Research and Reports in Biology

Dovepress
open access to scientific and medical research

Open Access Full Text Article

HYPOTHESIS

Biological genesis: the first step from dead matter to life. A contribution to the nature of DNA, RNA, and the genetic code

This article was published in the following Dove Press journal:
Research and Reports in Biology
3 April 2013
Number of times this article has been viewed

Friedrich H Schmidt

Retired, Schramberg, Germany

Abstract: Information is understood semantically in the special case of the genetic code as the contents of news-bearing and genetically acting molecules. The connection of single molecules to groups and molecule chains can be referred to as syntactic. Well-defined information is not only exchanged between molecules in biology like nucleic and amino acids cooperating in the genetic code: the topic of this article is that an exchange of information could also occur between inorganic and organic substances, eg, mineral crystals interacting with organic molecules. This may have played a role in the origins of life on earth. As the origin of the genetic code and the mechanism of its translation is still an unresolved problem, so is the interaction of inorganic substances and organic substances still an open question. Stereochemical similarities existing between code and amino acids cannot explain the relationship completely and are not present between inorganic and organic molecules at all. Symmetry is a structural entity in organic chemistry and organisms, and Δ-values calculated by a mathematical algorithm and introduced in this article give an estimate of symmetry and transferred information. Symmetric Δ-values exist in minerals as well as in genetic molecules, and could thus bring dead material to life before DNA, RNA, and enzymes were developed. The fact that symmetry is important as a quality of organic matter with the function of the genetic code is pointed out in the works of other authors, who are cited in this paper.

Keywords: genetic information, genetic code, symmetry in inorganic and organic molecules, calculation of Δ-values

Introduction

> The Greeks interpreted symmetry as the harmony of different parts of an object, the good proportions of its constituents, a symbol of seeking for perfection.[1]

The author aims to find common symmetry patterns between even very differently constructed molecules, reflecting similarities or equalities of body plan and function with respect to the source of genetic information and its transport and genetic coding.

> Symmetry is defined mathematically by the maxims of geometry and set theory.[1]
>
> The perfect code theory […] believes that the symmetry of genomes and the symmetry of their codes of translation [...] represents a key component of quantifying molecular information.[2]
>
> A concept of macrobiocrystalloid with inherent complex symmetry is proposed for the description of the human organism in its integrity […] the symmetrical pattern of plants, animals and crystals follow similar mathematical solutions.[3]

Correspondence: Friedrich H Schmidt
Langenberg 72, 78144 Schramberg, Germany
Tel +49 7729 8247
Email fhsdelta@aol.com

> The symmetries of crystals are decided by the electronic properties of their constituent atoms. Minerals with quite different chemical composition display the same symmetry and the proteins, consisting of different amino acid sequences, result in the same structural pattern and the same function.[4]

Biogenesis

The origin of information and its continuation is, from the syntactic view, ensured by symbols (bits, digits, letters, signs), semantically bound to its assessment. In biology, information must be understood generally as woven into the existence and heredity of all organisms. Genetic information hereby must be regarded as invariant and materially determined but subject to a dynamic process. Qualitative means of development are provided by influences of stereochemistry, conformation of biological molecules, and epigenetic aspects.

The use and classification of symbols is characterized overall by their amount and distribution. The symbols must not be ambiguous in their representation, and there must be, in this instance, an unequivocal correlation between them and the elementary signs that they code or signal.

The genetic code originated either during the early years of the earth's development, or later in an even more unsettled manner, from an inherent matter quality, or in a kind of arbitrary arrangement between encoding and decoding molecules.

The year 2012 represents the 50th year since the awarding of the Nobel Prize in Medicine for the discovery of the double-helix structure of DNA, and the function of RNA, and the engaged-transcription processes of the origin of proteins by Watson and Crick.[5] Before then, Gamow and Ycas[6] showed that a triplet of RNA nucleotides encodes an amino acid.

Originally, it was thought that the origin of life was seemingly exclusively bound to the DNA. Today, however, it is known that organized DNA did not exist in the form of genes at the beginning of life.[7]

Information held in the nucleic acids and amino acids, however, contains an independent architectural plan and is identical to the sequence in which the single building blocks are arranged.

The helical structures of the molecules constituting the genetic coding sequence also have another function: they form a text, which can serve to produce a mirror-like copy. This reflection is based on the fact that base pairs can complement each other to form a pair: adenine with thymine (or uracil in RNA) and guanine with cytosine. With the help of enzymes, a complementary molecule can thus be formed from a strand of nucleic acids. Moreover, two complementary molecule segments recognize themselves in the cell system and can functionally cooperate.

