Autocatalysis, Information and Coding

Peter R Wills

Department of Physics, University of Auckland, Private Bag 92019, Auckland,

New Zealand

fax: +64 - 9 - 373 7445

email: p.wills@auckland.ac.nz

URL: http://www.phy.auckland.ac.nz/staff/prw/

This paper has been prepared for the "Physics and Evolution of Symbols and Codes" special issue of BioSystems.

Abstract: Autocatalytic self-construction in macromolecular systems requires the existence of a reflexive relationship between structural components and the functional operations they perform to synthesise themselves. The possibility of reflexivity depends on formal, semiotic features of the catalytic structure-function relationship, that is, the embedding of catalytic functions in the space of polymeric structures. Reflexivity is a semiotic property of some genetic sequences. Such sequences may serve as the basis for the evolution of coding as a result of autocatalytic self-organisation in a population of assignment catalysts. Autocatalytic selection is a mechanism whereby matter becomes differentiated in primitive biochemical systems. In the case of coding self-organisation it corresponds to the creation of symbolic information. Prions are present-day entities whose replication through autocatalysis reflects aspects of biological semiotics less obvious than genetic coding.

Keywords: autocatalysis, self-construction, coding, reflexivity, information, prions

1. Introduction

Pattee (1995; 1996; 1997) has challenged theoretical biologists to elaborate the deep connection between alternative conceptions of biological processes as material mechanisms and as semiotic operations of information, symbols and codes. He urges us to investigate the organizational requirements that would allow molecules to function as descriptions. In essence we are being sent back to look at the question "What is Life?" anew. The current trend in experimental biology is to claim that the processes used to dissect, analyse and even alter organisms can be explained solely on the basis of mechanistic physical principles. However molecular biology has been unable to emulate the disciplines of physics and chemistry in which elementary quantities like “length” and “force” and entities like "electron" and "molecule" have unambiguous definitions both as theoretical concepts and as primitives of observation. Molecular biology is an intellectually closed, reductionist enterprise. We must return to an examination of the basic concepts used in biology in order to find a way of expanding the conceptual horizon and of illustrating the intrusion of semiotics into what is taken to be the purely material world of biology.

In spite of our impression that the genome of an organism provides a "blueprint" for its construction, we find on reflection that we have been deceived by the overwhelming generality, near universality, of the mechanisms of genetic information processing. In fact, the means of interpreting genetic information are as vital to molecular biological systems as the information itself. Although we talk in familiar terms about "the gene" (or genes) for some phenotypic characteristic, and even find that characteristics can often be transferred by means of the transposition of genes between diverse taxa, it is quite impossible to demarcate completely either

the peripheral genetic context or the biochemical preconditions that must remain approximately constant for the perceived association between genes and traits to be maintained. Our knowledge of the cellular components involved in the interpretation of genetic information, ribosomal RNA and proteins, amino-acyl tRNA synthetases and so on, and the manner in which the construction of those components depends on genetic information, allows us to confuse the physical and semiotic levels of description and say that the structure of those components is genetically encoded. However it is equally true to say that genetic information would be meaningless, that is, no coding would be possible at all, without the prior construction of those very components.

This paper is an exploration of the dual aspects, physical and semiotic, of selection in prebiotic evolution: the differentiation of molecular forms that is necessary for the definition of information in physical systems; and the self-organisation of coding necessary for the interpretation of information into functional form, a subject also treated by Bedian (2000). The role of autocatalysis in both the replication of genetic information and the maintenance of operational functions will be emphasised, illustrating that there is no solely physical basis for the separation of information and function.

2. Macromolecular Self-Construction

If we were to consider, in purely material terms, any replicating biological system, from a prion to a human, then the simplest description we could give would refer only to the molecular components and mechanical operations required for its construction and maintenance. The most

significant feature we would expect to emerge in our description of the system would be the capacity for reproduction of the original form, but this would probably involve the reproduction of various smaller components along the way. If we ignore the fact that the different components of a cell require very specialised spatial ordering and that different regions of cells have very different kinds of processes going on in them, some progress can be made. It is possible to specify the formal, semiotic relationships that must be satisfied for a group of component molecules (or physical states) to be able to reproduce themselves, especially in the manner of cooperative cross-catalysis that one imagines took place early in the prebiotic phase of evolution.

