic acid cap (GpppN) at the 5′-end, generated by mRNA nicking just upstream of the codon, in every round of peptide bond formation. The cap function could then be used for re-ligation of the two mRNA halves, in a way causing ribosome movement along the mRNA (or mRNA movement along the ribosome) to form a new complex with a new incoming tRNA.

read codon is a possibility that would explain the appearance of mRNA fragments in stalled ribosomes (Sunohara *et al*., 2004; Richards *et al*., 2008; Dreyfus 2009), a case in which enough time is given to the fragments to be freed from the ribosomal complex. The GTP necessary for moving the ribosome to the next codon could be actually used to cap the 5′-end generated by the nicking. A cap structure similar to that found at the 5′-end of eukaryotic mRNA is proposed or of the type GppN. The cap 5′-pp(p)-5′ group could be used for religation of the two mRNA halves, in a way causing the ribosome movement. The involvement of nicking would agree therefore with the appearance of mRNA fragments during ribosome stalling and with the necessity for GTP. It is easily explained also as an evolutionary trace of the proposed primitive mechanism of protein-mRNA co-synthesis, where the codon-anticodon triplet complex was independent. However, cap methylation, as commonly occurs in eukaryotic mRNAs, is expected to block the cap participation in ligation reactions. Prokaryotic and certain viral newly synthesized mRNAs have no cap structures but a triphosphate group. Yet, a similar translation mechanism involving transitional mRNA nicking and capping could well be operating in prokaryotes as well.

A point to be addressed is if during modern translation the tRNAs need to assume a conformation different than the normal one of the Greek capital gamma letter (Γ). More specifically, the question is if the tRNAs must indeed be rearranged into a viroid-like, circular ssRNA for the model of Figure 8 to be workable. All the model requires is a pairing between the codon on the new-coming mRNA and the anticodons on the tRNA. This can occur without any significant change in the tRNA conformation and transition of the tRNA into its proposed primitive stretched format is not necessary. Yet minor structural changes, related to the nicking at the 5′-end of the codon, and the guanylyl capping of the codon cannot be excluded.

A POSSIBLE MECHANISM FOR INDUCING MUTATIONS IN RESPONSE TO CHEMICAL AND PHYSICAL FACTORS

It has long been established that the early genetic material was RNA rather than DNA (Barbieri, 1981; 1985). DNA is a more stable storage format of the genetic information RNA is far more amenable to structural changes and a better candidate for mutational events. Once mutated, an RNA chain can be stored as DNA by the action of reverse transcriptase. The still unresolved question is if the mutational changes incurring on RNA are accidental or somehow induced in a directed way by chemical, physical or other unknown factors. In this section, a probable mechanism for non accidental changes in the base sequence of RNA will be discussed, in accordance with the hy-

