Thermal Proteins in the First Life and in the "Mind-Body" Problem

Sidney W. Fox

1 Introduction

Our general understanding is that biological information first arose in pre-biological molecules. The two kinds of macromolecule that are suitable candidates for initial bioinformation are nucleic acids and proteins, since these are the informational molecules of present-day biological systems. A logical case with some qualifications can be made *a priori* for each of these types of macromolecule as the original source of information (Schmitt, 1962; Lehninger, 1975). It is nevertheless true that special difficulties arise when one seeks to understand how either DNA or RNA could originally have served not only as a source of information but also as processing agents for that information, e.g., to synthesize nucleic acids and proteins, and cellular systems with or without template. Also, no one has explained satisfactorily by experiment how DNA or RNA could have arisen without prior protein (Fox, 1988, Waldrop, 1989).

Proteins are easily understood as processing agents because they serve such functions so richly in modern systems, and those abilities are found in isolated systems in the laboratory. A main question has been, rather, how proteins could

function initially also as a source of information conceptually as well as in processing agents, e.g., enzymes. When looked at in a strictly evolutionary way, one could test the origin of information in proteins from precursors of proteins, amino acids. The origin of information from nucleic acids, which are indeed storage molecules for information in modern systems, has not had comprehensive testing. Proteins have been thoroughly tested in this role (Fox, 1988) although not as the result of the above line of reasoning. The total results indicated that one kind of protein (thermal protein; protoprotein) is able to function as various processing agents: enzymes, inhibitors, direct precursors of cells, etc. and also as an adequate initial source of information.

A pragmatic way of looking at the alternative possibilities is to recognise that in the period 1960 to 1990 the route of protein-first has yielded a complete theory of the transition from inanimate matter to polybiofunctional cells, whereas in the same period nothing comparable has emerged from experiments with nucleic acids of any kind, especially not with primordial types. In order to explain an initial DNA Crick (1981), for example, has invoked a catalytic mineral on an

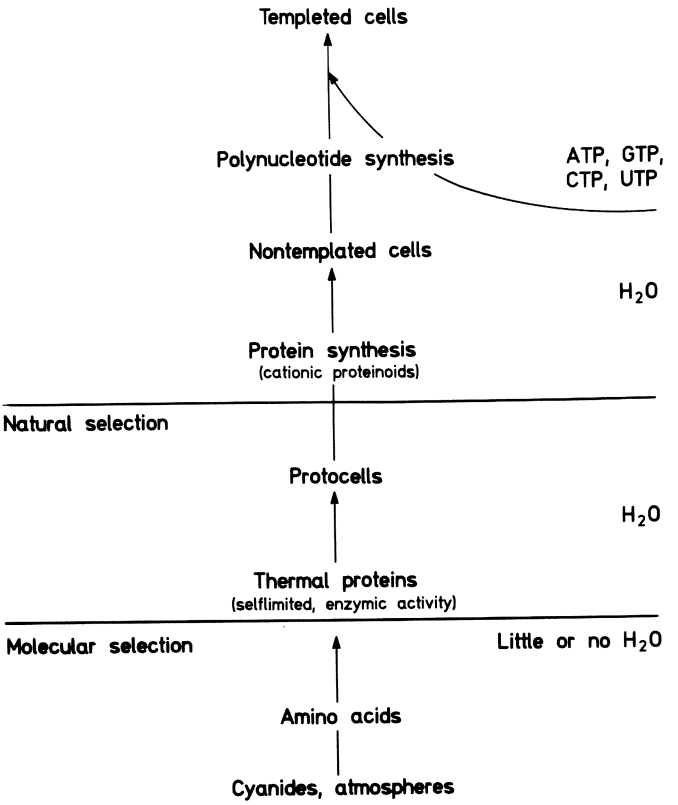


Fig. 1. Flowsheet of the proteinoid (thermal protein) emergence of life leading into its evolution. The self-ordering of amino acids (molecular selection) has made it possible

unnamed planet. While a modern RNA has been shown to have a kind of catalytic self-repair activity (Cech, 1985), there has been no support for the origin of metabolism from RNA broadly speaking, no satisfactory explanation for how the original RNA arose, no explanation for how cells appeared and no broad explanation for how various RNAs could have served as processing agents for information other than storage (to which, however, nucleic acids are well adapted because of their stability). The flowsheet from proteins-first is in Fig. 1.

