BioSystems 185 (2019) 104024

Contents lists available at ScienceDirect

BioSystems

journal homepage: www.elsevier.com/locate/biosystems

Evolution of the genetic code: The ambiguity-reduction theory

Marcello Barbieri

Dipartimento di Morfologia ed Embriologia, Via Fossato di Mortara 64a, 44121 Ferrara, Italy

ARTICLE INFO

Keywords:
Genetic code
Stereochemical theory
Code arbitrariness
Code ambiguity
Error minimization
Statistical proteins
Ribosomes
Adaptors
Biological specificity

ABSTRACT

The experimental evidence has shown that the genetic code is based on *arbitrary*, or *conventional*, rules, in the sense that any codon can be associated to any amino acid, and this means that there is no deterministic link between them. This is in sharp contrast with the traditional paradigm of the *stereochemical theory*, which claims that the rules of the genetic code were determined by chemistry, and more precisely by stereochemical affinities between codons and amino acids. The discovery that the genetic code is based on arbitrary rules, on the other hand, raised a formidable problem: how can such rules exist in Nature? In order to deal with this problem, it has been pointed out that the rules of an arbitrary code could not come fully formed into existence. The first genetic code, in other words, was necessarily ambiguous, and its evolution took place with a mechanism that systematically reduced its ambiguity and eventually removed it. The concept of ambiguity-reduction has been repeatedly mentioned in the scientific literature, but very few papers have actually addressed the problem of its mechanism. One of these papers was published with the name of *ribosome-oriented model* in order to underline the key role that the ribosomal proteins had in that process, but later on it became clear that other factors had to be taken into account. This is why the *ribosome-oriented model* had to be extended and here a more general version is proposed with the name of *ambiguity-reduction theory*.

1. Introduction

In *Chance and Necessity* (1971) Jacques Monod wrote that there are two alternative explanations for the genetic code. The first is chemical, or more precisely stereochemical: "… *if a certain codon was 'chosen' to represent a certain amino acid it is because there existed a certain stereochemical affinity between them*". The second is that "… *The code's structure is chemically arbitrary: the code as we know it today is the result of random choices which gradually enriched it*" (Monod, 1971, p. 135).

Monod had no hesitation in saying that the first hypothesis is far more appealing because "… it would explain the universality of the code and because it permits us to imagine a primitive translation mechanism based on direct interactions between codons and amino acids". He added, however, that "the numerous attempts to verify this hypothesis have up to now proved negative … Pending the unlikely confirmation of this first hypothesis we are reduced to the second one, displeasing from the methodological viewpoint because it does not explain the code universality, and because it does not provide any model of primitive translation" (Monod, 1971, p. 136).

Ten years later, in Life Itself (1981) Francis Crick wrote that "… the genetic code is as important for biology as Mendeleev's Periodic Table of the Elements is for chemistry, but there is an important difference. The Periodic Table would be the same everywhere in the universe. The genetic code appears rather arbitrary, or at least partly so … If this appearance of arbitrariness in the genetic code is sustained, we can only conclude that all life on earth arose from one very primitive population" (Crick, 1981, p 46–47).

The '*appearance of arbitrariness*' envisaged by Francis Crick became a certainty only a few years later, because it was shown that any codon can be associated with any amino acid (Schimmel, 1987; Hou and Schimmel, 1988; Schimmel et al., 1993) thus proving that there are no deterministic links between them. It is an experimental fact, in other words, that the genetic code is based on *arbitrary rules*, and the idea of descent from a common ancestor does explain its presence in all living organisms.

The first problem raised by Monod – how to explain the universality of the genetic code – does have a solution, but the second problem is still open: we still need to figure out how to "*provide a model of primitive translation*" with a code based on arbitrary rules.

An arbitrary genetic code is one where a codon can code for many amino acids, and this means that a sequence of codons is translated sometimes into one protein and some other time into a different protein. This amounts to saying that the first genetic codes were *ambiguous*, and that the primitive apparatuses of protein synthesis could only produce *statistical proteins*. It is a fact, on the other hand, that a fully *non-ambiguous* genetic code did appear on Earth and this means that the

E-mail address: brr@unife.it.

https://doi.org/10.1016/j.biosystems.2019.104024
Received 6 August 2019; Received in revised form 26 August 2019; Accepted 26 August 2019
Available online 06 September 2019
0303-2647/

ambiguity of the first codes was steadily reduced until it reached a point where any codon could code for one and only one amino acid[1] . That was the point when *biological specificity* came into existence: the apparatus of protein synthesis of the common ancestor ceased to make *statistical* proteins and started producing *specific* proteins. What we need to find out is the mechanism that brought that about.

2. Three groups of theories

The arbitrariness of the genetic code implies that the first code was ambiguous and its evolution was necessarily a process of ambiguity reduction (Woese, 1965; Fitch and Upper, 1987; Osawa, 1995). In such a situation, the theories that have been proposed so far can be divided into three great groups.

(1) The theories of the first group maintain that the genetic code is arbitrary *today*, but was not arbitrary *at the beginning*, and this means that we can continue to look for a stereochemical explanation of its origin. The first stereochemical theory was proposed by George Gamow (1954) with the idea that the amino acids fit with a lock-and-key mechanism into 'holes' formed by four nucleotides, and that it is the three-dimensional shape of each hole that determines which amino acid binds to which quartet of nucleotides. This model was quickly abandoned when it became clear that in protein synthesis there is no direct contact between codons and amino acids, but the *logic* of the stereochemical mechanism survived and has been re-proposed in many different forms (Dunnill, 1966; Melcher, 1974; Yarus, 1988,1998; Yarus et al., 2005). After decades of research, however, there still is no evidence in favour of the stereochemical theory, and what keeps it alive is the possibility that stereochemical interactions between codons and amino acids might have been important at some early stages of evolution (Koonin and Novozhilov, 2009).