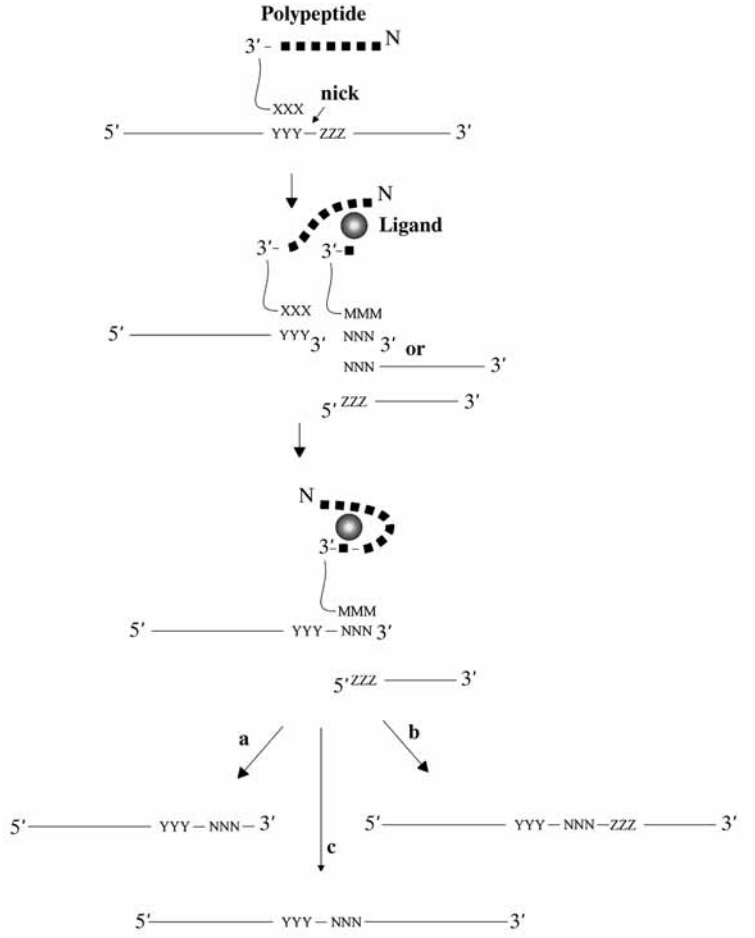

FIG. 10. A theoretical model for a translational mechanism allowing the introduction of point or large sequence mutations in a decoded transcript and simultaneously creating a mutated polypeptide fitted to the new needs of the cell. Incorrect folding of already synthesized part of a polypeptide, caused by ligand binding or by physical factors could cause ribosome stalling. Nicking of the transcript allows insertion of a new codon sequence, whose complementary amino acid-tRNA interaction with the incorrectly folded polypeptide restores its function. One or a few such new codons can be introduced before the ribosome resumes decoding of the rest of the original transcript (cases a and b). Alternatively, the ribosome can continue with the decoding of a new transcript (transcript shifting) resulting in the simultaneous formation of new hybrid mRNA and polypeptide containing fused parts of two or more different transcripts and of the corresponding parts of the polypeptide fragments they code. Genetically fixing the mutated transcript is done by reverse transcriptase.

pothesis of the existence of an early protein-mRNA co-synthesis apparatus in very primitive life and its evolution into the modern ribosomal translation process. According to this directed mutation mechanism, the alterations in the base sequence of an mRNA occur during its translation, when it is stalled by chemical factors interacting with the already synthesized part of the polypeptide.

A diagram of the model is given on Figure 10. The already synthesized part of the polypeptide could bind to a new non-self compound that accidentally entered the cell or that has been introduced as an antigen-antibody (receptor) complex. A change in the folding of the polypeptide portion is expected by ligand binding or even by physical factors, such as high or low temperatures, osmotic conditions inside the cell, ultraviolet radiation. Depending on the configuration the polypeptide assumes under the influence of such chemical or physical factors, ribosome stalling could occur if the next incoming amino acid-tRNA is stereospecifically hindered from positioned at the A site or if the ligand attracts preferentially a different amino acid. The 'improper' amino acid-tRNA will bring with itself its complementary codon triplet, which could be trailed by a whole new mRNA. This codon will be incorporated to the already translated mRNA. If it is a whole new mRNA we have a case of trans-translation in which the ribosome changes the transcript it translates (case c). This could be a mechanism by which protein, transcript and gene fusion (via reverse transcriptase) occurs. If it is a simple codon triplet inserted into the mRNA under translation, the ribosome could continue reading the left over 3′-half of the mRNA (case b) or it could terminate the translation after a few more 'improper' codons (case a). In short, the already synthesized part of a polypeptide starts undergoing folding before the completion of the entire polypeptide synthesis. Uncommon folding induced by a variety of factors, that could produce an enzymatically or structurally nonfunctional polypeptide, leads to translational stalling. Overcoming the blockage is done by abandoning the decoding of the regular transcript, temporarily for the length of one or more codons or for the rest of the

message and instead shifting to the use of new tRNA-triplet codon complexes, whose amino acid part interacts with the blocking ligand or the incorrectly folded nascent part of the polypeptide and which corrects the folding mode to a functional format.

The model requires mRNA nicking just upstream of the codon under translation and use of GTP for capping the 5′-ends before ligation. Triphosphorylated prokaryotic mRNAs and eukaryotic mRNA bearing non-methylated cap structures, if selected, could enter the ligation step, resulting in mRNA fusion hybrids that become hybrid genes via reverse transcriptase.

