Stepwise evolution of molecular biological coding

Peter R Wills[1,2]

[1] Department of Physics, University of Auckland, Private Bag 92019, Auckland 1020, New Zealand
[2] Fraunhofer Gesellschaft, Biomolecular Information Processing (BioMIP), Schloss Birlinghoven, D-53754 Sankt Augustin, Germany
p.wills@auckland.ac.nz

Abstract

Two principles that embody necessary characteristics of self-sustaining physical systems are espoused. These principles are then used to analyze the dynamics of coding self-organization in the process of nucleic acid sequence-dependent protein synthesis. An artificial system is constructed in which catalysis of codon to amino-acid assignments is embedded hierarchically in protein sequence space. An example is provided of a system that evolves through execution of a coarse-grained binary code to execution of a more refined quarternary code. General implications for the construction of ALife systems are considered.

Introduction

We take as our starting point the fact that all molecular biological systems involve the translation of genetic information. Strings of nucleotides (RNA copies of DNA genes) serve as templates for the construction of strings of amino-acids that fold to give functional proteins. During translation, molecular recognition of a nucleotide triplet ("codon") causes the addition of a specified amino-acid onto a chain which, when completed, is a protein molecule. A set of proteins which are themselves products of the synthetic process, the amino-acyl synthetases, catalyze the physical process that assigns a particular amino-acid to a particular codon. The universal genetic code is often represented as a table listing a set of unique assignments from each of the 64 codons $X \equiv \{A, B, \ldots\}$ onto the set of 20 standard amino-acids $y \equiv \{a, b, \ldots\}$. When any of the 1216 assignments $X \rightarrow y$ not found in the coding table is made during translation, as occurs rarely in reality, it is considered by human observers to be an error.

Three basic problems confront our understanding of how such a complex self-contained information-processing system can maintain physical stability and accuracy, let alone have emerged from molecular disorder in the first place. These problems arise in one form or another in any consideration of systems that are autocatalytic – systems whose components are involved obligatorially in their own production.

The first problem is the "error catastrophe" predicted to arise due to amplification of dysfunctionality. Macromolecules that turn out to be partially dysfunctional due to random errors in their synthesis will participate in the synthetic processes required to produce further macromolecules that will be even more likely to be incorrectly synthesized and dysfunctional. Although first outlined by Orgel (1963) in relation to the process of translation, the essential solution to the error catastrophe problem was provided by Eigen (1971) who showed how Darwinian selection can lead to the accumulation of genetic information, rather than its loss, in populations of replicating macromolecules. Self-organization occurs when the accuracy of symbol replication is above a certain system-determined threshold.

Eigen (1971) did not solve the second problem, the instability of error-prone translation, but Hoffmann (1974) was able to show that stable translation can be achieved if the specificity of the adaptors assigning codons to amino-acids is likewise above some threshold. Bedian (1982) provided the first demonstration of thermodynamically driven coding self-organization. Wills (1993) and Nieselt-Struwe and Wills (1997) subsequently established that any translational mapping from nucleic acid to protein sequences can be established in practice only when codon-to-amino-acid assignment capability (adaptor functionality) is distributed among proteins in certain defined ways. That is to say, the way in which catalytic capability is embedded in protein sequence space, the so-called "protein structure-function relationship" must meet certain formal requirements before translation can ever emerge.

The third problem is the establishment and maintenance of functional coupling between the individually error-prone processes of information replication and translation. It is one thing to demonstrate, as Eigen (1971) has, that Darwinian selection takes place when the means of genetic replication is already provided, or as Bedian (1982) and Wills (1993) have, that coding self-organization occurs spontaneously when appropriate genetic information is provided, but it is quite another thing to demonstrate that the two processes can be mutually self-sustaining when they are coupled together and the effect of errors can seep from one process to the other. Only recently have Füchslin and McCaskill (2001) shown that simple reaction-diffusion coupling is sufficient to allow disordered processes of both kinds to condense into a fully-fledged operational system

Copyrighted Material

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/268191/9780262291392_cai.pdf by guest on 01 October 2023

of self-encoded translation in a cell-free system involving genetic replication.