The passing on of information and its storage is based on the reflection principle between DNA and DNA, or with the interactions between DNA, RNA, and amino acids, in the forward or reverse direction.

The origin of the first life on earth is based on the cooperation of two kinds of molecules – RNA and proteins. Both molecule types are made up of linearly aligned single blocks: RNA of four and proteins of 20 different types of molecules. Such molecule rows can produce themselves mutually, transforming and storing architectural plans on the basis of information exchange. The key to facilitating the protein's role as a building material for organisms is to specify the sequence of amino acids in the protein through the RNA template, derived from DNA codes, in a process known as transcription.

The simplest known forms of life already assume the concurrent coexistence of molecules that must fit well enough together, so that they can cooperate chemically with each other.[8]

Life is not primarily produced by a specific type of molecule, but by information stored in molecules that drives the production of other molecules. Insofar as life depends on DNA, this means that the use of genetic information stored in this molecule is of determining importance.

It is the huge capacity of information made possible by biological molecules that strains our imagination, because it makes a chance production so unlikely. Even if all necessary substances in the right amount are given in an in vitro system, somewhat like a "soup of Miller," the enzymatic process "must form in one-thousandth of a second" and all functioning proteins must be "situated in the right position with functional building blocks."[9]

If it were only left to chance, with a primordial soup carrying an immense number of existing molecule classes that could combine over and over again to form a living substance, "life would have had no chance to develop."[9] Neither the available time nor the available material in the whole universe would have been sufficient to bring about the completely unlikely rare and extremely short event that through testing and recycling created a divine spark at one given single moment, which gave way to life that would last for eternity.

Some scientists believe that the laws of physics and chemistry would have to be revised before the query of life

could be solved. The biologist does not need to deal with the chemical construction and the spatial structure of the macromolecules. For him/her, just the almost incalculable complexity and variety of life is its essential property. By the sequence analytics of the latest molecular genetics, we know that even very differently appearing living beings, like humans and bacteria, have basic qualities in common, like the genetic code and a relatively large part of the genome. On the other hand, the extraordinarily divergent peculiarity of living beings is not to be dismissed. Despite the desire to understand life as a special but generally valid organized form of matter, both phenomena cannot be explained together so simply.

Today's biology denies the basic question of whether there is a clear demarcation of the animated and the inanimate. Indeed, the crossing is blurred. A virus owns nucleic acids and a protein coat; however, it must be considered only as a conditional living nanoorganism, as it needs to be hosted to develop itself (autoreplication). In fact, viruses are commonly considered as opportunist organisms. It should be noted that viruses are only a particle of a given composition and shape that can be transformed analogously to mineral materials in a crystal structure. In this form, they are not regarded as living.

Moreover, viruses can be disassembled and composed again without losing their peculiarity and virulence. Thus, life and death must be understood differently from their traditional definitions in this circumstance.

The manner in which inanimate entities became animated also remains unclear. The production of life-forms is bound to the availability of entire cells in which, and above all, not only a complete genetic apparatus but also a smoothly functioning enzyme system exists. Organisms can reproduce themselves completely.

The first cells that developed (prokaryotes) were seedless; however, they already possessed various special components, among them ribosomes. The next-generated core-containing cells (eukaryotes) dispose of regular life functions, in particular with regard to storage of information and inheritance and cell organelles. According to a theory, eukaryotes incorporated prokaryotes (the "endosymbiotic" theory of Mereschkowski[10] and Margulis[11]) and used their components as cell organelles.

The big outstanding question in biology is whether the cells needed a directing effect and to receive information by means of external forces, or if the ability to grow is innate. The building of a codon out of three nucleotides to encode an amino acid provides a universal arrangement in the whole animated world. This gives rise to the formation of secondary and tertiary structures by drilling and interlinking of linearly arranged information and function-bearing molecules. All of this works in the same manner over and over again as a parallel running production of tissues and organs in individuals who retain recurring properties and inheritance mechanisms, in addition to differentiation. That the existence of all of this has occurred by chance seems incredible. One must assume a structure that already exists in the inanimate, where the origin of life takes over to consolidate a process.