THE POSSIBILITY OF REVERSE TRANSLATION

Since the discovery of reverse transcriptase activity, researchers have been wondering if the second part of the central dogma, the translation, could also be reversed. Published articles by Mekler (1967), Cook (1977), Craig (1981), Nashimoto (2001) and Biro (2004) describe in detail the biological question involved. Undoubtedly, as usually happens, a great number of other scientists have thought about it without having ever written on the subject. Yet, a hypothetical reverse translation mechanism and a reverse translatase complex have been sought without success. By reverse translation is here defined the synthesis of an mRNA by a cellular enzymatic complex that uses as template an existing polypeptide and back translates it into the corresponding transcript. It was thus hoped that a complete reversal of the central dogma would be proved.

According to Mekler (1967) the sequential binding of amino acids along an antigen, equivalent to the ligand proposed in this polypeptide-mRNA co-synthesis theory, brings the anticodons of their tRNAs in a linear arrangement that can be used as template by a polymerase for the synthesis of the complementary mRNA. The mRNA can then be reverse transcribed into a new DNA coding sequence for an antibody recognizing the antigen. In this sense, the Mekler theory is not proper reverse translation involving mechanistic reversal of the translation process, because it does not involve the degradation of the back-translated polypeptide, neither synthesis of mRNA in the reverse 3′ to 5′ direction. This theory differs significantly from the protein-mRNA co-synthesis hypothesis presented in this article in that it does not involve simultaneous synthesis of the antibody polypeptide and assumes mRNA synthesis by some nucleic acid polymerase using nucleotide triphosphates, as all nucleic acid polymerases do. In contrast, the polypeptide-mRNA co-synthesis theory proposes the simultaneous synthesis of the polypeptide by the ribosomes or by a ribosome-like entity using independent codon triplets or cutting them off from already existing mRNAs.

It must however be noticed that reverse transcription is not mechanistically reverse to transcription. Reverse transcriptase is using RNA as a template to synthesize DNA but it is still a polymerase as are all RNA polymerases and like them catalyzes polymerization of nucleoside triphosphates by a very similar enzymatic mechanism. It would be a reverse operating enzyme if it could act in reverse and depolymerize a RNA transcript into its nucleoside triphosphates. Such backward enzymatic mechanism has not been observed and reverse transcriptases are a relative of the RNA polymerases and other nucleic acid polymerases, with common ancestor and differentiated to use RNA templates instead of being of an opposite activity. In a similar way, a real reverse translatase activity, mechanistically reverse to the ribosomal translation mechanism, that would be able to read along a polypeptide from its carboxyl end, depolymerizing it or not and synthesizing the corresponding mRNA starting from its 3′-end, has never been detected. A semi-artificial reverse translation system using engineered tRNA has been reported by Nashimoto (2001) and was successful with one type of codon at least. Yet, the real question is if a natural intracellular reverse translatase exists or not. Its existence could have great consequences for the evolutionary theory. It could mean that the origin of genetic variability could be at the polypeptide level and they could then pass to the DNA through the sequential action of reverse translatase and reverse transcriptase. However, it is very difficult to imagine how one or more amino acids along a pre-existing polypeptide, free from ribosomes, could be replaced by the action of a chemical of physical factor. It is more plausible to assume that the amino acid changes are co-occurring with changes in the corresponding mRNA codon during the synthesis of the polypeptide in the ribosomes, by a mechanism that evolved from a primitive protein-mRNA co-synthesis system.

If a reverse translatase exists, either as backward ribosomal activity or catalyzed by a different complex, related to ribosomes, it would have to add a triplet codon to the synthesized mRNA for each amino acid

read along the polypeptide template. This would necessitate the preexistence of triphosphorylated tri-ribonucleotides in the cell, as is proposed to have been in existence in early life by the protein-mRNA co-synthesis hypothesis presented in this article. Backward movement of ribosomes by one codon along the mRNA has been proved to occur under certain conditions (Qin *et al*., 2006; Konevega *et al*., 2007; Youngman & Green, 2007). However, what really has to be proved is backward movement along the polypeptide chain and addition of a triplet codon to the 5′-end of the mRNA.