None of the theories that have been proposed for these three problems is complete enough to give a plausible description of the pathway or timescale of prebiotic events. This is due in large part to the complexity and extreme specificity displayed by the final products of that early stage of evolution: 64 codons mapping onto a 20-member set of amino-acids. On the other hand, the universal genetic code displays regularities that offer potential clues to not only the history of its evolution but also some basic features of how the function of a protein varies with its amino-acid sequence, its primary structure. It has long been noted that the distribution of redundancy in the coding table is far from random and that amino-acids with similar properties are represented by similar codons. These regularities are at least partly explained in terms of chemical and structural similarities among codons or amino-acids. On this basis, a group theoretical decomposition of the coding table was provided by Hornos and Hornos (1993), but Nieselt-Struwe and Wills (1997) reframed the question to ask whether progressive decomposition from relatively coarse-grained to increasingly fine-grained translation might be a feature of the real-world historic events that gave rise to the universal code. In this paper I take the argument one step further by asking under what conditions dynamic processes could have led to the progressive refinement of the genetic code, all the way to the final form that has since survived to our own very late evolutionary epoch.

The concept of structure-function decomposition has been presented in several separate contexts (Nieselt-Struwe and Wills 1997; Wills and Henderson 2000; Wills, 2001). The basic idea is that the macromolecules that perform any function are selected in a progressive fashion such that inexact information is first used to specify a general class of catalysts from which those whose functions are more optimally refined emerge through a process of mutual selection. There is a close analogy to the progressive increase in Shannon information that Adami *et al.* (2000) report in the Avida replication system, except that we are concerned here with the way in which components that carry out elementary algorithmic operations cooperate to optimize their function, rather than the overall competitive advantage being seized by a whole quasi-species of algorithms.

I begin by enunciating two basic principles of complex autocatalytic systems and then show how those principles guide the construction of an artificial system in which the self-organization of genetic coding takes place in a stepwise fashion.

Theory

Scientific principles, like Newton's three laws of motion, express empirically valid generalizations that seem logically self-evident when the relevant phenomena are perceived from within some rigorously prescribed framework. The virtue of principles lies in the extent to which they limit the terms of phenomenological description but capture some universal aspect of the empirical domain under consideration. Here we are interested in how genetic information sustains the existence of complex molecular biological systems or their algorithmic counterparts.

The *principle of reflexivity* expresses a truism: all of the complex components of a self-sustaining system must be constructible as a result of executing the operations that their structures functionally enable; or, all of the elementary chemical reactions required within living systems must be catalyzed by the molecular components that are produced by execution of those reactions. The principle applies to all autocatalytic systems, but it has special application to organisms because their existence relies on genetically encoded information: the genetic sequences in a living cell must, *via* translation, be reflexive *vis-à-vis* the catalytic functions of the products of their translation, especially coding assignments. Let us consider the process of coding self-organization (Wills 1993).

Suppose there existed a very primitive mechanism, a completely indiscriminate process analogous to ribosomal protein synthesis, whereby proteins could be produced from RNA templates by joining amino-acids together in strings as a result of some sequential tri-nucleotide docking process. Such indiscriminate "translation" of sequences of triplet codons would produce completely random proteins. However, suppose that proteins with particular structures exerted an influence on the process, some proteins capable of biasing the assignment of particular amino-acids to particular codons during the synthetic process. For example, some proteins could cause a bias in the assignment of GYN, UUN and AUN codons to hydrophobic amino-acids. [The 4 standard nucleotide units, A, C, G and U, are categorized such that N stands for any of them and Y stands for the pyrimidines C and U. The particular codons mentioned here are actually mapped onto hydrophobic amino-acids in the universal coding table.] Suppose that the ability to cause such a bias in the assignment of amino-acids to codons depended on a preponderance of hydrophobic residues at particular positions, perhaps relative positions, in the proteins responsible for the assignment bias.

The presence of RNA templates with a preponderance of the specified codons in appropriate relative positions would allow specific "biasing" proteins to generate themselves autocatalytically in the manner of coding self-organization (Wills 1993). A primitive code in which GYN, UUN and AUN codons were assigned predominantly to hydrophobic amino-acids could evolve as a result of dynamic symmetry-breaking, induced by the matching of RNA template sequences to the structure-function properties of proteins and not in any way dependent on the chemical mechanism of peptide bond formation or direct chemical interactions between amino-acids and the codons to which they became assigned. The maintenance of coding would be based on the presence of specifically selected molecules with special

Copyrighted Material

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/268191/9780262291392_cai.pdf by guest on 01 October 2023

functions, namely, proteins that guaranteed the required assignment bias by virtue of the reflexive relationship of their amino-acid sequences, *via* translation, to the available genetic information.