The reason for attempting to make a protein before producing a cell was initially simply the attempt to understand how protein *per se* might have come into existence, cell or no cell. The answer to how the cell could have arisen thrust itself upon the experimental scene. When water entered the flasks containing thermal protein, as it almost inevitably did, the thermal protein *organized itself* into cells. The self-organization of thermal protein was one of two essential advances that had to precede a laboratory protoorganism with qualities of life as modern biologists list them.

2 Essential Molecular Advances

The two chemical advances were:

1. *The self-sequencing of amino acids to thermal proteins.* Some chemists recognised that the composition of mixtures of amino acids heated to yield thermal proteins would necessarily differ from the compositions of the resultant proteins. They were however not prepared for the idea that the sequences in the polymers might be ordered by the compositions in the mixture. Three reasons for delay in thinking can be given:
   (a) there is no sequence in a mixture of free amino acids to compare with,
   (b) the thinking is preconditioned by knowledge of modern systems, which rely on nucleic acids, and
   (c) the sequences in the polymers formed had first to be analysed.
2. *The self-organization of thermal proteins to cellular structures.* In his lecture at the Sorbonne in 1864, Pasteur had asked, "Can matter organize itself? In other words are there beings that can come into the world without parents, without ancestors? That is the question to be resolved."

In 1956, Schmitt reported the self-organization of precipitating collagen to form microfibrils. In 1959, quite independently, our laboratory (Fox et al., 1959) reported the ready conversion, by contact with water, of thermal protein to cellular structures. This tendency is more characteristic of laboratory protoprotein than of modern protein, probably because of (limited) branching in the thermal protein molecules (Harada and Fox, 1975).

The production of cellular structures from thermal proteins by steps that were themselves understood allowed determination of the properties of such units

through the next 20 or more years (Fox, 1981, 1989). Since the cell-like structure possesses properties of metabolism, growth, reproduction, evolvability, and many other biofunctions it deserves to be regarded as a synthetic *proto*organism. Demanding special attention was the finding that growth of a cell from protein is quite strictly programmed in size, and that this limitation is not mediated by nucleic acids. Once that is recognised, one can understand that the formation of buds, a necessary step in primitive reproduction, would be a continuation from parental growth.

The significance of self-organization for the construction of flowsheets beginning with the origin of cellular life was recognised early (Fox, 1960a, b) and even explained by request to scientists associated with the Vatican (Fox, 1985). That a greater significance of self-organization in wider vistas of biology and computerology would arise was not foreseen.

The synthetic production of a protoorganism from noncellular precursors validates the conceptual flowsheet representing the route from amino acids to the first organism (Young, 1984). The amino acids are much like those identified as precursors in the Apollo program, especially the dicarboxylic amino acids that enable polycondensation (Fox, 1960b). Because of the experimentally established properties of this organism, as determined in the laboratory between 1960 and 1980 and beyond, no other experimental or theoretical flowsheet has been proposed. Nor has any other sequence of studies recited the need for self-organization, although Eigen (1971) suggested it for RNA. No self-organization of RNA has been demonstrated.

The initial information base for the reaction sequence in the flowsheet is provided by the method of self-sequencing of amino acids. The self-sequencing mechanism is the one that Ivanov and Förtsch (1986) have inferred as being "universal" for modern organisms. They have reached this conclusion from study of 2898 protein sequences in the databank at the Max Planck Institute for Biochemistry in Martinsried, in association with the late Professor Gerhard Braunitzer. In this connection, it should be noted that the starting point for the entire study that developed into an investigation on the emergence of life was out of the laboratory efforts that are credited with stimulating the present era of amino acid sequence determination (Fox, 1945) before Sanger's work (history in Rosmus and Deyl, 1972; Oparin and Rohlfing, 1972, and others). That line of study was aimed at elucidating the informational, i.e., bioinformational, beginning of evolution based on proteins. Ivanov and Förtsch used data *derived* by statistical analysis from amino acid sequence data. The study has thus come full circle methodologically and this fact can introduce a unity of thinking into the larger view. It supports the existence of biospecificity derived from varied amino acid sequences, by an integrated approach of both chemical analysis and synthesis, leavened with evolution.