In addition to our knowledge of the genetic triplet as a result of experimental investigations Eigen and Winkler proved that the genetic code is not based on chance. They realized from kinetic investigations in different nucleotide combinations of defined composition that the genetic code is an optimum solution of the nature above all "for reasons of the dynamic stability and flexibility."[9]

If one interprets the above statements properly, one must suppose that nature has no other possibility than the genetic code to store and transmit information because of its structural, functional, and stable properties for the purpose of producing organic life.

The question of inanimate/animate remains, as does the question of general validity/individual specific feature, which is associated with the investigation of the genetic code. The syntactic arrangement and the semantic meaning of genes are based on the universal use of the genetic code everywhere in biology. The laws of physics and chemistry are equally as universal. However, biology is aimed at the individual appearance of individuals. Without this, biology would be a tasteless imitation of chemistry and physics. It does not seem that individuality can be explained by physical and chemical laws. In general, the linear strand of the peptides is considered universally valid, while the more perpendicular arrangement of the side chains of the amino acids is what determines the individuality, form, and function. Information is attributed by assigned syntax and semantics. Absolute information is characterized by the amount of symbols and their distribution.[12] The symbols must be free of any other meaning and be transferable by physical signals. In addition, there must be an unequivocal code, so an unequivocal correlation must exist between the number of symbols and their assigned meaning.

It has become accepted that the transmission of information is based on physical strength effects between nucleotide sequences. The absolute information content of a given number of symbols can be expressed quantitatively by the likelihood distribution for the totality of all possible

arrangements. One labels the number of the elementary signs of a given amount with ν, the number of the classes of signs with Λ. Then the totality of the possible arrangements is denoted by Λ^{ν}.

It is well known, due to the pioneering work of Woese,[13] that each of the four naturally occurring bases of the biological code act as coding substances by the substitutes it carries on the 2 and 6 positions in the ring, adenine (A) being 2-H, 6-NH_2; cytosine (C) being 2-OH, 6-NH_2; guanine (G) 2-NH_2, 6-OH; and uracil (U) (and thymine [T] in DNA) 2-OH, 6-OH. It appears probable that coding will be performed exclusively in terms of the 2 and 6 positions. It means that the attention on the carried genetic code information should be confined to the 2 and 6 positions rather than to the four bases as a whole.

If all possible sequences of groups were equally encoded, we could assign each the same a priori likelihood, and then this number could be used to characterize the information content. If not, ie, if one deals with different sequence likelihoods, then each segment with the same sequence must be calculated separately and summarized to be able to meet an unequivocal allocation from a dissimilar outgoing likelihood distribution.

Likelihood has the quality to be multiplicative. That is, the likelihood for the concurrent presence of several independent (characterized by the likelihood of $p_1, p_2 \ldots p_k$ single events) is the product $p_1 \times p_2 \times \ldots p_k$.

On the other hand, the information content should have additive qualities as an amount measure. The easiest mathematical function that fulfills this condition is the logarithm:

$$\mathrm{ld}\,(a \times b) = \mathrm{ld}\,a + \mathrm{ld}\,b \tag{1}$$

For reasons of usefulness, and because of mathematically specific features in physics, which are not discussed within the scope of this article, one chooses the logarithm with base 2, which is referred to here as ld. Then, in the simplest case of the same, a priori likelihood of all orders:

$$H = \mathrm{ld}\,(\Lambda^{\nu}) = \nu\,\mathrm{ld}\,\Lambda. \tag{2}$$

In the case of a mixture of different sequences the information content is calculated as follows:[14]

$$H = -\sum_{i=1}^{n} p_i\,\mathrm{ld}\,p_i \text{ with } \sum_{i=1}^{n} p_i = 1, \tag{3}$$

considering ν as the total number of all possible and Λ^{ν} as the total numbers of arrangements. The p_i are concurrently the weight factors of the single elements.

H (designated by Shannon and Weaver[14] as entropy) represents an average value of all information ld p_i. If these are all even, then H gets the simple expression $H = -\mathrm{ld}\,p$.

H would be zero if all the arrangements except one possess the likelihood zero ($H = -1 + [n - 1]0 + 1 = 0$). In the case of one circumstance having the likelihood of one, this one would not need an alternative, because it would not need an information amount.

An example of this would be a molecule structure that appears from a stochiometric mixture of the elements on the grounds of specific interchange efficacies. The spatial arrangement of the nucleic and polypeptide chains and their folding can be stated as an additional example. It is unambiguously determined by the sequence of the molecules; "no change of information originates from the secondary structural processes."[4]

Divergences in the same distribution of the a priori likelihood of genetically active molecules, irrespective of the genetically active groups in them as they cooperate in the genetic code, are designated as redundant. By them the so-called Δ-values shall be calculated.