How might the primitive code in such a system become refined so that the genetic information was interpreted in a more differentiated fashion and the assignment of amino-acids to codons was more selective? All we need do to answer this question is apply the principle of reflexivity again. Suppose in our example that the selective presence of alanine rather than other hydrophobic amino-acids at some positions in proteins conferred the function of biasing the assignment to alanine of GCN codons relative to other GYN, UUN and AUN codons. [This particular choice is made by way of example because GCN codons happen to stand for alanine in the universal coding table.] In the presence of genetic information with GCN codons in appropriate relative sequence positions, the concentration of the alanine-containing proteins could be autocatalytically amplified relative to the overall concentration of the original set of generally hydrophobic proteins. GCN codons would then be assigned to alanine as a result of the dynamic state of the system.

The recapitulation of reflexivity at increasing levels of functional specificity constitutes the *principle of decomposition*: a class of structures can become effectively decomposed into narrower categories through the operation of more refined recognition processes only if the products of those processes have the specified capability of selective recognition. Only when the effect of some new bias in functional specificity results in the synthesis of components that reinforce that bias can a system evolve as a result of autocatalytic selection. When the appropriate condition is met, more refined translation of the genetic information supplied to a system produces assignment catalysts that carry out translation assignments more selectively. In the example considered above, the emergence of the GCN-to-alanine from the GYN-to-hydrophobic assignment depended on the supposition that proteins containing alanine at sequence positions corresponding to GCN codons in the available genetic information were instrumental in biasing the placement of alanine at those positions in the sequences of proteins being produced.

Model System

I now describe a self-organizing coding system in which the embedding of assignment functions in the protein sequence space decomposes in a rigorous hierarchical fashion. The system first evolves to execute the assignment functions defined by a binary code that subsequently decomposes into a quarternary code. The design of the system is based on the principles of reflexivity and decomposition. First, the information supplied to the system is constructed so that components that carry out the assignments of a code are produced when they themselves are used exclusively for the translation process. Second, the requirements of decomposition are met by constructing a hierarchical embedding of assignment functions in the protein sequence space.

The first step in the construction of the system is to specify classes $\{K, L\}$ of codons and $\{k, l\}$ of amino-acids. In keeping with previous practice (Wills, 1993; Füchslin and McCaskill, 2001), single points in protein sequence space are chosen at random to represent "catalytic centers", optimal sequences for the performance of each of the 4 binary assignment functions $\{K, L\} \rightarrow \{k, l\}$. These arbitrary sequences are shown in the second column of Table 1.

K→k	llkkkklkkllk	A→a	ddbbabcaadca*
		A→b	cdbbbbcabcda
		B→a	dcaaaadaaccb
		B→b	cdababdbadca*
K→l	kkkllllllkkll	A→c	baacddddbadc
		A→d	abacdddcbbdc
		B→c	bbbcdcccbacc
		B→d	baadcddcbbdc
L→k	lkklkkllllkkl	C→a	dbadabcddbbc
		C→b	dabdabdcdabd
		D→a	daacbbcccaad
		D→b	cabdbbdcdbac
L→l	klkkkkkkllllk	C→c	bcbaabbbccdb*
		C→d	bcababaaddcb
		D→c	bdbbabbaccdb
		D→d	acabbabaddca*

Table 1. Protein sequences for optimal assignment activity. The nominal binary assignments, $\{K, L\} \rightarrow \{k, l\}$, were first embedded in the nominal binary sequence space of length 12. The sequences of the 4 optimal catalysts, chosen at random, are shown in column 2. The 16 possible quarternary codon to amino-acid assignments, $\{A, B, C, D\} \rightarrow \{a, b, c, d\}$, were then embedded in the defined subspaces of the quarternary sequence space. Assignments and optimal catalytic sequences for the code $A \rightarrow a$, $B \rightarrow b$, $C \rightarrow c$ and $D \rightarrow d$ are indicated*.