(2) The theories of the second group accept that the first genetic codes *could* have been ambiguous but call attention to other properties of the coding system. A typical representative of this group is the *coevolution theory* proposed by Jeffrey Wong (1975, 1981). In this case the starting point was the fact that only about *half* of the 20 canonical amino acids are formed spontaneously in experiments that simulate the conditions of the primitive Earth. They are referred to as 'primary' (or 'precursor') amino acids, whereas the other canonical amino acids can only be produced by living systems and are referred to as 'secondary' (or 'product') amino acids. Wong proposed that the first genetic code contained only primary amino acids and all codons were assigned to their transfer-RNAs. Later on, when secondary amino acids were introduced in protein synthesis they too entered the genetic code and received their codons from the t-RNAs that transported their precursor amino acids. Wong did not say how the coding rules of the primary amino acids came into being, but this issue was taken on by Di Giulio (2008) who proposed an extension of the theory that applied the same mechanism to all codons and all amino acids. The key idea of the coevolution theory, in both the original and the extended form, is the hypothesis that in the early stages of evolution the synthesis of amino acids was taking place on transfer-RNAs, and there was therefore a *metabolic link* between them: a codon was associated to an amino acid because it belonged to the transfer-RNA that directed the synthesis of that amino acid. The coevolution theory is *compatible* with the existence of ambiguous codes in the ancestral systems, but does not address the evolutionary problem of the reduction of that original ambiguity.

(3) The theories of the third group acknowledge that the evolution of the genetic code was essentially a process of ambiguity reduction, and what distinguishes them are the mechanisms that they propose. The most common are feedback loops, self-organization and symmetry-breaking mechanisms, and in many cases these have been illustrated by computer simulations (Hoffmann, 1974; Bedian, 1982, 2001; Wills, 1993, 2001, 2009, 2016; Ardell and Sella, 2002; Delarue, 2007; Carter and Wills, 2018). Computer simulations are undoubtedly a powerful instrument, but the main problem is the underlining biological process, and in our case this process is protein synthesis. The genetic code is a component of the apparatus of protein synthesis, and the ancestral apparatuses could only produce *statistical proteins* so it is possible that the code evolved in order to improve the synthesis of *those* proteins. This idea was first proposed with the name of *ribosome-oriented model* (Barbieri, 2015) because it was argued that the ribosomal proteins had a key role, but other factors were also involved so let us see how that model can be generalized.

3. Two distinct mechanisms of code evolution

The translation of a sequence of nucleotides into a sequence of amino acids is subject to various types of errors that have been extensively studied in the laboratory. In particular it has been found that the error rate in the third position of a codon is about 100 times greater than that in the first position, which in turn is about 10 times greater than the error rate in the second position. The third position, in other words, is the most error-prone, whereas the second position is the most stable (Woese, 1965).

On the basis of these results Carl Woese proposed a theory on the evolution of the genetic code according to which the codons of the ancestral code were 'readjusted' in order to minimize the effects of the translation errors. More precisely, the ancestral code evolved in such a way that the codon resulting from a translation mistake would code either for the same amino acid or for an amino acid with very similar chemical properties, an idea that has become known as the '*error-minimization theory*'.

Woese pointed out that the ancestral apparatus of protein synthesis was bound to be far more rudimentary and prone to errors than the modern apparatus, and concluded that "… *errors in translation were extreme, to such an extent that the probability of translating correctly any given messenger-RNA was essentially zero. From this concept of error-ridden translation in the primitive cell it follows that the proteins produced by any given gene will have to be statistical proteins*" (Woese, 1965, p. 1548).

This is one of the first cases in which the term '*statistical proteins*' appeared in the scientific literature, and Woese underlined that such proteins were produced either because the ancestral translation apparatus was error-ridden, or because the ancestral code was ambiguous. Woese expressed this concept in no uncertain terms "… *the evolution of the genetic code starts with a primitive cell possessing random, ambiguous codon assignments, and a very error-ridden translation process, and it was the 'necessity' of minimizing the effects of translation errors that led to the highly ordered code that we observe today*" (Woese, 1965, p. 1550).

In his 1965 paper, furthermore, Woese insisted that the ancestral genetic code was necessarily *ambiguous*, and that the ancestral translation apparatus was necessarily *error-prone*, so there were two distinct types of statistical proteins in the ancestral systems: those produced by code ambiguity and those produced by translation errors. Some forty years later, Woese underlined again this difference and repeated that codon ambiguity has nothing to do with translation errors: "… *Ambiguity is therefore not the same thing as error*" (Vetsigian et al., 2006, p. 10696).

The existence of two different types of statistical proteins in the ancestral systems implies that the evolution of the genetic code took place by two distinct mechanisms, one that reduced code ambiguity and

[1] There are a few cases in which some ambiguity is present in the genetic code (Moura et al., 2009) but they are probably due to the '*rediscovery*' of ambiguity as a new mechanism of innovation, not to the survival of the ancestral ambiguity (that would have affected all living systems).

one that minimized the effects of translation errors. It follows that a comprehensive theory should take both mechanisms into account, but Woese concentrated exclusively on error-minimization and for a long time nobody proposed a mechanism of ambiguity-reduction.