Divergences of the same distribution of a priori likelihood do not originate from a different frequency of symbolic classes, but from a different number of elementary signs of a given amount in the course of sequences.

Essential knowledge of molecular biological research is the "Chargaff rule"[15] from the pairing of the bases, the cooperation of the nucleotides, and finally in the formulation of the pairing of nucleic acid chains into the double helix and the production of proteins through interpretation of the genetic code by Watson and Crick.[5] This rule states that the base adenine is always paired with thymine in the DNA (and with uracil in the RNA) and guanine with cytosine for both nucleic acids in a complementary manner. In this way, the production of double-stranded chains can proceed.

The Chargaff[15] rule, and the use of the base side groups, presumably as the actual encoding units, should be understood in chemical detail if one wants to entirely understand the molecular biological events.

If the information content from the amount and arrangement of the side groups of the nucleic and amino acids is to be calculated, then the base atoms must be considered, on the grounds that by them the values of the side groups could be different. In addition, the position and the sequence of a side group in the respective molecule conformation must also be considered.

If one wants to calculate the information content of a side group, all the atoms within a molecule have to be

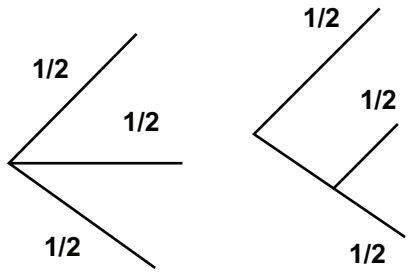

Figure 1 Diversity in logic.

considered, and in different kind according to their distance to the requested side group after division of a molecule, like the tree in informatics.

The procedure of such a classification by dichotomy returns to Plato. In addition to the binary logic, such a graphic tree can possess one or a multitude of logical values. These circumstances are illustrated in Figure 1. Both results possess the same likelihood.

A similar process occurs with the calculation of Δ-values in relation to the minimum model of the theory of combinations as illustrated in Figure 2 and Table 1.

It is easily recognized that, in a symmetrically and harmoniously arranged molecule constellation as in Figure 2, the sum of all information that possesses single elements according to Figure 3 (atoms) is essentially zero. The substitution of additional side groups would change this sum and lead to circumstances of dissonant, asymmetrical molecules.

The use of a multitude valued logic is of interest in the calculation of information in multiple atomic molecules. Individual values should be referred to as Δ-values, which are an indicator of the chemical–biological information contained in a molecule, for symmetries and reactive forces.

To the atomic weight of a given atom, the value +1, to its next neighbor −1/2, the very next +1/4, the following −1/8, and so on shall be assigned. The total number of the values of the atoms of a class will be summarized. In addition, the atomic weight will be used as its dual logarithm.

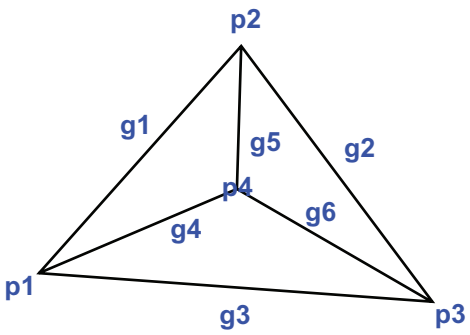

Figure 2 Minimum model of combinations.

Table 1 Elements p_n/quality of relations g_n (+ and −)

	g1	g2	g3	g4	g5	g6
p1	+	−	+	+	−	−
p2	+	+	−	−	+	−
p3	−	+	+	−	−	+
p4	−	−	−	+	+	+

The Δ-values will be calculated according to the following algorithm:

$$H = (-1)^n \sum_{i=1}^{n} 2^{-i} \operatorname{ld} AW, \tag{4}$$

where AW is atomic weight.