The process of protein synthesis requires that amino-acids be assigned sequentially to codons in genes that are supplied to the system. Which amino-acid is assigned to a particular codon in the process of stringing together amino-acids to form a protein depends on the catalytic capabilities of proteins already in the system. When supplied with the two gene sequences *LLKKKKLKKLLK KLKKKKKKLLLK* a population comprised exclusively of proteins with the sequences *llkkkklkkllk klkkkkkkllllk* corresponding to the assignment activities $K \rightarrow k$ and $L \rightarrow l$ for a binary code can clearly maintain its absolute purity because no $K \rightarrow l$ or $L \rightarrow k$ assignments can be made in a system lacking proteins with those catalytic capabilities.

Copyrighted Material

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/268191/9780262291392_cai.pdf by guest on 01 October 2023

Error-proneness is introduced into the model system by allowing that sequences near to optimal centers for each assignment function are assigned relative catalytic capability R that diminishes rapidly with Hamming distance h from that center according to a Gaussian profile

$$R = \exp(-h^2 / h_0^2) \qquad (1)$$

where h_0 is a constant. This also means that a protein with a randomly chosen sequence is most likely to catalyze all assignment functions in the set $\{K, L\} \rightarrow \{k, l\}$ at an equally minimal rate. Coding self-organization from an initially random population of minimally active proteins in a system of this sort has been extensively documented (Wills 1993).

The second step in the construction of the system is to assume that the classes of codons $\{K, L\}$ and amino-acids $\{k, l\}$ are decomposable into more refined categories, specified as $\{A, B\} \equiv K$ and $\{C, D\} \equiv L$ for codons, and $\{a, b\} \equiv k$ and $\{c, d\} \equiv l$ for amino-acids, generating 16 refined assignment functions, $\{A, B, C, D\} \rightarrow \{a, b, c, d\}$. In keeping with the supposition that a and b are alternative forms of the amino-acid of sort k, and in like manner for all of the decomposed codons and amino-acids, we should expect that the optimal sequence for catalysis of any one of the functions $\{A, B\} \rightarrow \{a, b\}$ will be a specialized example of *llkkkklkkllk* that is optimal for the assignment $K \rightarrow k$. So, the optimal protein sequences for catalysis of assignments $\{A, B\} \rightarrow \{a, b\}$ were chosen by randomly substituting a or b for k and c or d for l in the optimal $\{k, l\}$-level sequence for $K \rightarrow k$ assignment catalysis. The arbitrary $\{a, b, c, d\}$-level sequences generated in this way are shown in the fourth column of Table 1. Suboptimal catalytic activities of sequences were calculated from the Gaussian profile described above with $h_0 = 0.7$ and with h defined in terms of digital representations $a \equiv 00$, $b \equiv 01$, $c \equiv 10$ and $d \equiv 11$.

Results

Simulations of protein synthesis involving the quarternary assignments $\{A, B, C, D\} \rightarrow \{a, b, c, d\}$ were carried out using the methods described in Wills (1993). At each timestep each gene in the system was used as a template for translation to produce a new protein which, when added to the population, contributed to assignment catalysis according to its capabilities as defined by its Hamming distance from each of the optimal sequences (Eq. 1). The genes supplied to the system had sequences *DDBBABCAADCA CDABABDBADCA BCBAABBBCCDB ACABBABADDCA*. When translated according to the assignments $A \rightarrow a$, $B \rightarrow b$, $C \rightarrow c$ and $D \rightarrow d$ constituting the quarternary code, these genes generate proteins with sequences *ddbbabcaadca cdababdbadca bcbaabbbccdb acabbbabaddca* that are optimal for those same coding assignments. Note that these genetic sequences are refinements of the $\{K, L\}$-level genetic sequences needed for elementary coding self-organization, as discussed above.

In an illustrative example, the system started with 3000 proteins whose sequences had been chosen at random. It underwent well-separated transitions after about 3000 and 10000 generations. The binary code first evolved from the initially random assignments of codons to amino-acids (Figure 1). In the initial situation all possible assignments $\{A, B, C, D\} \rightarrow \{a, b, c, d\}$ were carried out at equal rates and random proteins were synthesized. Specified coding assignments, $\{A, B\} \rightarrow \{a, b\}$ and $\{C, D\} \rightarrow \{c, d\}$ representing $K \rightarrow k$ and $L \rightarrow l$, became selectively amplified as a result of the reflexive information supplied to the system for translation. Correspondingly, components that catalyze non-coding assignments, $K \rightarrow l$ and $L \rightarrow k$, initially equally represented, effectively disappeared from the system.