Consider, quite generally, the class of operations $\{O_i: i = 1, 2, 3 \ldots N\}$ and molecular components $\{M_j: j = 1, 2, 3 \ldots M\}$ constituting a biological system, from the point of view of a material, mechanical description. Any operation O_i must be carried out by a component or components of the system, no immaterial agency being admitted. Thus, the available functional capacity of the system, representing the rates $\mathbf{w} = (w_1, w_2, \ldots. w_N)$ at which operations O_i can be carried out, are functions of the population numbers $\mathbf{x} = (x_1, x_2, \ldots. x_M)$ of the various components:

$$w_i = f_i(\mathbf{x}) \tag{2}$$

Also, any component can only be synthesised as a result of a number of operations being performed:

$$dx_j/dt = g_j(\mathbf{w}) \tag{3}$$

While the f_i represent the embedding of functional operations O_i in a field of material components M_j, the g_j represent sequences $n_j(k: k = 1, 2, \ )$ of operations needed to construct the components. Detailed consideration has been given to the special case of a field comprising

polymers of length ν, assigning $w_i = \Sigma\, w_{ij}\, x_j$ and $dx_j/dt = \Pi w_{n_j(k)}$ with $n_j(k) \in \{1, 2, \ldots N\}$ and $k = (1, 2, \ldots \nu)$ ligation steps. With $w_{ij} \in \{0, 1\}$ assigned according to some probability rule, the threshold for selection in a homogeneous biochemical system is well characterised for autocatalytic ligation systems (Wills & Henderson, 1997) and coding systems (Wills, 1993; Wills, 1994). Beyond the dynamic instability the system comprises a subset of components and a subset of operations that are related to one another reflexively: all of the components required to carry out the selected operations can be constructed by carrying out operations from within the selected subset of operations; all of the operations required to construct the selected components can be carried out by components from within the selected subset of components. The existence, *a priori*, of the semiotic relationship of reflexivity between operations and components is a necessary condition for the emergence, through a dynamic process of selection, of a self-constructing molecular system. Kauffman (1986) has argued that reflexivity is an inevitable collective property of sufficiently complex sets of protein catalysts.

Whether a collectively autocatalytic set of molecular components can emerge from a system and maintain itself stably depends primarily on the embedding of functional operations in the material field. If the field of material components is the sequence space of proteins, then defined catalytic functionalities may be seen to be embedded at different locations within the space. This embedding is better known as the protein "structure-function relationship". Any particular functional operation within a system, such as the capability to catalyse a particular reaction, can be determined, in principle, from fundamental physical principles, that is, from quantum mechanics. However the possibility of self-construction derives from a formal, semiotic relationship of reflexivity between structures and functions, which is at most only implicitly dependent on the physical principles that underlie the structure-function relationship. The

mathematical forms of f and g might be calculable from quantum mechanics, but a dynamic system described by f and g will evolve into a self-constructing system only if f and g impose a selective advantage on some subsets of operations and components that are related to one another reflexively. This is seen most easily in the case of molecular components that are constructed out of a class of more elementary components, like a few amino acids or nucleotides (Wills & Henderson, 1997).

3. Coding Systems

A genome encodes an organism only within the context of well-specified molecular biological functions and no organism can be generated *de novo* solely from the symbolically encoded genetic information that is utilised in reproduction. This raises the question of what must be given, *a priori*, besides genetic information, in order to make a living organism.

Consider a system in which specific macromolecules are constructed on the basis of a genetic information source, akin to the manner in which proteins are synthesized by ribosomes by using information encoded in messenger RNA molecules. In thinking about the origin of life, we have a "bootstrap problem". The amino-acyl-tRNA synthetases that assign amino-acids to codons are themselves proteins whose sequences are encoded in genes. Ribosomes cannot produce functional proteins from genetic information unless the tRNAs are charged with the correct amino-acids, so where did the first functional synthetases, able accurately to distinguish within classes of amino acids and codons, come from? The correct genetic information encoding the

synthetases needed to make the assignments defining a specified code is useless without the proteins themselves.