One of the first models that described such a mechanism was published in 2015 with the name of *ribosome-oriented model* in order to underline the key role of the ribosomal proteins (Barbieri, 2015). The translation errors, on the other hand, cannot be ignored and we need therefore to keep in mind that *error-minimization* and *ambiguity-reduction* have been two distinct mechanisms and that both contributed to the evolution of the genetic code.

4. A primitive apparatus of protein synthesis

If the first genetic code was ambiguous and the ancestral systems were producing statistical proteins, as the evidence suggests, we need to address two major problems: (1) how could an arbitrary code appear in a primitive apparatus, and (2) how could a primitive apparatus work with statistical proteins?

(1) An ambiguous code in a primitive apparatus

The ribosomal RNAs are among the most conserved molecules in evolution (Woese, 1987, 2000) and contain regions that have the ability to form peptide bonds (Nitta et al., 1998). This suggests that the ribosomal-RNAs appeared very early on the primitive Earth and that some of them could create peptide bonds and had the *potential* to join amino acids together. To this purpose, however, the amino acids must be activated by the addition of ATP, and then the activated amino acids are transported by transfer-RNAs to the ribosomes. When the transfer-RNAs reach the site of protein synthesis, on the other hand, it is necessary that they are kept at a close distance for a long enough time to allow the formation of a peptide bond (Wolf and Koonin, 2007; Fox, 2010). This means that the transfer-RNAs must have temporary anchoring sites, and in primitive systems these were provided by *anchoring-RNAs*, the ancestors of the modern *messenger-RNAs*.

The combination of ribosomal-RNAs, transfer-RNAs and anchoring-RNAs gave origin to an apparatus of protein synthesis where the transfer-RNAs were automatically creating a *mapping* between codons and amino acids, and any such mapping is, by definition, a *genetic code*. This is how the genetic code came into being: that code appeared on Earth when ancestral transfer-RNAs and anchoring-RNAs joined the ancestral ribosomal-RNAs and became an integral part of the ancestral apparatus of protein synthesis.

But what type of code was it? All modern transfer-RNAs are small molecules (75–90 nucleotides long) with a basic cloverleaf structure that has been highly conserved in evolution, which strongly suggests that they descended from a common ancestor. This means that the primitive apparatus of protein synthesis was using transfer-RNAs that were still little diversified and could associate a codon with any number of amino acids. The ancestral genetic code, in other words, was *ambiguous* and it took a complex evolutionary process to get rid of its original ambiguity.

(2) A primitive apparatus with statistical proteins

It is known that the ribosomal proteins can vary from species to species, and this means that their functions can be performed by slightly different proteins. In bacteria, furthermore, the ribosomal proteins are fewer and smaller than in eukaryotes, which means that their number and their molecular weights are not crucial to their functions. The ribosomal proteins, in other words, can have different molecular weights, different numbers and different compositions and yet all can take part in fully operational translation machines. The ancestral ribosomes, in short, did not require specific proteins and could well have worked with *statistical* ribosomal proteins.

Another family of proteins that play a crucial role in protein synthesis is the *aminoacyl-tRNA-synthetases*, the molecules that activate the amino acids by attaching ATP to them, and then transport the activated amino acids to the transfer-RNAs. Today there are 20 different synthetases, one per amino acid, and all of them are highly specific proteins that could not have been present when the genetic code was ambiguous. It has been shown, however, that the synthetases belong to two distinct superfamilies which have completely different structures and yet are all capable to catalyze the same reaction (Rodin and Ohno, 1995; Woese et al., 2000). This implies that different molecules could have been able to perform similar synthetase functions and their ancestors could well have been a family of *statistical* synthetase proteins.

5. Evolving the ribosomes

The ribosomal RNAs appeared very early on the primitive Earth and together with statistical ribosomal proteins could start evolving the first machines of protein synthesis. Specific ribosomal proteins could not exist before the modern genetic code, but statistical ribosomal proteins could be manufactured, so what do we know about them?

A particularly illuminating information has come from the discovery that ribosomes are formed by the self-assembly of their components, and it has been possible to find out the contribution of individual ribosomal proteins by studying what happens when ribosomes are reassembled without anyone of them in turn. These experiments have shown that the ribosomal proteins fall into three major categories: some are necessary for function, others are required for self-assembly, and those of the third group have a stimulating effect but are fundamentally disposable (Kurland, 1970; Fox, 2010).

At first sight there does not seem to be a reason for the presence of disposable proteins, but in reality an explanation does exist. It comes from a general principle in engineering that Burks (1970) expressed in this way: *"there exists a direct correlation between the size of an automaton – as measured roughly by number of components – and the accuracy of its function"*. In our case, this principle means that there was an evolutionary advantage in increasing the number of ribosomal proteins because that was making the ribosomes more heavy, more resistant to thermal noise and therefore less prone to errors.

A similar principle accounts for the evolution of an increasing number of functional ribosomal proteins. Any complex system can improve its efficiency by increasing the number of controlling operations (Ashby, 1962), and it is probably for this reason that the number of ribosomal proteins with functional roles did increase in evolution. The same is true for the proteins involved in self-assembly: by increasing their number it was possible to produce ribosomes that could reassemble more easily and more efficiently from their components.

By increasing the number of the ribosomal proteins, in short, it became possible to reduce the translation errors and to improve the performance of the apparatus of protein synthesis, and this does explain why the number of those proteins did increase in evolution. In effect, the number of ribosomal proteins is 57 in Bacteria, 68 in Archaea and 78 in Eukaria, which clearly show there has been a tendency to increase their number (Lecompte et al., 2002). On the other hand, there are 34 ribosomal proteins which are universally conserved in all organism and they are probably the ribosomal proteins that evolved in the primitive systems before the common ancestor split into Bacteria, Archaea and Eukarya.