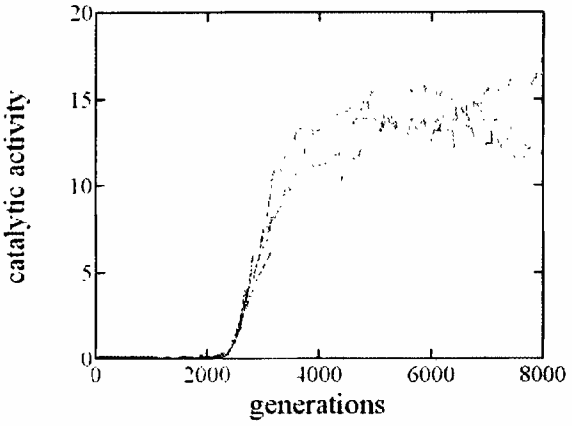

Figure 1. Transition to a binary code. The 8 quarternary assignment functions $\{A, B\} \rightarrow \{a, b\}$ and $\{C, D\} \rightarrow \{c, d\}$, equivalent to the two binary coding assignments $K \rightarrow k$ and $L \rightarrow l$, are selected from the complete set of 16 possible assignments $\{A, B, C, D\} \rightarrow \{a, b, c, d\}$. Each of the traces represents the net rate at which one codon-to-amino-acid assignment $X \rightarrow y$ is catalyzed in the system.

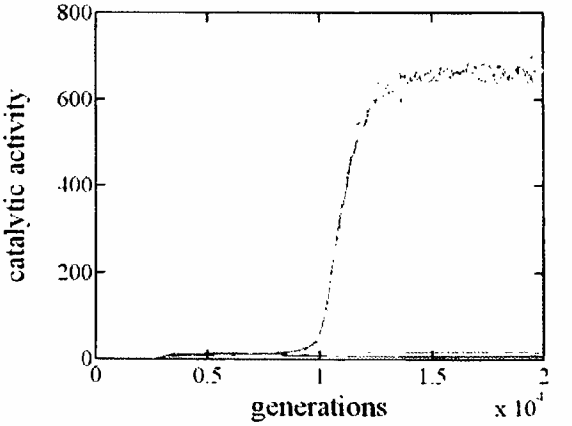

Figure 2. Further transition to a quarternary code. After the first transition to a binary code, the 4 assignment functions $A \rightarrow a$, $B \rightarrow b$, $C \rightarrow c$ and $D \rightarrow d$ are selected from the binary coding sets $\{A, B\} \rightarrow \{a, b\}$ and $\{C, D\} \rightarrow \{c, d\}$. Each of the traces represents the net rate at which one translation assignment $X \rightarrow y$ is catalyzed in the system.

A second stage of evolution resulted in selection of the quarternary code, based on the refined assignments of

Copyrighted Material

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/268191/9780262291392_cai.pdf by guest on 01 October 2023

codons from decomposed sets, $\{A, B\} \equiv K$ and $\{C, D\} \equiv L$ to amino-acids from subclasses $\{a, b\} \equiv k$ and $\{c, d\} \equiv l$ (Figure 2). Specified coding assignments, $A \rightarrow a$, $B \rightarrow b$, $C \rightarrow c$ and $D \rightarrow d$, became selectively amplified and the concentration of components that catalyze the other four quarternary assignments, $A \rightarrow b$, $B \rightarrow a$, $C \rightarrow d$ and $D \rightarrow c$, still represented in the coarse-grain binary code, were reduced to a relative minimum.

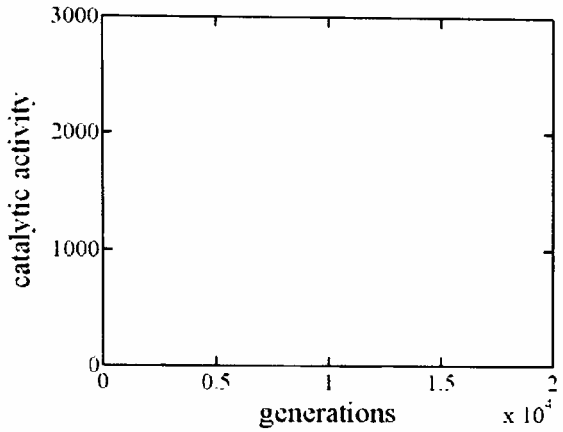

Figure 3. Selective driving force. During each coding selection transition there is a radical increase in total catalytic activity, a measure of the general rate at which proteins are synthesized.