As an example, the Δ-value of the atom N1′ shall be calculated:

$$\begin{aligned} \Delta N &= 1 + \tfrac{1}{4} + \tfrac{1}{16} = \tfrac{21}{16} \operatorname{ld} AW(N) = 5.00 \\ \Delta C &= -\tfrac{1}{2} + \tfrac{1}{4} - \tfrac{2}{8} = -\tfrac{1}{2} \operatorname{ld} AW(C) = -1.79 \\ \Delta O &= \tfrac{1}{16} \operatorname{ld} AW(O) = 0.25 \\ \sum \Delta N1' &= 3.46. \end{aligned} \tag{5}$$

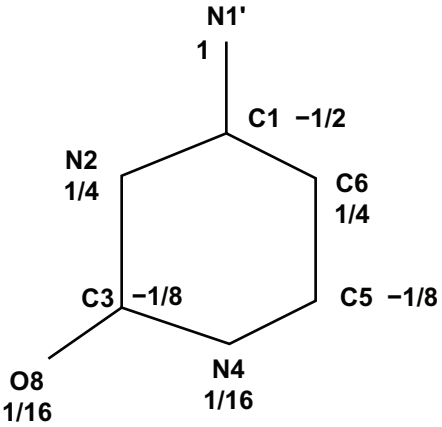

Figure 3 Calculation of Δ-values of single atoms.

The aim of this action is to find common characteristics in the orders of encoding units and create an understanding of how all genetic material, like nucleic acids and possibly also other material regardless of genetic code, can be of importance for the creation of life.

Stewart writes on mathematical interpretation of life: "It also has become clear that epigenetic information not written in DNA, and possibly not coded in any obvious symbolic fashion, is also vital to life on Earth."[16] He also believes that mathematics can "[…]provide significant insights into the science (biology) itself, to help explain how life evolved, how it works."[16] So the mathematics of

biology of our day is more than a tool to work up the results of scientific investigations. The mathematical modeling of the molecular units used in genetic coding is of particular interest.

We find regularities in the chemical interaction necessary for the base pairing and the formation of two-chain models of polynucleotide and amino acid chains, related to molecular and helical symmetries between them, which must be mathematical. The two-chain model in macromolecules is one example of repeated findings of two-ness in biological systems. Important biological objects come in pairs, such as the symmetrical A and B form of the helices.[12]

The idea exists that the Chargaff rule has its origin in a symmetrical pattern of the genetic code. Adenine equaled thymine (uracil), and guanine equaled cytosine, because of a yet-undiscovered role in the ordering of the bases (see in Figure 6 the base pairing of guanine and cytosine as an example).[15] If the complementary scheme was right, one might find attractive forces between bases with different structures. This idea opposes the typical view that symmetries between the most complicated of all molecules are made out of forces existing by which like attracts like.

Gene replication starts with the separation of symmetrical but not structurally identical chains. New daughter strands are created on the parental templates, thereby forming two DNA molecules with a reciprocal identity to the very first strands. After pairing of strands with complementary bases, two newly synthesized helices are generated.[9]

One supposes that the reason for the Chargaff rule is similar to that of the genetic code. The mating of the bases with complementary ones, in addition to helix formation and gene encoding, is managed primarily by the side groups. This is symmetrical correspondence of great importance. Figure 4 demonstrates the symmetrical patterns with the pairing of the side groups of complementary molecules of amino acids, DNA and RNA and between RNA and amino acids.

The symmetry of molecules can be disturbed by the availability or the addition of side groups. Although unexpected, molecular pairing and mutual genetic coding can occur between substances of different symmetry and structure.

Herein it should be described how the passing on of genetic information proceeds between carriers. It has been shown here how the calculation of Δ-values is performed. Δ-Values of the atoms involved in the genetic coding (of the functional side groups) from amino acids are matched to the nucleic acids.

Table 2 Side groups of amino acids in ratios −1/2 of their differences

	N3	C2	C1	O2
Thr	1.86	1.99	−1.31	3.65
Val	1.91	1.88	−1.26	3.62
Gly	1.91	1.88	−1.26	3.62
Pro	1.91	1.88	−1.26	3.62
Isl	2.14	1.43	−1.04	3.53
Cys	2.18	1.34	−0.99	3.49
Ser	2.31	1.09	−0.86	3.43
Gls	2.34	1.04	−0.84	3.42
Gln	2.34	1.03	−0.84	3.41
Arg	2.46	0.80	−0.71	3.35
Lys	2.53	0.65	−0.64	3.32
Try	2.55	0.61	−0.62	3.31
Met	2.56	0.58	−0.61	3.30
His	2.59	0.53	−0.59	3.29
Tyr	2.61	0.49	−0.56	3.28
Phe	2.64	0.43	−0.53	3.26
Ala	2.81	0.09	−0.36	3.18
Leu	2.81	0.09	−0.36	3.16
Asn	2.85	0.02	−0.33	3.16
Asa	2.86	0.01	−0.31	3.15

Abbreviations: Thr, threonine; Val, valine; Gly, glycine; Pro, proline; Is, isoleucine; Cys, cysteine; Ser, serine; Gls glutamic acid; Gln glutamine; Arg arginine; Lys lysine; Try, tryptophan; Met, methionine; His, histidine; Tyr, tyrosine; Phe, phenylalanine; Ala, alanine; Leu, leucine; Asn, asparagine; Asa, aspartic acid.