3 Premises

The premises of information processing in the evolutionary context include the assumption that all molecules contain information. Molecules possess this information due to their individual stereoelectronic configurations. Each molecular type has its own shape, including those in some arrays of tautomers.

The informed system is capable of making choices because it possesses information. The conversion of one type of molecular system to another is information processing. We recognise the result of interaction of informed macromolecules with other molecules, other macromolecules, or with other systems as manifestations of making choices.

Although premises of randomness or determinism could not at the outset of this research in the 1940s and 1950s be analysed as clearly as they can now be, they provided contexts to include randomism, near-randomism, a state approximately equidistant between randomism and determinism, near-determinism, and determinism. The premise of randomism in evolutionary beginnings, for example, was assumed by Eigen (1971), although he has more recently departed somewhat from that position (Eigen, 1986) on the basis of theoretical analysis.

More emphatically, Tyagi and Ponnamperuma (1990b) have studied the formation of polymers from amino acyl adenylates which Krampitz and Fox (1969) showed reacted selectively (Tyagi and Ponnamperuma, 1990a). Ponnamperuma's study of the condensation compared the selectivity in polymerization of amino acid residues with that of polynucleotides from the same starting compounds, various amino acyl nucleotide anhydrides. All of the specificity was found to be in the amino acids, without influence from the nucleotide residues in the same molecules. Thus, amino acids select their own sequences whether polymerized by heat, by reaction through the activated adenylates, guanylates etc., or via the more complex modern protein synthesis mechanism (Ivanov and Förtsch, 1986).

Since his graduate student days (1938 to 1940), the author of this paper has believed in genetic determinism (Wilson, 1978), due especially to discussions with the chairman of his Ph.D. committee, Thomas Hunt Morgan. The premise of genetic determinism has now been extended forward from molecular determinism, a view somewhat easily arrived at by a chemist who is concerned with transferring the high degree of repeatability of chemical reactions into a more philosophical, evolutionary context. The most recent synthesis of these views has led to the suggestion of a "daisy-chain" of (a high degree of) determinism throughout the hierarchy and throughout the evolutionary sequence (Fox and Nakashima, 1984).

As a result of this deterministic context, one may not speak strictly of the "origin of the genetic code", but rather of the "origin of the genetic coding mechanism". The genetic code is an expression of the selective interactions of nitrogen bases in triplets for amino acid residues; this attraction is inherent in the molecules. The evolutionary sequences developed in this study are expressed,

in fact, more fully in the sequence of molecular structures than in the processes; this is believed to be due to the fact that the processes are derived from the molecular structures.

Evolutionary Continuity. The concepts of evolution underlying conversion of information gain validity as they develop into a lengthening sequence of processes having connections that have been disciplined by testing, especially experimental demonstration. The basic assumption is that evolution of systems or processes must be stepwise as in a staircase.

4 The Self-Sequencing of Amino Acids As the Initial Bioinformational Process

Biological molecules sufficiently large to contain the amount of information needed for processes in organisms appear to be nucleic acids (DNA and RNA), proteins, and polysaccharides. Polysaccharides are readily eliminated from consideration because they tend each to be composed of a single type of monomer, e.g., glucose (in starch or cellulose) or mannose (in mannans). The essential conditions for informational nucleic acids and informational proteins is the variety of information derived from variety of monomer types in polymer chains.

Since proteins and nucleic acids are polymer types that contain monomers in variety and do not resemble monotonous chains, they are richly informational. Lehninger (1975) has treated both nucleic acids and proteins as informational molecules although it is possible to hear or read treatments by others in which nucleic acids seem to be the only informational biomacromolecules, by a kind of default of consideration. In some of the thinking in this field, in fact, the focus of attention on DNA or RNA is so intense that proteins seem to have been tacitly disqualified by being sloughed off.