The increase in number of the ribosomal proteins, on the other hand, was accompanied by a parallel increase in size of the ribosomal RNAs, and the ancestral ribosomes steadily expanded their dimensions and eventually gave origin to enormous machines with molecular weights of over 2 million in prokaryotes and over 4 million in eukaryotes.

6. Evolving the adaptors

The evolution of the genetic code has been illustrated by Jacques Ninio (1982) with a beautiful metaphor. He pointed out that in any hotel, in addition to the familiar keys that open individual doors, there is a pass-key that opens all doors. At first, one may think that the pass-key is the most complex of all, but the truth is exactly the opposite. The pass-key is the simplest because what is complex in a key is not the ability to open a door but the ability to open one particular door *and not all the others*.

Ninio remarked that the transfer-RNAs of the modern genetic code can be compared to keys that open individual doors, whereas their common ancestor was like a pass-key that could open all doors. The evolution of the genetic code, in other words, can be described as a process of diversification of the transfer-RNAs that steadily increased their complexity by *decreasing* the number of amino acids that they could associate to each codon.

The amino acids are attached to the transfer-RNAs by synthetases (more precisely by *aminoacyl-tRNA synthetases*) that perform two distinct operations: on one side they recognize a specific amino acid, and on another side they recognize a specific structure of a transfer-RNA. The result is that each transfer-RNA gets attached to a specific amino acid because it contains a region that is recognized only by the synthetase that is carrying that amino acid. This means that the evolution of the genetic code consisted in two parallel evolutions: one that differentiated the transfer-RNAs by evolving individual features in each of them, and one that differentiated the synthetases in such a way that they could recognize those individual features.

The transfer-RNAs and the synthetases, in other words, evolved in parallel, very much like a set of locks that evolved in parallel with a set of keys until the point was reached in which any key could fit into one and only one lock. In many cases the features recognized by the synthetases are the anticodons of the transfer-RNAs, but in some cases the synthetates do not interact at all with the anticodons. Each synthetase, in other words, has evolved its own idiosyncratic way of recognizing a unique feature in a transfer-RNA, and this is why the synthetases are so different from one another (Schimmel et al., 1993; Carter and Wills, 2019).

The genetic code is a mapping between codons and amino acids that is realized by *adaptors* which are formed by transfer-RNAs and synthetases. The evolution of the genetic code was therefore the evolution of ancestral transfer-RNAs and ancestral synthetases that became increasingly *interdependent* and increasingly *selective* until the point was reached in which the anticodon of each transfer-RNA was associated with one and only one amino acid.

This is a description of what happened in the early history of life, but of course a description is not an explanation because it does not tell us *why* it happened. Why did the adaptors evolve the way they did? What were the *causes* that fuelled the evolution of the genetic code?

7. The mechanism of ambiguity-reduction

Before the origin of the modern genetic code the ancestral systems could only produce statistical proteins and yet life went on and evolved even in those times. There were two main reasons for this. The first is that the primary functions were performed by the RNAs and these molecules were faithfully transmitted from one generation to the next by molecular copying. The second is that the same protein functions could be implemented by slightly different molecules, and life could continue even if the proteins of the descendants were slightly different from those of the progenitors. More precisely, life could continue if the progenitors transmitted to the descendants the same RNAs and the same *families* of statistical proteins. There was however a condition that had to be met. The statistical proteins of a progenitor could reappear in a descendant only if the statistical differences between them were not cancelled out by the ambiguity of the genetic code.

The ancestral systems, in other words, could produce viable descendants only if the ambiguity of the genetic code was low enough to allow the same families of statistical proteins to reappear in each generation. This amounts to saying that the ambiguity of the genetic code could not exceed a prefixed limit, but within that limit the ancestral systems could go on indefinitely producing descendants that were statistically similar to the progenitors.

Evolution was bound to favour any improvement in the translation apparatus of the ancestral systems, and we have seen that the translation errors could be reduced by increasing the number of the ribosomal proteins. This increase, on the other hand, could be *perpetuated* only if a higher number of protein families could reappear in the descendants, and this was possible only if the ambiguity of the genetic code was reduced. The ambiguity of the code, in turn, could be reduced only by increasing the number and the diversity of the synthetases that were attaching amino acids to the transfer-RNAs.

An increase of the ribosomal proteins, in short, was favoured by evolution because it was reducing the translation errors, but could be achieved only by reducing the ambiguity of the genetic code, and this in turn could be achieved only by increasing the number of the synthetase proteins.

The evolution of the ribosomal proteins and the evolution of the synthetases, in other words, were two interdependent processes and both were favoured because the first was reducing the translation errors and the second was reducing the ambiguity of the genetic code.

The synthetases and the ribosomal proteins, in conclusion, evolved in parallel and the mechanism at the heart of their evolution was a systematic reduction in the ambiguity of the genetic code, a reduction that went on until any ambiguity was completely erased. At that point a sequence of codons was translated into one and only one protein and *biological specificity* came into existence.

The above scenario may look entirely speculative, at first, but in reality it does have consequences that can be tested. It implies, for example, that the universal ribosomal proteins and the synthetases were the first *specific* proteins that appeared in the history of life, and this is in agreement with the molecular phylogenies (Woese, 2000; Fox, 2010; Petrov et al., 2015). It also implies that the ribosomal proteins and the synthetases evolved in parallel and this too is a feature that could be documented by a comparative study of their phylogenetic records. Finally, we should keep in mind that the computer simulation studies are a powerful aid of investigation, and the mechanism of the ambiguity-reduction theory does provide the basis for a new computer simulation study of the genetic code.