FINAL REMARKS

It was the purpose of this article to present a possible molecular mechanism by which mutations could be created and inherited in a directed way and in response to the specific chemical or physical factor that elicits them. It has been argued that a stereochemical correspondence of the anticodons to their amino acids exists, if we accept that the early anticodon triplets were wrapped around their amino acid. Then a hypothetical early life mechanism that used pairs of anticodon-codon triplets and their amino acids for simultaneous synthesis of a polypeptide and its mRNA was presented. It was explained how modern tRNAs and translation could have evolved from such a primitive protein-mRNA co-synthesizing system. Finally, it was proposed that mutations are created simultaneously at the polypeptide and its corresponding mRNA level and a hypothetical scheme of the involved mechanism was given. Certainly a great deal of experimental work is needed to fully verify these theoretical consideration that were based on and were attempting to consolidate the available (until now) experimental data. Theories have never fully covered all the details from their beginning, but only attempt to set new directions for solving a biological problem by further investigation. Even today, technological advances might not be sufficiently progressed to experiment with some of the points of the proposed theory. For example, chemical or enzymatic synthesis of anticodon/codon tri-phosphorylated triplets is possible at a large cost and synthesis of the same triplets with non-methylated guanylic acid caps is impossible.

ACKNOWLEDGMENTS

The author would like to thank his colleagues for critically reading the manuscript.

REFERENCES

Barbieri M, 1981. The ribotype theory on the origin of life. *Journal of Theoretical Biology*, 91: 545-601.

Barbieri M, 1985. *The semantic theory of evolution.* Hardwood Academic Publishers, London.

Biro JC, 2004. Seven fundamental, unsolved questions in molecular biology. Cooperative storage and bi-directional transfer of biological information by nucleic acids and proteins: an alternative to ‘central dogma’. *Medical Hypotheses*, 63: 951-962.

Biro JC, Benyó B, Sansom C, Szlávecz Á, Fördös G, Micsik T, Benyó Z, 2003. A common periodic table of codons and amino acids. *Biochemical and Biophysical Research Communications*, 306: 408-415.

Chiusano ML, Alvarez-Valin F, Di Giulio M, D’Onofrio G, Ammirato G, Colonna G, Bernardi G, 2000. Second codon positions of genes and the secondary structures of proteins. Relationships and implications for the origin of the genetic code. *Gene*, 261: 63-69.

Cook ND, 1977. The case for reverse translation. *Journal of Theoretical Biology*, 64: 113-135.

Craig R, 1981. The theoretical possibility of reverse translation of proteins into genes. *Journal of Theoretical Biology*, 88: 757-760.

Crick FHC, 1968. The origin of the genetic code. *Journal of Molecular Biology*, 38: 367-389.

Darwin C, 1868. *The variation of animals and plants under domestication.* John Murray, London.

De Vries HM, 1901. *Die Mutationstheorie. Versuche und Beobachtungen über die Entstehung von Arten im Planzenreich von Hugo De Vries. Erster Band. Die Entstehung der Arten durch Mutation*. Verlag von Veit & Comp., Leipzing.

De Vries H, 1910. *Intracellulare Pangenesis.* The Open Court Publishing Company, Chicago.

Di Giulio M, 1992. On the origin of the transfer RNA molecule. *Journal of Theoretical Biology*, 159: 199-214.

Di Giulio M, 1995. Was it an ancient gene codifying for a hairin RNA that, by means of direct duplication, gave rise to the primitive tRNA molecule? *Journal of Theoretical Biology*, 177: 95-101.

Di Giulio M, 1999. The non-monophyletic origin of tRNA molecule. *Journal of Theoretical Biology*, 197: 403-414.

Di Giulio M, 2004. The origin of the tRNA molecule: implications for the origin of protein synthesis. *Journal of Theoretical Biology*, 226: 89-93.

Di Giulio M, 2006a. *Nanoarchaeum equitans* is a living fossil. *Journal of Theoretical Biology*, 242: 257-260.