One part of this prebiotic bootstrap problem has a simple solution which was first elaborated by Bedian (1982). Some genetic sequences are reflexive with respect to the mechanism of their interpretation: when they are interpreted by a certain subset of all the proteins that can act as assignment catalysts, the result is the construction of just that subset of assignment catalysts. Such reflexive genetic sequences create a dynamic instability in the synthesis of catalysts that enables the selective construction of a minute subset from a very large class of polymers (Wills, 1993; 1994). Polymers from the selected subset make coding assignments that lead to their own synthesis. However, there is no intrinsic feature of a particular genetic sequence that confers on it the property of reflexivity. The differential equations describing the dynamics of such systems require specification of the embedding of catalytic functions in polymer sequence space, in effect, the structure-function relationship for the polymeric assignment catalysts. Only within the context of a specified catalytic structure-function relationship can it be determined which genetic sequences have the semiotic property of reflexivity.

Consider a representation of coding relationships in terms of a set of assignment reactions, for example, {A→a, A→b, B→a, B→b} between two alphabets, of codons {A, B} and amino acids {a, b}. Under the assumption that there exist linear polymers, proteins comprised of a and b, that constitute a set of catalysts, each of which carries out just one of the assignment reactions, a requirement for the evolution of coding is that the assignment catalysts be distinguished from one another in a minimum of two positions. A representation of a minimal set of assignment catalysts would be {-a-a-, -a-b-, -b-a-, -b-b-} and any embedding can be represented as a one-to-one mapping between these assignment catalysts and the assignment reactions. Take, as an example, the following embedding:

-a-a- catalyses the assignment A→a

-a-b- catalyses the assignment A→b

-b-a- catalyses the assignment B→a

-b-b- catalyses the assignment B→b

The exclusive presence of the genes -A-A- and -B-B- will allow selection of the proteins -a-a- and -b-b- to occur, leading to the establishment of the code [A→a, B→b], because the genes -A-A- and -B-B- constitute a reflexive representation of the catalysts -a-a- and -b-b- within the context of this embedding. Similarly, exclusive presence of the alternative pair of genes -A-B- and -B-A- will allow selection of the proteins -a-b- and -b-a- to occur, leading to the establishment of the code [A→b, B→a].

The very possibility of dynamically stable coding is constrained by the form of the structure-function relationship of the relevant assignment catalysts. Some embeddings of catalytic functions (amino-acid to codon assignments) in polymer sequence space give rise to a situation of "ambiguity" in which competing codes that have no assignments in common require identical genetic information for the specification of their respective assignment catalysts (Nieselt-Struwe & Wills, 1997). Stable coding is not possible in such a system because the genetic information needed for a code is ambiguous with respect to the subset of possible assignment catalysts specified by and specifying that code. The salient point is that the physical instantiation of any symbolic encoding of protein sequence information in genes is constrained by purely formal, semiotic features of the way in which the catalytic properties of proteins are related to their sequences. For coding to be able to evolve it is not enough for quantum mechanics to allow that a set of proteins that execute the rules of a genetic code can be constructed. Stability also demands that there does not exist any other subset of possible assignment catalysts that can specify its own synthesis from the same genetic information. The reflexive genetic sequences required by potentially competing codes must be distinct. This is a semiotic rule with which any physical system must comply for genetic coding to be possible.

4. Functional and Structural Decomposition

A physical system can be said to contain information only after certain possible states have been defined as constituting a class, according to some criterion of membership. In the case of information which is stored in linear polymers like nucleic acids or proteins, there is an elementary class of distinguishable substates known as an *alphabet* of monomers, either

nucleotides or amino acids, small molecules comprised of covalently bonded atoms. The criterion for membership of the alphabet is that a molecule possess a chemically defined set of "functional groups". All members of the alphabet have these features in common, but their structures show individuality elsewhere. The members of an alphabet can be grouped together as a class because they all undergo certain reactions on which their individuality has no great effect. On the other hand, the hyper-astronomical number of possible polymers of even modest length, all with certain individual properties, is what provides the extraordinary variability in protein structure and function. In the same way, the existence of the recognition processes involved in molecular biological translation and replication provides the means whereby nucleic acid sequences constitute heritable genetic information.