8. The optimization phase

A number of computer simulation studies have shown that the modern genetic code performs better than most of its potential alternatives (Haig and Hurst, 1991; Freeland and Hurst, 1998; Bollenbach et al., 2007). There is however some disagreement about the characteristics that have been optimized.

According to Carl Woese, the genetic code has been optimized for minimizing the impact of the translation errors (Woese, 1965; Woese et al., 1966), whereas Gilis et al. (2001) have proposed that the modern code is optimal in respect to the stabilization of protein structure. Itzkovitz and Alon (2007) have argued that the modern code is nearly optimal for the acquisition of additional information into genetic sequences, whereas Drummond and Wilke (2008) have suggested that the modern code is ideally suited to favor the process of protein folding. On the whole, the computer simulations support the idea that the genetic code went through processes of optimization, and we cannot exclude the possibility that it was simultaneously optimized for a variety of different properties.

It must be underlined that some authors have warned against reaching overoptimistic conclusions on this issue. Novozhilov et al. (2007) have pointed out that there are 10^{84} possible codes and many of

them are more robust than the modern code. Their computer simulations have revealed that the genetic code did go through processes of optimization but apparently it went only half way up the optimality ladder.

Today there is a large consensus that the genetic code has been *at least partially* optimized and this has two outstanding theoretical implications.

The first is that there has been an period in the early history of life in which the rules of the genetic code went through a selection procedure that transformed the ancient genetic code into the modern one. In that historical period, in other words, *the coding rules were actually changed*, and this is an argument in favour of their arbitrariness because they could not have been tested and modified in a variety of different ways if they had been generated by deterministic processes.

The second implication concerns the regularities that have been discovered in the genetic code. More precisely, it has been found that in the table of the genetic code there are regularities both along the *columns* and along the *rows* of the table, but they seem to have different biological implications; it has been suggested that the regularities along the columns favour the stereochemical theory (Nelsestuen, 1978; Wolfenden et al., 1979; Sjostrom and Wold, 1985) whereas those along the rows seem in agreement with the coevolution theory (Taylor and Coates, 1989; Freeland et al., 2000; Di Giulio, 2017, 2018).

It has been assumed, in other words, that the regularities of the genetic code allow us to reconstruct the steps that gave origin to the code, but there is also a completely different explanation: they might have been produced in the later phase of code optimization and not in the earlier phase of code origin. Perhaps a familiar example can be useful here. The rules of the Morse code have been optimized by associating the most frequent letters of the alphabet with the shortest combinations of dots and dashes, but these regularities were introduced in a second phase of code refinement and tell us nothing about the first phase that gave origin to the Morse code.

9. The 'mega transition' of the genetic code

When the modern genetic code came into being, the ancestral systems acquired the ability of producing specific proteins. At that point they could have limited themselves to replace the previous statistical proteins with specific proteins, but in reality they did much more than that. They started using specific proteins for entirely new functions or for functions that were previously performed by the RNAs, and in this way initiated a revolution that eventually transformed the ancient RNA world into the modern protein world. Once in existence, in other words, the genetic code set in motion a massive exploration of the protein universe that changed the whole course of the history of life.

An outstanding example of that revolution comes from the machines of DNA replication. One can hardly imagine the ancestral systems without replicating DNA molecules, and yet most of the enzymes involved in DNA replication are radically different in bacteria and archaea, which means that the modern mechanisms of DNA replication have been re-invented at least twice by the descendants of the common ancestor (Woese, 2002; Koonin and Martin, 2005).

A second example of the protein revolution comes from *ATP synthase*, a molecular machine that represents one of the few universals of biology, like DNA, ribosomes and the genetic code. Virtually all cells obtain energy from proton currents associated with *ATP synthase*, a system made of some two dozen proteins organized into four functional elements (headpiece, rotor, shaft and stator) (Harold, 2014). These *specific* proteins could be synthesized only after the appearance of the modern genetic code, and yet even the previous ancestral systems needed energy and could only get it by similar energy-transduction means. This is why it has been proposed that the modern ATP synthases took the place of previous RNA machines that were operating in the ancestral world (Mulkidjanian et al., 2007).

The transition from the ancient RNA world to the modern protein world started with the appearance of the modern genetic code in the last universal common ancestor, and went on until the origin of the first modern cells, when life as we know it came into being.

The characteristics of the first modern cells have been reconstructed by molecular phylogenies and are traditionally represented by the three primary kingdoms that Carl Woese called Archaea, Bacteria and Eukarya (Woese, 1987, 2000). It must be underlined, however, that a major contribution to the origin of the first cells came also from the development of different types of cell membranes in the descendants of the common ancestor.

The cell membrane has a special role in life because it is never constructed *de novo*. Membranes always grow from pre-existing membranes, and this has lead to the concept of *membrane heredity*, the idea that membranes are passed down from one generation to the next in an uninterrupted chain of descent (Blobel, 1980; Sapp, 1987; Cavalier-Smith, 2000; Wächterhäuser, 2003; Harold, 2005). Chromosomes are reproduced from pre-existing chromosomes and membranes from pre-existing membranes but they carry two very different types of instructions. Chromosomes transmit genetic information, whereas membranes transmit architectural order.

We realize in this way that the mega transition of the protein revolution went on in parallel with the membrane revolution, and both took place in the descendants of the common ancestor after the origin of the modern genetic code.