First, it must be recognized in Table 2 that, within the amino acids, the differences of the Δ-values always behave according to the −1/2 ratio between the oxygen atoms (O2) of the side groups involved in the peptide bindings to the differences of their basic carbon atoms (C1) and the differences of the nitrogen atoms also involved in the peptide bindings (N3) to those of their basic carbon atoms (C2). These relations are a part of the symmetry, not only in the amino acid molecules but also in the nucleic acids (Table 3). This will be explained by Figure 5. In A the difference, ie of Asa N3 2.86 with Adenine N6′3.19 and C2 0.01 with C6 −0.67 of the same molecules and in B the difference of thr O2 3,65 with guanine O6′3,62 and C1 −1.31 with C6 −1.24.

With the coding of the amino acids by the nucleic acids, the atoms of chemically similar side groups communicate in stereo with each other and exchange information. That is, one

Table 3 Side groups of the nucleic acids (in clips of their basic atoms; numbering according to Figure 6)

Cytosine	N6′: 3.46	C6: −1.20	O2′: 3.67	C2: −1.35
Uracil	O6′: 3.64	C6: −1.29	O2′: 3.69	C2: −1.38
Guanine	O6′: 3.62	C6: −1.24	N2′: 3.50	C2: −1.30
Adenine	N6′: 3.19	C6: −0.67		
Thymine	N6′: 3.20	C6: −0.39	O2′: 3.58	C2: −1.16

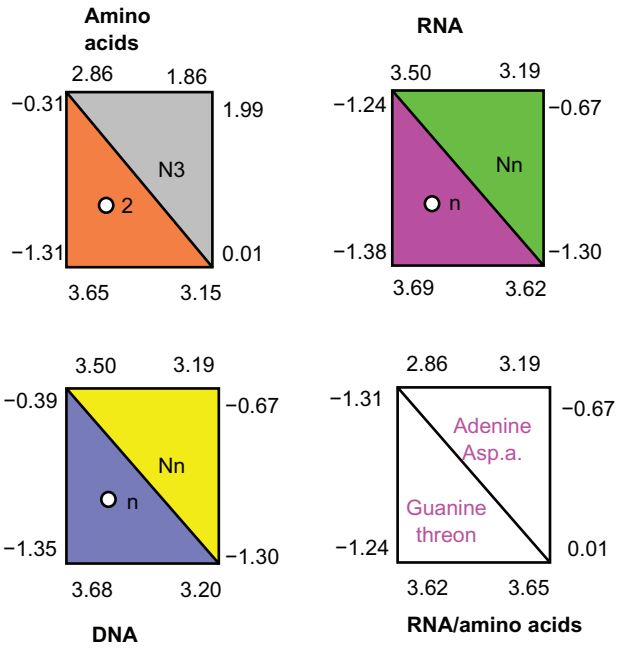

Figure 4 Symmetry as basis of the interaction of coding molecules in the genetic code (based on the subtraction of values technique).
Abbreviations: Asp.a., aspartic acid; threon, threonine.

amino acid atom (O2) forms, eg, with the atom (O6′) from guanine, a Δ-value difference that is bigger or smaller than the difference of another amino acid atom (O2) with the same atom (O6′) from guanine. This then guarantees, in the context of all values in the codon, that the amino acids concerned find their place in the originating protein chain.

If the same atomic classes are not found in both molecule classes (two encoding O-atoms as well as two encoding N-atoms), the classification is carried out only by the atomic type, which exists with both molecule classes. Moreover, besides the difference forming between the groups within the same molecules, this also takes place between groups of different molecules.

The linear course of the Δ-values of the functional groups and their basic carbon atoms proceeds with and in between them symmetrically. The next step shows how in each case,

A	B
2.86 − 3.19 = −0.33	−1.31 + 1.24 = −0.07
0.01 + 0.67 = 0.68	3.65 − 3.62 = 0.03
−1:2	−2:1
aspartic acid - adenine	threonine-guanine

Figure 5 The symmetrical relations between coding molecules.