If we set aside the arbitrary perspective of the means that we use to represent chemical structures and ask how a class or alphabet of molecules is defined intrinsically within a biochemical system we must do so in terms of a relationship between components and operations. Molecules on which a common operation is performed irrespective of their individuality can be said to form a class relative to that operation. However, in order to make this definition stick we still have to draw an arbitrary boundary circumscribing the physical sphere being considered, as well as the level of precision beyond which any differences within the sphere are to be regarded as insignificant. Herein lies the fundamental problem for anyone interested in the evolution of molecular biological functionality. It is precisely the apparently progressive increase in the precision with which entities that are otherwise identical come to be distinguished from one another that is in need of explanation.

The decomposition of functions and structures resulting from the selection of autocatalytically self-constructing systems provides the necessary explanation (Wills & Henderson, 1997). By way of illustration, consider a class of polymers which are constructed by ligation from a binary alphabet of monomers {a, b}. Suppose that polymers are synthesised by operation of the ligation reactions {a¬a, a¬b, b¬a, b¬b} where x¬y represents the reaction …-x + y-… → …-x-y-… between the terminal monomers of two directed polymers …-x and y-… for x, y ∈ {a, b}. If some of the polymers constructed by concatenation of a and b catalyse each ligation reaction in the system then it is possible for a self-constructing subsystem to be selected, under the condition that some subsets of reactions and polymers bear a reflexive relationship to one another. How might it then be possible for one letter of the alphabet, b for example, to be differentiated into two recognisable forms, β and γ? Should such differentiation occur, the binary alphabet {a, b} would be decomposed into the ternary alphabet $\{\alpha, \beta, \gamma\}$ according to the mapping a→α and b→$\{\beta, \gamma\}$. Such a decomposition could occur only if a relationship of reflexivity prevailed between some subset of structures involving $\{\beta, \gamma\}$ and selective catalysis of reactions from within one of the sets $\{\alpha\neg\beta, \alpha\neg\gamma\}$, $\{\beta\neg\alpha, \gamma\neg\alpha\}$ or $\{\beta\neg\beta, \beta\neg\gamma, \gamma\neg\beta, \gamma\neg\gamma\}$. Decomposition of the alphabet {a, b} into $\{\alpha, \beta, \gamma\}$ and the concomitant selection of reactions from within the enlarged sets can be regarded as a re-enactment of the initial selection, from within an initially undifferentiated population, of a collectively autocatalytic set of polymers comprised of differentiated monomers a and b on the basis of their reflexive relationship with a subset of reactions from the set {a¬a, a¬b, b¬a, b¬b}.

5. Evolution of Molecular Symbolism

Having outlined the conditions under which the precision of recognition within a self-constructing systems of polymers can potentially increase, it is pertinent to ask how such an increase in precision might be represented and stored genetically. The problem is by no means trivial. We simply ignored it when we discussed how a coding system can evolve when a certain subset of suitably embedded assignment catalysts is able to catalyse its own formation by interpreting an appropriately reflexive genetic sequence. We did not discuss what conditions must be satisfied for the initial selection, let alone the stable maintenance, of a reflexive genetic sequence.

Eigen (1971) described the process of Darwinian selection in a population of individually replicating macromolecular genetic information carriers, but the more elaborate theory of the "hypercycle" (Eigen and Schuster, 1979) did not give a satisfactory account of the establishment of the genotype-phenotype relationship in prebiotic evolution. While this second problem is addressed by theories of coding based on information-dependent macromolecular self-construction (Wills 1993; Nieselt-Struwe & Wills, 1997; Wills & Henderson, 1997), the way in which coding self-organisation can stabilise the selection and maintenance of the reflexive genetic information needed for an orderly genotype-phenotype relationship remains something of a puzzle. We can highlight the problem by applying the idea of functional and structural decomposition to the evolution of coding.