A demonstrable model for sequence formation of nucleotide monomers in polynucleotides without prior processing by already informed protein has not been provided. Such fundamental questions as how DNA and RNA came into existence without prior informed protein, how such nucleic acids were informed, how they produced cells, how they were translated into protein, and many other questions have not been answered experimentally. Nor have they been answered theoretically. On the other hand, the emergence of informed protein has been explained experimentally, as has the emergence of informed nucleic acids as a result of the action of informed protein. The answer stemming from informed protein has developed many interdigitated spinoffs, and meets the requirement of evolutionary continuity – in an extended series.

Two kinds of experiment support the rigour of the protein-first perspective for information. One is the structural nonrandomness. The other is the finding of biofunctions in many of the thermal proteins: antiaging of neurons when treated in culture (Hefti et al., 1991), memory enhancement when injected into

mice (Fox and Flood, 1991), etc. Random polymers could not be expected to yield biofunctions and to yield them repeatably (Eigen, 1971b).

Among the biofunctional results from the proteins-first investigations are: an array of enzymatic activities, specificities of interactions between thermal proteins and substrates, linking of enzymatic events to provide metabolic chains, self-assembled cells, experiments indicating how an initial genetic coding mechanism could have arisen, the onset of bioelectricity, comprehensive membrane properties, reproduction, assemblies of cells, and others (Fox and Dose, 1977; Fox, 1980; Fox, 1988).

Related to the primary question of which kind of informational biomacromolecule, protein or nucleic acid, arose first, is the question of whence came that information? The one answer, which has several significances, is that informational content for the first protein molecules was derived from the amino acids that reacted to form thermal proteins. This follows from the finding that amino acids can be reacted to form proteins directly (Fox and Harada, 1958) (Fig. 2).

It is of interest that Fritz Lipmann (Kleinkauf et al., 1988) found he could produce a presumably primitive polypeptide without direct involvement of DNA or RNA; only protein and amino acids were used. Incidentally, Lipmann several times (1972, 1984) acknowledged motivation from the present program for his active interest in his studies, which became a principal interest in his latest years.

Lipmann, however, attributed the directive effects on amino acids in large part to a protein template that functioned instead of the nucleic acid template usually

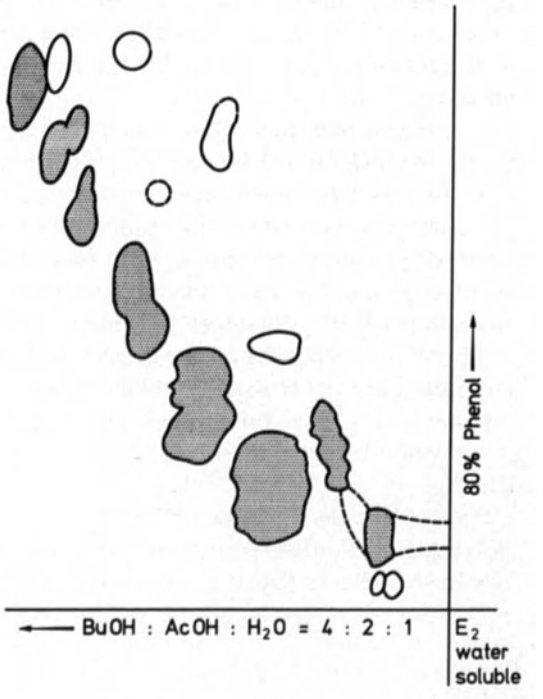


Fig. 2. Chromatogram of hydrolyzed sample of polymer of the common amino acids

invoked. If he had looked further into the mechanism, Lipmann might have as well or better attributed the main "patternization" effects to the self-sequencing of the free amino acids, much as Ivanov and Förtsch (1986) later did.