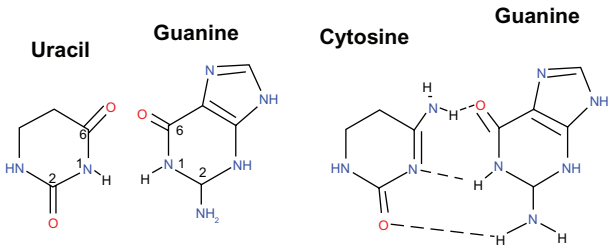

Figure 6 Base pairing is only possible for the pairing cytosine–guanine.

atoms of the RNS and amino acids meet, so that with symmetrical relations the coding can take place.

With identical symmetries, as illustrated in Figures 4 and 5, an undisputed choice can take place, and each amino acid can find its place on the grounds of its Δ-values with the translation in the genetic code. The formation of functioning protein chains is guaranteed. The author has, in his monograph *The Genetic Code: A Computational Interpretation*,[17] developed the Δ-values and the coding of symmetrically running molecule parts in the genome first described.

The Chargaff rule

This is based on specific features of chemical structures. The generations of pairs leading to the double helix only take place between bases if a hydrogen bridge formation is possible between them, ie, if the necessary H-atom is made available by a side group (with uracil and thymine 1-N-1 ′H). Then the formation of an H-bridge with the complementary base (adenine) is possible after a tautomere relocation of 1-N takes place. In a similar manner, a tautomere relocation takes place with O6′ of uracil and thymine, so that an additional H-bridge is formed to 6′-N-6″-H by adenine. The corresponding interaction takes place between 6′-O, 1-N-1′-H, and 2′-N-2″-H of guanine and 6′-N-6″-H, 1-N, and 2′-O from cytosine, so that between these pairs three H-bridges originate.

Such a generation of pairs is not possible between uracil or thymine and guanine, or between cytosine and adenine (or insufficiently); it cannot lead to a double-helix formation in such a case.

The structural relations should be made clear with help of uracil and guanine and cytosine and guanine (Figure 6).

How everything started: the step from the inanimate to life

From investigations by Wächtershäuser,[18] we know that iron oxide, copper oxide, nickel oxide, and nickel sulfides catalyze the formation of organic hydrocarbon molecules, but also the

Table 4 Δ-value relations

Cu-O on C-O	C	1.58	O	2.21
	Cu	3.98	O	1.01
Difference		**−2.40**		**1.20**
Fe-O (magnetite) on N-O	N	1.90	O	2.05
	Fe	3.81	O	1.10
Difference		**−1.91**		**0.95**
Cu-O on threonine (As)	C1	−1.31	O2	3.65
	Cu	3.98	O	1.01
Difference		**−5.29**		**2.64**
Cu_2O on threonine	N3	1.86	O2	3.65
	Cu	5.47	O	1.83
		−3.61		**1.82**
Fe-O on guanine (Ns)	C3	−1.24	O11	3.62
	Fe	3.81	O	1.10
Difference		**−5.05**		**2.52**
Fe_2O_3 on threonine	N3	1.86	C2	1.99
	Fe	2.76	O	0.19
Difference		**−0.90**		**1.80**

accumulation of carbonyl, hydroxyl, and amino groups in already-existing molecules under suitable conditions.

Between such molecules, the 1:−2 ratios in relation of their Δ-values are shown in Table 4.

If such catalyzing metal compounds are given as crystals, they form in themselves closed and symmetrically regular bodies and form in this arrangement from organic molecules symmetrically regular chains, in secondary processes helices, while also following the laws of symmetry.

Summary

Proteins, as building blocks of all living beings, are composed out of sequences of amino acids. By the cooperation of three nucleic acids at a time out of four different kinds, called the genetic code, one amino acid is chosen and added to a chain of amino acids. This procedure is outlined, controlled, and directed by a kind of information through its transport from storage molecules: the DNA, to the RNA, and from there to the amino acids.

The discovery of the chemical structure of DNA and RNA and their arrangements as helices by Watson and Crick[5] in 1952, with contributions of X-ray diffraction studies by Wilkins et al,[19] was rewarded with the Nobel Prize in Medicine in 1962. In 2012, the 50th anniversary of this award was commemorated.