10. The conservation imperative

The genetic code is an integral part of the apparatus of protein synthesis, and yet there is a profound difference between the evolution of that code and the evolution of that apparatus. The genetic code has been conserved since its origin, almost 4 billion years ago, whereas the apparatus of protein synthesis has continued to change. In prokaryotes, for example, the ribosomes have molecular weights of about 2 million whereas in eukaryotes their weights exceed 4 millions. In eukaryotes, the ribosomal RNAs are much heavier than in prokaryotes, and the biogenesis of the ribosomes is much more complex.

In fact, virtually everything has been modified in the apparatus of protein synthesis in the course of evolution and the sole outstanding exception is the rules of the genetic code. They are the sole entities that have been conserved for billions of years while everything else has changed.

This extraordinary process of conservation is usually accounted for by saying that the genetic code is a set of *constraints* (Pattee, 2001; Gould, 2002) and that constraints cannot be changed, an idea that appears to explain why the genetic code has been *frozen* since the origin of life.

The statement that the genetic code is a set of constraints is *formally* correct because its rules impose severe limitations on a virtually unlimited number of possibilities. It must be underlined, however, that they are not *physical* constraints. A big piece of rock on a road is a physical constraint, but a traffic light is not; a traffic light is a code, a totally different type of constraints.

The rules of the genetic code are *biologically generated constraints* that in no way can be assimilated to physical constraints because the genes of the molecules that implement the code are constantly subject, like all other genes, to mutation and neutral drift. They are in a continuous state of flux and the fact that they have been conserved in evolution means that there is a biological mechanism that *actively and continuously* restores their original structure.

The conservation of the genetic code, in other words, is not the result of physical constraints but of active biological mechanisms that are continuously at work. It is tempting to say that these mechanisms consists in the standard processes of gene replication and gene repair, but these processes work in the same way on all genes and do not explain why some genes are more conserved than others.

From a theoretical point of view, the conservation of the genetic

code has been regarded as the result of *autopoiesis*, the process by which living systems continuously fabricate their own components and conserve them in time (Maturana and Varela, 1980). Autopoiesis, however, implies that all components of the living systems are equally conserved, whereas the reality is that in the long run most of them have been changed and only a few have been conserved. More precisely, it is the coding rules that have been highly conserved in evolution, so what we have is not autopoiesis but *codepoiesis* (Barbieri, 2012).

The conservation of the genetic code, in conclusion, is an experimental reality but it is also a major theoretical problem that has not yet been solved.

11. Conclusions

The traditional theories on the genetic code are based on what we actually know about its properties, but there is another approach that should be considered. Rather than studying the genetic code on its own, as an isolated case, we could study it together with the many other biological codes that have appeared in the history of life (Barbieri, 2003, 2018).

This approach is possible because all biological codes have one thing in common: they are all based on arbitrary rules that create mappings between independent worlds and this allowed them to give origin to absolute novelties, including the novelties that we associate with the major transitions in evolution. Among the major transitions, furthermore, three have a special place because they gave origin to entirely new worlds and for this reason can be referred to as "*mega transitions*". The three mega transitions were the origin of life, the origin of mind and the origin of language and their codes were respectively the genetic code, the neural code and the language code.

If we study the genetic code in this larger framework, we realize that the theory of ambiguity-reduction is applicable to many other cases. Any other arbitrary code was necessarily ambiguous at the beginning and this means that its evolution was necessarily a process of ambiguity reduction. Another characteristic of the genetic code that can be found in many other codes is the existence of five distinct phases in their history (Barbieri, 2019).

The first phase is the appearance of an ambiguous code in a biological system; the second is the phase in which the ambiguity of the code is systematically reduced until it is completely eliminated. The third phase is the optimization of the coding rules. The fourth phase is the major transition set in motion by the code and the fifth is the conservation phase, the development that put an end to any other change in the code.

It has been shown that these five phases can also be recognized in the history of the neural code (Barbieri, 2019) and this opens the doors to a completely new approach to evolution.

Let us summarize: the genetic code evolved as a means of solving a local problem – how to improve the synthesis of statistical proteins – but then it became the tool of a much larger change that transformed the ancient RNA world into the modern protein world. The key point is that this was not an isolated case because there have been many other similar episodes in the history of life: other biological codes started as a means of solving local problems and then became the tools that set in motion other great events of macroevolution.

The mechanism of ambiguity reduction, in conclusion, is not limited to the genetic code. It is a mechanism that we find in various other codes and may well represent a general mechanism of evolution.

Declaration of Competing Interest

There is no conflict of interest.

Acknowledgments

I am deeply grateful to Jannie Hofmeyr and to Peter Wills because their comments induced me to rethink a previous model on the evolution of the genetic code and to start the changes that eventually transformed that model into the more general theory proposed in this article.