If protein and nucleic acid template were each superfluous, as the experiments suggest, how did amino acids line up if there was no template for them to line up on? The answer, as has already been mentioned, is that as many as twenty kinds of amino acid order, or sequence, themselves into a line. This mechanism is an extremely economical one, and seems most appropriate to prebiotic times. It can be best understood on the basis that the family of amino acids consists of a remarkable group of siblings, to which there is nothing comparable in the biochemical world, and the further inference that there was nothing comparable in the protobiochemical world. The amino acids are all related in having both amino and carboxyl groups (which makes polycondensation possible) but in being varied by having different side chains and different configurations of electrons. Their condensation results in predominantly linear molecules, while the stereoelectronic effects of the side chains builds in specificities of interaction in the growing polymer chain.

For the self-sequencing of amino acids there are several kinds of evidence (Fox and Nakashima, 1984); the best is undoubtedly that obtained by sequence analysis of peptides formed (work from which this entire project began; Fox, 1945, 1956). They include the difference in composition between the reacting amino acid mixture and the composition of the polymers obtained, sharply limited heterogeneity in the components of the polymers obtained by fractionation, large differences between compositions of the polymers and the compositions of the N-terminal and the C-terminal portions, and nonrandom sequences in the components as determined by standard methods (Fox, 1945; Fox and Dose, 1977).

The results of fractionation of such polymers on DEAE-cellulose as compared to fractionation on DEAE-cellulose of turtle serum proteins are seen in Fig. 3.

In Fig. 4 a two-dimensional chromatogram of the product of heating of glu, gly, and tyr is presented. This result was surprising in that the individual spots (stained by α-nitrosonaphthol) were found to be as discrete as they are. More surprising was that each one of them showed upon hydrolysis and analysis stoichiometric ratios between the contained amino acids.

Spot 3–3 contains by far the largest amount of any of the peptides formed. It contains all of the tripeptides (Nakashima et al., 1977) except for a very small amount of <glu-tyr-tyr subsequently found in Martinsried (1988). Otherwise 3–3 is only two *tri*peptides: <glu-tyr-gly and <glu-gly-tyr (<signifies pyro or cyclo-).

In Table 1 the peptides obtained are listed, and also those expected if the polycondensation were random, as would usually be anticipated. The results obviously indicate a great nonrandomness. Dr. John Jungck calculated the ratio nonrandomness/randomness to be 19/1 (Nakashima et al., 1977). *This was pro-*