Consider again the phenomenon of coding self-organisation among a population of assignment catalysts whose structure-function relationship is specified by the embedding of reactions in an {a, b} sequence-space as discussed in Section 3, above. The protein catalysts made of

monomers from the alphabet {a, b} could be regarded as being made of monomers from an enlarged alphabet {α, β, γ} if there existed chemical processes that differentiated between the β and γ forms of b. Within the context of coding this would also require the alphabet of codons to be enlarged from {A, B} to {A, B, Γ} in which there could be B and Γ forms of the codon B. Let us suppose that a subset of -B-B- genes, -B-B- and -Γ-Γ-, were reproduced more rapidly than the other forms, -B-Γ- and -Γ-B- , and came to dominate the population as a result of natural selection. We can now recapitulate the argument concerning the relationship of reflexivity between the information and the functional embedding made previously in respect of the alphabets {A, B} and {a, b}. Provided the embedding were such that the -β-β- and -γ-γ- forms of the -b-b- proteins catalysed $B\rightarrow\beta$ and $\Gamma\rightarrow\gamma$ assignments, then a further process of coding self-organisation could result in the binary coding assignment B$\rightarrow$b being replaced by the differentiated assignments $B\rightarrow\beta$ and $\Gamma\rightarrow\gamma$. The final outcome would be the decomposition of the binary [A$\rightarrow$a, B$\rightarrow$b] code into the ternary [$A\rightarrow\alpha$, $B\rightarrow\beta$, $\Gamma\rightarrow\gamma$] code. This entire process is depicted pictorially in Figure 1.

FIGURE 1 HERE

What is missing from this analysis is any explanation of the natural selection of genetic sequences possessing the special property of reflexivity in respect of the embedding of coding assignments in the polymer sequence space. Why should genetic sequences with the property of reflexivity *vis-à-vis* the protein structure-function relationship have higher reproductive fitness than other genetic sequences? One could propose two quite different answers to this question. The first would be to invoke a biological version of the anthropic principle and say that the

coincidence between reflexivity and high reproductive fitness is a necessary condition for the evolution of coding. In any world in which the necessary condition was not fulfilled there would be no molecular biological coding and no organisms whose biochemistry needed explaining. The second approach to the question would be based on the assumption that the reproductive fitness of any entity, including prebiotic genes, depends on a network of interactions between different species in the system and that there is a mechanistic connection, albeit probably indirect and excursive, between the reproduction of genetic sequences and the products of their translation. In that case we would want to propose a dynamic coupling, perhaps as simple as that which the compartmentalization of related processes imposes, between the kinetically driven self-organisation of coding and the rate of replication of the associated genes. It is reasonable to think that the improved functional specificity achievable through more highly differentiate coding could allow for the more efficient replication of associated genes, but it may be necessary to demarcate a set of physically bounded entities, constituting something more like individual organisms than abstract conglomerates of quasi-biochemical processes, before the connection between reflexivity and fitness of genes became really plausible. As yet these speculations remain unelaborated.

In spite of incompleteness in our understanding of all the related mechanisms that underlie the evolution of precise molecular biological coding, we have outlined two preconditions for the spontaneous emergence and storage of useful symbolic information in self-constructing molecular systems. The first requirement is for reflexivity in respect of the mapping from "information carriers" to "functional catalysts" so that the particular catalysts needed to define an orderly mapping can be selected. The second requirement is for an association between the

property of reflexivity and reproductive fitness in the population of information-carrying molecules.

7. Discussion

Autocatalysis, by virtue of its very definition, requires the identification of some entity as both the product and the effector of a physical process. In its simplest manifestation, autocatalysis is the mechanism of macromolecular replication. Replicating systems are subject to Darwinian selection, the dynamics with which we are familiar from studies of inheritance; how autocatalysis gives rise to genetic inheritance is well understood (Eigen, 1971). Detailed features of complex macromolecules may be said to carry information because molecules with many alternative forms of those features may be replicated. Perhaps we should not then be surprised that entities as simple as single protein molecules appear to be able to carry their own heritable biological definition, as in the case of prions and their strains.