References

Ardell, D.H., Sella, G., 2002. No accident: genetic codes freeze in error-correcting patterns of the standard genetic code. Phil. Trans. R. Soc. Lond. B 357, 1625–1642.
Ashby, W.R., 1962. Principles of the self-organizing system. In: Von Foerster, H., Zopf Jr.G.W. (Eds.), Principles of Self-Organization: Transactions of the University of Illinois Symposium. Pergamon Press, London, UK, pp. 255–278.
Barbieri, M., 2003. The Organic Codes. An Introduction to Semantic Biology. Cambridge University Press, Cambridge, UK.
Barbieri, M., 2012. Codepoiesis – the deep logic of life. Biosemiotics 5 (3), 297–299.
Barbieri, M., 2015. Evolution of the genetic code: the ribosome-oriented model. Biol. Theory 10 (4), 301–310.
Barbieri, M., 2018. What is code biology? BioSystems 164, 1–10.
Barbieri, M., 2019. A general model on the origin of biological codes. BioSystems 181, 11–19.
Bedian, V., 1982. The possible role of assignment catalysts in the origin of the genetic code. Orig. Life 12, 181–204.
Bedian, V., 2001. Self-description and the origin of the code. BioSystems 60, 39–47.
Blobel, G., 1980. Intracellular membrane topogenesis. Proc. Natl. Acad. Sci. U. S. A. 77, 1496–1500.
Bollenbach, T., Kalin Vetsigian, K., Kishony, R., 2007. Evolution and multilevel optimization of the genetic code. Genome Res. 17, 401–404.
Burks, A.W., 1970. Essays on Cellular Automata. University of Illinois Press, Urbana.
Carter, C.W., Wills, P.R., 2018. Interdependence, reflexivity, fidelity, and impedance matching, and the evolution of genetic coding. Mol. Biol. Evol. 35, 269–286.
Carter, C.W., Wills, P.R., 2019. Class I and II aminoacyl-tRNA synthetase tRNA groove discrimination created the first synthetase*tRNA cognate pairs and was therefore essential to the origin of genetic coding. IUBMB Life 71 (8), 1088–1098.
Cavalier-Smith, T., 2000. Membrane heredity and early chloroplast evolution. Trends Plant Sci. 5, 174–182.
Crick, 1981. Life Itself. Simon and Schuster, New York.
Delarue, M., 2007. An asymmetric underlying rule in the assignment of codons: Possible clue to a quick early evolution of the genetic code via successive binary choices. RNA 13, 1–9.
Di Giulio, M., 2008. An extension of the coevolution theory of the origin of the genetic code. Biol. Direct 3, 1–37.
Di Giulio, M., 2017. The aminoacyl-tRNA synthetases had only a marginal role in the origin of the organization of the genetic code: evidence in favor of the coevolution theory. J. Theor. Biol. 432, 14–24.
Di Giulio, M., 2018. A discriminative test among the different theories proposed to explain the origin of the genetic code: the coevolution theory finds additional support. Biosystems 169, 1–4.
Drummond, D.A., Wilke, C.O., 2008. Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. Cell 134, 341–352.
Dunnill, P., 1966. Triplet nucleotide-amino-acid pairing; a stereochemical basis for the division between protein and non-protein amino-acids. Nature 210, 1267–1268.
Fitch, W.M., Upper, K., 1987. The phylogeny of tRNA sequences provides evidence for ambiguity reduction in the origin of the genetic code. Cold Spring Harbor Symp. Quant. Biol 52, 759–767.
Fox, G.E., 2010. Origin and evolution of the ribosome. Cold Spring Harb. Perspect. Biol. 2, a003483.
Freeland, S.J., Hurst, L.D., 1998. The genetic code is one in a million. J. Mol. Evol. 47, 238–248.
Freeland, S.J., Knight, R.D., Landweber, L.F., Hurst, L.D., 2000. Early fixation of an optimal genetic code. Mol. Biol. Evol. 17, 511–518.
Gamow, G., 1954. Possible relation between deoxyribonucleic acid and protein structures. Nature 173, 318.
Gilis, D., Massar, S., Cerf, N.J., Rooman, M., 2001. Optimality of the genetic code with respect to protein stability and amino-acid frequencies. Genome Biol. 2, 41–49.
Gould, S.J., 2002. The Structure of Evolutionary Theory. Harvard University Press, Cambridge MA.
Haig, D., Hurst, L.D., 1991. A quantitative measure of error minimization in the genetic code. J. Mol. Evol. 33, 412–417.
Harold, F.M., 2005. Molecules into cells: specifying spatial architecture. Microbiol. Mol. Biol. Rev. 69, 544–564.
Harold, F.M., 2014. In Search of Cell History. The Evolution of Life's Building Blocks. The University of Chicago Press, Chicago and London.
Hoffmann, G.W., 1974. On the origin of the genetic code and the stability of the translation apparatus. J. Mol. Biol. 86, 349–362.
Hou, Y.-M., Schimmel, P., 1988. A simple structural feature is a major determinant of the identity of a transfer RNA. Nature 333, 140–145.
Itzkovitz, S., Alon, U., 2007. The genetic code is nearly optimal for allowing additional information within protein-coding sequences. Genome Res. 17, 405–412.
Koonin, E.V., Martin, W., 2005. On the origin of genomes and cells within inorganic compartments. Trends Genet. 21, 647–654.
Koonin, E.V., Novozhilov, A.S., 2009. Origin and evolution of the genetic code: the universal enigma. IUBMB Life 61 (2), 99–111.
Kurland, C.G., 1970. Ribosome structure and function emergent. Science 169, 1171–1177.