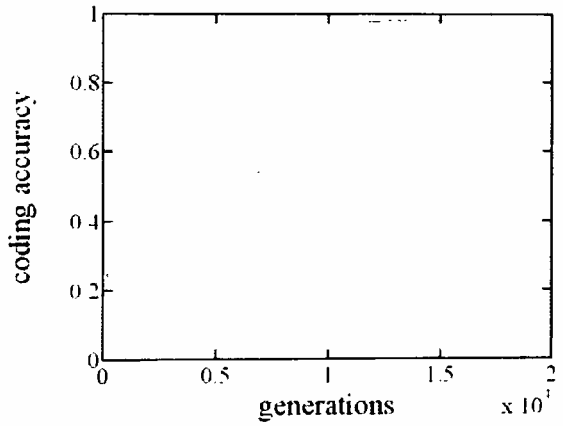

Figure 4. Coding selection. The accuracy with which the final code $A \rightarrow a$, $B \rightarrow b$, $C \rightarrow c$ and $D \rightarrow d$ is executed undergoes two transitions. The initial probability of approximately 0.25 corresponds to the random assignment of codons to amino-acids, the intermediate probability of about 0.5 corresponds to the binary code $\{A, B\} \rightarrow \{a, b\}$ and $\{C, D\} \rightarrow \{c, d\}$ and the final probability close to 1.0 corresponds to the fully-fledged quarternary code.

The two transitions in the dynamics of the system are evident in further results. Figure 3 displays what is in effect the thermodynamic driving force behind the transitions, the improvement in the overall rate of assignment catalysis derived from the selection of catalytically active components that mutually stimulate protein synthesis. Figure 4 demonstrates evolution from initially random protein synthesis, through the partially resolved binary stage to the quarternary code. In the final stationary state most of the proteins in the system are highly active for the four assignments that constitute the full code.

Discussion

One of the criticisms voiced in the community of researchers interested in the origin of life is that the systems studied are artificially constructed so as to display exactly the behavior being sought as an indication of evolutionary self-organization. Thus, it is argued, rather than enlightening us as to how life actually evolved out of disordered chemical events and processes, we are left with no more than a description of how a system behaves when it is designed to behave that way, the biochemical equivalent of a tautology. Such criticism must be taken seriously when assessing the significance of hypothetical prebiotic scenarios. It is also true that ALife systems are completely hypothetical constructs conforming to preconceived design principles. However, the description and study of ALife systems can demonstrate quite generally that certain preconditions must necessarily be fulfilled before life can evolve. The principles of reflexivity and decomposition describe rather obvious preconditions of that sort, but they have not been generally applied to our understanding of the accumulation of biological information and its exploitation in the precise specification of biochemical function. This is because it is very difficult, after the event of evolution, to reconstruct either the context within which information has promoted the circularity of processes needed for the maintenance of autocatalysis or the classes of primitive processes that were subsequently refined as biochemical complexity increased. Further discouragement from seeking explanations for the origin of biological complexity comes from the confident neo-Darwinian retort "Selection takes care of all of that!" The retort may be true, but it is not complete.

The principle of reflexivity is significant because the complexity and specificity of the most fundamental molecular biological structures and functions is such that their sudden appearance out of a morass of undifferentiated chemical processes and events defies simplistic explanations. As has been shown elsewhere (Nieselt-Struwe and Wills 1997; Wills and Henderson 2000) the structure-function relationship of catalysts must be such as to contain the possibility of reflexive closure between components and operations in the first place. Of course it is important that there exist components that carry out selected operations, but it is also important that the functional specificity of the generated population of inexactly synthesized components keeps the system above the error-catastrophe threshold. Whether this is possible depends in the end on how operational function varies with structure.