The most elementary possible form of autocatalysis is represented by the chemical equation $A + B \rightarrow 2A$ which describes the replication of the proteinaceous infectious agents (prions) responsible for spongiform encephalopathies in mammals, as well as various forms of non-Mendelian inheritance in yeast and fungi (Wills, 1989; Eigen 1996). These proteinaceous infectious agents display phenomena, such as mutation, selection and strain stability, typical of more complex biological agents that carry and reproduce genetically encoded information in nucleic acid form, like viruses and bacteria. A slightly more complicated system of autocatalysis involving a ligation reaction, $A + B + C \rightarrow 2A$, displays dynamics typical of even richer

biological phenomena, the stable coexistence of apparently competing entities (Wills *et al*., 1998).

Prion strain information, though sometimes transmissible from one species to another (as appears to be the case with bovine spongiform encephalopathy and new variant Creutzfeldt-Jakob Disease), resides to quite some extent in other phenotypic features of the cellular environment in which replication proceeds. In contrast to genetic information, for which the existence of a biologically universal means of coding serves to define the character of the symbolic alphabet, there is a subtle interplay between what we take to be prion strain information and the means whereby it is encoded. Therefore, it is not surprising that the idea of prions was widely rejected by molecular biologists when it was mooted in its modern form by Prusiner (1982). According to the Central Dogma and the demands of Darwinian (as opposed to Lamarckian) evolution, the information specifying stable phenotypic differences between similar entities must be encoded in nucleic acid genes (Crick, 1970).

When a system or an entity reproduces through a mechanism of autocatalysis involving many different operations and components we may be at a loss to delineate how and where information *per se* is at play; a self-constructing system may emerge as a result of a selection process in which the idea of individual "fitness" has no definition. What is required is that the operations and components are related to one another "reflexively", a term whose definition belongs to the realm of semiotics rather than physics. Pattee (1995) has noted the relevance of theories of the origin of genetic coding to the matter-symbol distinction. Indeed, when the synthetic operations required for autocatalysis are conditionally dependent on an ordered set of molecular features, as in the case of assignment catalysts for the formation of linear polymers, the selection of a self-

constructing set of catalysts corresponds to the emergence of a code. The ordered set of molecular features upon which molecular synthesis is dependent can be regarded as a source of information, but selection of a code requires that the information be reflexive *vis-à-vis* the structure-function relationship of the catalysts.

An increase in the precision with which features of components in a self-constructing system can be differentiated depends on there being a relationship of reflexivity between molecular sub-features and their role in the mechanism of autocatalysis. When such a relationship exists it is possible for autocatalysis to lead to the decomposition of recognition operations so that the sub-features of the molecules can be differentiated from one another. In the case of coding operations, information carriers with the property of reflexivity *vis-à-vis* the more precisely differentiated structure-function relationship of assignment catalysts must replicate at a higher rate than their competitors. There is currently no satisfactory explanation of the conditions under which selective replication of reflexive information is likely to occur. The emphasis which Pattee (1997) places on "the interdependence of the semiotic domain of the heritable genetic memory and the dynamic domain of construction and function" is highly apposite for those interested in this fundamental problem in theoretical biology; and our relative lack of understanding of such matters should warn us against having any confidence in predictions concerning the ultimate outcome of effecting arbitrary transfers of genes among diverse taxa.

Acknowledgment I am grateful to Kay Nieselt-Struwe, Leah Henderson and many other colleagues for enlightening discussions, to Stuart Kauffman for his enduring support and to Gordon Mangan for his perpetual encouragement.

References

Bedian, V., 1982, "The possible role of assignment catalysts in the origin of the genetic code", Origins of Life 12, 181-204

Bedian, V., 2000, "Self-description and the origin of the code", BioSystems [same number].