Di Giulio M, 2006b. The non-monophyletic origin of the tRNA molecule and the origin of the genes only after the evolutionary stage of the last universal common ancestor (LUCA). *Journal of Theoretical Biology*, 240: 343-352.

Di Giulio M, 2008a. Permuted tRNA genes of *Cyanidio-*

schyzon merolae, the origin of the tRNA molecule and the root of the Eukarya domain. *Journal of Theoretical Biology*, 253: 587-592.

Di Giulio M, 2008b. An extension of the coevolution theory of the origin of the genetic code. *Biology Direct*, 3: 37. doi: 10.1186/1745-6150-3-37.

Di Giulio M, 2009. Formal proof that the split genes of tRNA of *Nanoarchaeum equitans* are an ancestral character. *Journal of Molecular Evolution*, 69: 505-511.

Dreyfus M, 2009. Killer and protective ribosomes. In: Condon C, ed. *Molecular Biology of RNA Processing and Decay in Prokaryotes, vol. 85.* Academic Press, London, Amsterdam, Burlington, San Diego: 423-466.

Freeland SJ, Hurst LD, 1998. The genetic code is one in a million. *Journal of Molecular Evolution*, 47: 238-248.

Ho CK, Wang LK, Lima CD, Shuman S, 2004. Structure and mechanism of RNA ligase. *Structure*, 12: 327-339.

Ivanova N, Lindell M, Pavlov M, Schiavone LH, Wagner EGH, Ehrenberg M, 2007. Structure probing of tmRNA in distinct stages of *trans*-translation. *RNA*, 13: 713-722.

Jungck JR, 1978. The genetic code as a periodic table. *Journal of Molecular Evolution*, 11: 211-224.

Karasev VA, Stefanov VE, 2001. Topological nature of the genetic code. *Journal of Theoretical Biology*, 209: 303-317.

Keiler KC, Waller PRH, Sauer RT, 1996. Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science*, 271: 990-993.

Konecny J, Eckert M, Schoniger M, Hofacker GL, 1993. Neutral adaptation of the genetic code to double-strand coding. *Journal of Molecular Evolution*, 36: 407-416.

Konevega AL, Fischer N, Semenkov YP, Stark H, Wintermeyer W, Rodnina MV, 2007. Spontaneous reverse movement of mRNA-bound tRNA through the ribosome. *Nature Structural & Molecular Biology*, 14: 318-324.

Mekler LB, 1967. Mechanism of biological memory. *Nature*, 215: 481-484.

Nashimoto M, 2001. The RNA/protein symmetry hypothesis: Experimental support for reverse translation of primitive proteins. *Journal of Theoretical Biology*, 209: 181-187.

Qin Y, Polacek N, Vesper O, Staub E, Einfeldt E, Wilson DN, Nierhaus KH, 2006. The highly conserved LepA is a ribosomal elongation factor that back-translocates the ribosome. *Cell*, 127: 721-733.

Richards J, Sundermeier T, Svetlanov A, Karzai AW, 2008. Quality control of bacterial mRNA decoding and decay. *Biochimica et Biophysica Acta*, 1779: 574-582.

Sugino A, Snopek TJ, Cozarelli NR, 1977. Bacteriophage T4 RNA ligase: reaction intermediates and interaction of substrates. *Journal of Biological Chemistry*, 252: 1732-1738.

Sunohara T, Jojima K, Yamamoto Y, Inada T, Aiba H, 2004. Nascent-peptide-mediated ribosome stalling at a stop codon induces mRNA cleavage resulting in nonstop mRNA that is recognized by tmRNA. *RNA*, 10: 378-386.

Woese CR, 1967. *The genetic code*. Harper and Row, New York.

Yarus M, 2000. RNA-ligand chemistry: a testable source for the genetic code. *RNA*, 6: 475-484.

Youngman EM, Green R, 2007. Ribosomal translocation: LepA does it backwards. *Current Biology*, 17: R136-R139.

Zwieb C, Guven SA, Wower IK, Wower J, 2001. Three-dimensional folding of the tRNA-like domain of *Escherichia coli* tmRNA. *Biochemistry*, 40: 9587-9595.