The formulation of the awarded paper of Watson and Crick[5] was based on the preceding works of Chargaff[15] and Woese.[13] Chargaff recognized that pyrimidine bases always pair with a specific molecule of purines (uracil in RNA or thymine in DNA with adenine and cytosine with guanin), which lead to the typical double helix of two complementary strings of nucleic acids.

Woese also pioneered this topic when he recognized that each of the four naturally occurring bases of the biological code act as coding substances by the substituents they carry on the 2 and 6 positions in the ring, adenine being 2-H, 6-NH_2, cytosine being 2-OH, 6-NH_2, guanine 2-NH_2, 6-OH, and uracil (and thymine in DNA) 2-OH, 6-OH.[13] It appears probable that coding will be done exclusively in terms of the 2 and 6 positions only. It means that focus on the genetic code-carrying information should be confined to the 2 and 6 positions rather than to the four bases as a whole.

If only the side groups of the nucleic acids and nothing else were of importance for the coding of proteins, then only 1–2 side groups out of 7^2 (49) possibilities that were different in molecule and position would be sufficient to code for 20 different amino acids. But in reality, 64 triplets formed by three bases are necessary for this task. So other factors must be considered that have an influence.

Stereochemistry, symmetry, and mass seem to carry importance. So a mathematical approach, an algorithm, had to be developed to provide a deeper understanding of how genetic coding works and which necessary factors are involved. By an algorithm, so-called Δ-values are produced to give an estimate of chemical reactivity, the chemical–genetic information of a side group, and an indication as to the amount that it contributes to the symmetry of a compound. By this, it can be shown whether two functional groups, of the same or other molecules, can react with each other and if they exchange information. Therefore the encoding interaction between nucleic acids and amino acids can be explained by the presence of symmetric correspondences.

Furthermore, symmetric correspondences of Δ-values exist when inorganic molecules catalyze organic molecules, thus possibly generating nucleic acids and bringing dead material to life.

Disclosure

The author reports no conflicts of interest in this work.

References

1. Petitjean MS. A definition of symmetry. *Symmetry Sci Art*. 2007;10:99–119.
2. Rafiki [homepage on the Internet]. Bloomington, IN: Rafiki, Inc; 2003–2007 [updated 2008). Available from: http://www.codefun.com. Accessed January 7, 2013.
3. Lima-de-Faria A. The atomic basis of biological symmetry and periodicity. *Biosystems*. 1997;43(2):115–116.
4. Frank FC. The influence of dislocations on crystal growth. *Discuss Farady Soc*. 1949;5:48–54.

5. Watson JD, Crick FH. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature*. 1953;171:737–738.
6. Gamow GM, Ycas M. Statistical correlation of protein and ribonucleic acid composition. *Proc Natl Acad Sci U S A*. 1955;41:1011–1019.
7. Bauer J. *The Cooperative Gene*. Hamburg: Hoffmann und Campe Hamburg; 2008. German.
8. Calder N. *The Life Game*. London: BBC; 1973.
9. Eigen M, Winkler R. Ludus vitalis. In: *Mannheimer Forum*. Vol 73/74. Mannheim: Boehringer Mannheim; 1974;53:139. German.
10. Mereschkowski K. The theory of two plasms as the basis of symbiogenesis, a new study of the origins of organisms. *Biolog Centralbl*. 1910;30:353–367. German.
11. Margulis L. *The Origin of the Eucaryotic Cell*. New Haven: Yale University Press; 1970.
12. Flechtner HJ. *Fundamentals of Cybernetics*. Stuttgart: Wissenschaftliche Verlagsgesellschaft Stuttgart; 1970. German.
13. Woese CR. Nature of the biological code. *Nature*. 1962;194:1114–1115.
14. Shannon CE, Weaver W. *The Mathematical Theory of Communication*. Woods Hole: Urbana USA; 1949.
15. Chargaff E. Structure and function of nucleic acids as cell constituents. *Woods Hole: Fed Proc;* 1951;10:654–659.
16. Stewart I. *The Mathematics of Life*. Philadelphia, PA: Basic Books; 2011.
17. Schmidt FH. *The Genetic Code: A Computational Interpretation*. Marburg: Tectum; 1996. German.
18. Wächtershäuser G. Origin of life as we don't know it. *Science*. 2000;289(5483):1307–1308.
19. Wilkins MH, Stokes AR, Wilson HR. Molecular structure of deoxypentose nucleic acids. *Nature*. 1953;171:738–740.