Lecompte, O., Ripp, R., Thierry, J.C., Moras, D., Poch, O., 2002. Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale. Nucleic Acid Res. 30 (24), 5382–5390.
Maturana, H.R., Varela, F.J., 1980. Autopoiesis and Cognition: The Realisation of the Living. D. Reidel Publishing Company, Dordrecht, Holland.
Melcher, G., 1974. Stereospecificity and the genetic code. J. Mol. Evol. 3, 121–141.
Monod, J., 1971. In: Monod, J. (Ed.), Chance and Necessity. Alfred Knopf, New York, original edition. Le Hasard et la Nécessité. éditions du Seuil, Paris 1970.
Moura, G.R., Carreto, L.C., Santos, M.A., 2009. Genetic code ambiguity: an unexpected source of proteome innovation and phenotypic diversity. Curr. Opin. Microbiol. 12, 1–7.
Mulkidjanian, A.Y., Makarova, K.S., Galperin, M.Y., Koonin, E.V., 2007. Inventing the dynamo machine: the evolution of the F-type and V-type ATPases. Nat. Rev. Microbiol. 5, 892–899.
Nelsestuen, G.L., 1978. Amino acid-directed nucleic acid synthesis. A possible mechanism in the origin of life. J. Mol. Evol. 11, 109–120.
Ninio, J., 1982. Molecular Approaches to Evolution. Pitman books, London.
Nitta, I., Kamada, Y., Noda, H., Ueda, T., Watanabe, K., 1998. Reconstitution of peptide bond formation. Science 281, 666–669.
Novozhilov, A.S., Wolf, Y.I., Koonin, E.V., 2007. Evolution of the genetic code: partial optimization of a random code for robustness to translation error in a rugged fitness landscape. Biol. Direct 2 (24).
Osawa, S., 1995. Evolution of the Genetic Code. Oxford University Press, New York.
Pattee, H.H., 2001. The physics of symbols: bridging the epistemic cut. BioSystems 60, 5–21.
Petrov, A.S., et al., 2015. History of the ribosome and the origin of translation. Proc. Natl. Acad. Sci. U. S. A. 112, 15396–15401 doi:10:1073/pnas.159761112.
Rodin, S.N., Ohno, S., 1995. Two types of aminoacyl-tRNA synthetases could be originally encoded by complementary strands of the same nucleic acid. Orig. Life Evol. Biosph. 25 (6), 565–589.
Sapp, J., 1987. Beyond the Gene: Cytoplasmic Inheritance and the Struggle for Authority in Genetics. Oxford University Press, Oxford UK.
Schimmel, P., 1987. Aminoacyl tRNA synthetases: general scheme of structure-function relationship in the polypeptides and recognition of tRNAs. Ann. Rev. Biochem. 56, 125–158.
Schimmel, P., Giegé, R., Moras, D., Yokoyama, S., 1993. An operational RNA code for amino acids and possible relationship to genetic code. Proc. Natl. Acad. Sci. U. S. A. 90, 8763–8768.
Sjostrom, M., Wold, S.A., 1985. A multivariate study of the relationship between the genetic code and the physico-chemical properties of amino acids. J. Mol. Evol. 22, 272–277.
Taylor, F.J.R., Coates, D., 1989. The code within the codons. BioSystems 22, 177–187.
Vetsigian, K., Woese, C., Goldenfeld, N., 2006. Collective evolution and the genetic code. Proc. Natl. Acad. Sci. U. S. A. 103 (28), 10696–10701.
Wächterhäuser, G., 2003. MicroHypothesis: from pre-cells to Eukarya - a tale of two lipids. Mol. Microbiol. 47 (1), 13–22.
Wills, P.R., 1993. Self-organization of genetic coding. J. Theor. Biol. 162, 267–287.
Wills, P.R., 2001. Autocatalysis, information and coding. BioSystems 60, 49–57.
Wills, P.R., 2009. Informed generation: physical origin and biological evolution of genetic codescript interpreters. J. Theor. Biol. 257, 345–358.
Wills, P.R., 2016. The generation of meaningful information in molecular systems. Phil. Trans. R. Soc. A A374 20150016.
Woese, C.R., 1965. Order in the genetic code. Proc. Natl. Acad. Sci. U. S. A. 54, 71–75.
Woese, C.R., 1987. Bacterial evolution. Microbiol. Mol. Biol. Rev. 51, 221–271.
Woese, C.R., 2000. Interpreting the universal phylogenetic tree. Proc. Natl. Acad. Sci. U. S. A. 97, 8392–8396.
Woese, C.R., 2002. On the evolution of cells. Proc. Natl. Acad. Sci. U. S. A. 99, 8742–8747.
Woese, C.R., Dugre, D.H., Saxinger, W.C., Dugre, S.A., 1966. The molecular basis for the genetic code. Proc. Natl. Acad. Sci. U. S. A. 55, 966–974.
Woese, C.R., Olsen, G.J., Ibba, M., Söll, D., 2000. Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. Microbiol. Mol. Biol. Rev. 64 (1), 202–236. https://doi.org/10.1128/MMBR.64.1.202-236.2000.
Wolf, Y.I., Koonin, E.V., 2007. On the origin of the translation system and the genetic code in the RNA world by means of natural selection, exaptation, and subfunctionalization. Biol. Direct 2, 14.
Wolfenden, R.V., Cullis, P.M., Southgate, C.C.F., 1979. Water, protein folding, and the genetic code. Science 206, 575–577.
Wong, J.T., 1975. A co-evolution theory of the genetic code. Proc. Natl. Acad. Sci. U. S. A. 72 (5), 1909–1912.
Wong, J.T., 1981. Coevolution of genetic code and amino acid biosynthesis. Trends Biochem. Sci. 6, 33–36.
Yarus, M., 1988. A specific amino acid binding site composed of RNA. Science 240, 1751–1758.
Yarus, M., 1998. Amino acids as RNA ligands: a direct-RNA-template theory for the code's origin. J. Mol. Evol. 47 (1), 109–117.
Yarus, M., Caporaso, J.G., Knight, R., 2005. Origins of the genetic code: the escaped Triplet theory. Annu. Rev. Biochem. 74, 179–198.