From a design point of view, the structure-function relationship of catalytic polymers is created by embedding

Copyrighted Material

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/268191/9780262291392_cai.pdf by guest on 01 October 2023

catalytic capabilities in the relevant polymer sequence space, as in Table 1. It has been possible to design systems that support the stable execution of molecular biological translation (Wills, 1993) by ensuring that the embedding of assignment catalysts in the protein sequence space is consistent with the possibility of constructing genetic sequences that are reflexive *vis-á-vis* the production of a coding set of assignment catalysts *via* translation. In the present paper it is shown how this basic principle can be extended to produce nested realizations of reflexivity that support a gradual pathway for the development of complex structures and functions. The principle of decomposition is more subtle than its parent principle (Wills 2001) because it relies on the description of components of the system at different levels: "coarse-grained" like $\{K, L\}$ and $\{k, l\}$, "fine-grained" like $\{A, B, C, D\}$ and $\{a, b, c, d\}$, *etc.* In ALife we have the luxury of defining hierarchical levels *a prior*, but it is unclear how they can be chosen adequately to describe natural biological systems. The formalized principle of decomposition may have more application to the design of artificial systems than the description of living systems found in nature. Its general significance is apparent in the work of Watson and Pollack (). It remains to be seen whether systems constructed according to the principle of decomposition enlighten intermediate stages of evolution when coding self-organization is coupled to the error-prone replication of the required genetic information. The study of Füchslin and McCaskill (2001) could be extended in this direction.

A philosophical note is appropriate in closing. The Central Dogma of molecular biology (Crick 1957) specifies that molecular sequence information can flow between nucleic acids and into protein but "once it is in protein it cannot get out again" (Crick 1970). The principle of reflexivity complements the Central Dogma by requiring that the information specifying an organism be adapted to the structure-function relationship of proteins such that coordinated execution of the informationally specified functions of proteins produces the very proteins that carry out those functions. In many ways the non-algorithmic embedding of biological functions in the folded structure of proteins is a more characteristic feature of biological systems than the one-way flow of sequence information from DNA to RNA to protein. Algorithmic information flow, translation essentially, guarantees the stability of the processes of inheritance and the operation of natural selection. On the other hand, the non-algorithmic embedding of function within structure space can be exploited to produce the basic phenomena of life, replication, variation and selection, without informational encoding, as the existence of prions demonstrates (Wills 1988). Further articulation and discussion of these ideas is likely to contribute significantly to our understanding of bioinformatics as we attempt to build models of the relationship between complex assemblages of genes and the characteristics of the organisms to which they belong.

Acknowledgments. Support from the Alexander von Humboldt Foundation and helpful advice and encouragement from John McCaskill are appreciated. It is the author's wish that no person or agency should ever derive military or purely commercial benefit that depends in any way on the publication of this paper.

References

Adami, C., Ofria, C. & Collier, T. C. 2000. Evolution of Biological Complexity. Proc. Natl. Acad. Sci. USA, 97:4463-4468

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

Crick, F. H. C. 1958. On protein synthesis. Symp. Soc. exp. Biol. 12:138-163

Crick, F. H. C. 1970. Central Dogma of Molecular Biology. Nature 227:561-563

Eigen, M. 1971. Selforganization of Matter and the Evolution of Biological Macromolecules. Naturwiss. 58:465-523

Füchslin, R. M. & McCaskill, J. S. 2001. Evolutionary self-organization of cell-free genetic coding. Proc. Natl. Acad. Sci. USA, 98:9185-9190

Hoffmann, G. W. 1974. On the Origin of the Genetic Code and the Stability of the Translation Apparatus. J. Mol. Biol. 86:349-362

Hornos, J. E. M. & Hornos, Y. M. M. 1993. Algebraic model for the evolution of the genetic code. Phys. Rev. Lett. 71:4401-4404

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

Orgel, L. 1963. The Maintenance of the Accuracy of Protein Synthesis and its Relevance to Ageing. Proc. Natl. Acad. Sci. USA, 49:517-521

Watson, R. A. & Pollack, J. B. 2001. Symbiotic Composition and Evolvability. In Proceedings of Advances in Artificial Life, 6th European Conference, (ECAL 2001), Lecture Notes in Computer Science 2159, 480-490 (Springer)

Wills, P. R. 1988. Genetic Information and the Determination of Functional Organization in Biological Systems. Systems Research 6:219-226

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

Wills, P. R. 2001. Autocatalysis, Information and Coding. BioSystems 60:49-57

Wills, P. R. & Henderson, L. 2000. Self-organisation and information-carrying capacity of collectively autocatalytic sets of polymers: ligation systems. In Unifying Themes in Complex Systems: Proceedings of the First International Conference on Complex Systems, 613-623 (Perseus Books)

Copyrighted Material

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/268191/9780262291392_cai.pdf by guest on 01 October 2023