Crick, F.H.C., 1970, "Central dogma of molecular biology", Nature 227, 561-563

Eigen, M., 1971, "Self-organisation of matter and the evolution of biological macromolecules", Naturwissenschaften 58, 465-532

Eigen, M., 1996, "Prionics or the kinetic basis of prion diseases", Biophys. Chem. 63, A1-A18

Eigen, M. and Schuster, P., 1979, The Hypercycle (Springer-Verlag, Berlin)

Kauffman, S.A., 1986, "Autocatalytic sets of proteins", J. Theor. Biol. 119, 1-24

Nieselt-Struwe, K. and Wills, P.R., 1997, "The Emergence of Genetic Coding in Physical Systems", J. Theor. Biol. 187, 1-14

Pattee, H.H., 1995, "Evolving self-reference: matter, symbols and semantic closure", Communication and Cognition - Artificial Intelligence 12 (1-2), 9-27

Pattee, H.H., 1996, "The problem of observables in models of biological organizations", in: Evolution, Order, and Complexity, E.L. Khalil and K.E. Boulding (eds.), (Routledge, London)

Pattee, H.H., 1997, "Causation, control, and the evolution of complexity", in: Downward Causation, P.B. Anderson, P.V Christiansen, C. Emmeche, and N.O. Finnemann (eds.)

Prusiner, S.B., 1982, "Novel proteinaceous infectious particles cause scrapie", Science 216, 136-144

Wills, P.R., 1989, "Quantitative Constraints on Mechanisms of Replication for Information-Carrying Prions", Prog. Clin. Biol. Res. 317, 669-677

Wills, P.R., 1993, "Self-organisation of genetic coding", J. Theor. Biol. 162, 267-287

Wills, P.R., 1994, "Does Information Acquire Meaning Naturally?, Ber. Bunsenges. Phys. Chem. 98, 1129-1134

Wills, P.R. and Henderson, L., 1997, "Self-organisation and information-carrying capacity of collectively autocatalytic sets of polymers: ligation systems", NESCI Interjournal, <http://interjournal.org>, Ms 102

Wills, P.R., Kauffman, S.A., Stadler B.M.R. and Stadler P.F., 1998, "Selection dynamics in autocatalytic systems: templates replicating through binary ligation", Bull. Math. Biol. 60, 1073-1098

Figure 1

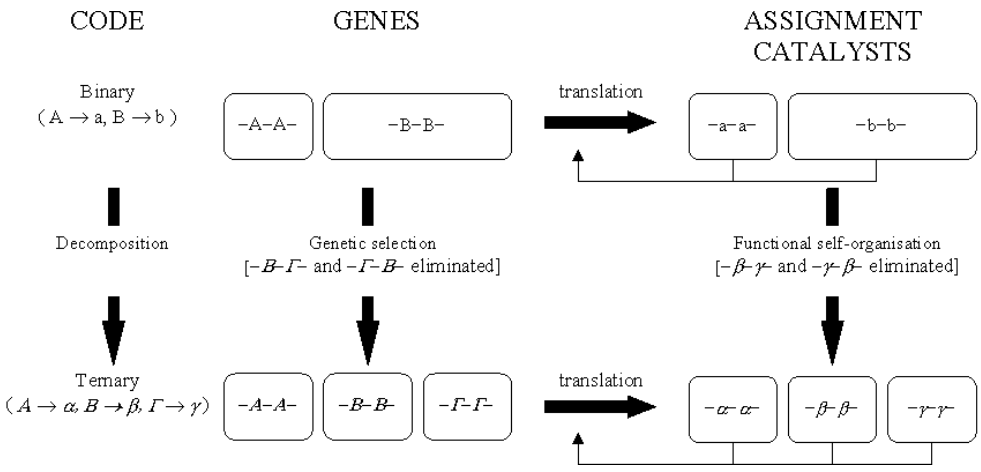

Figure 1. Decomposition of a binary code into a ternary code . Selection of a subset, {-*B*-*B*- , -*Γ*-*Γ*-} of the gene pool {-B-B-} allows corresponding selection of a subset {-*β*-*β*- , -*γ*-*γ*- } of the protein pool {-b-b-}