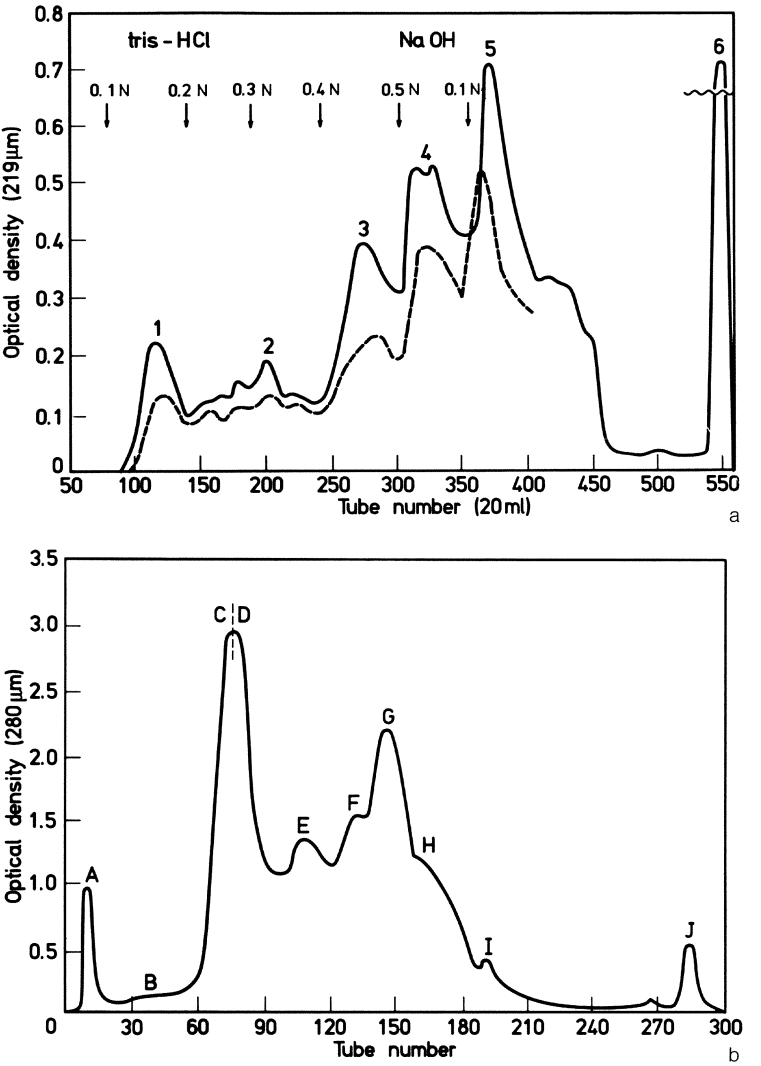


Fig. 3a. Fractionation of amidated 1:1:1-proteinoid on DEAE-cellulose column. A few major fractions are observed. Repetitions yield similar patterns. The dotted line represents one of these; **b** Distribution of turtle serum proteins on DEAE-SF-cellulose column in sodium phosphate buffer (Block and Keller, 1960, in Fox and Nakashima, 1984)

bably a first quantitative evaluation of nonrandomness. Since no template is present when the polycondensation begins, we infer that the reactant amino acids have constructed their own sequences. This is to be understood on the basis that each of the twenty kinds of amino acid has its own steric and electronic identity, as stated.

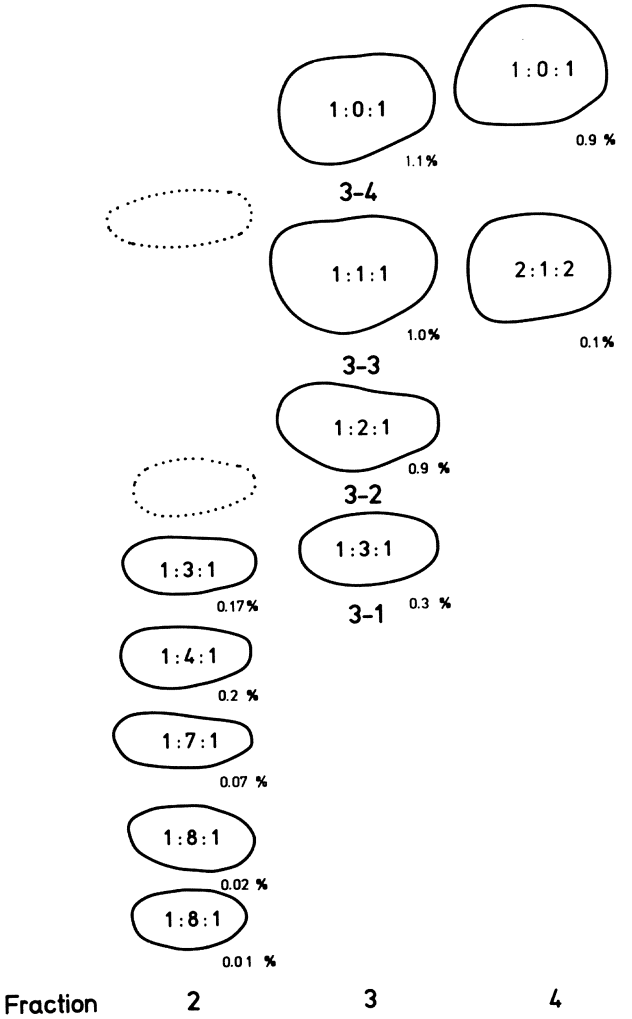


Fig. 4. Discrete peptides by paper chromatography of thermal product of glutamic acid, glycine, and tyrosine (Nakashima et al. 1977; Hartmann et al. 1981). Dominant fraction is 3–3; it represents all tripeptides formed

5 The Mechanism of Self-Sequencing of Amino Acids

The fundamental answer to the question of the outset of evolution of information and its processing is, as already noted, the self-sequencing of amino acids. The answer to the question of the necessary mechanism has been partially given, but it is not complete. What has been learned is that pyroglutamic acid, which results

Table 1. Tyrosine-containing tripeptides found vs. those expected on the basis of the random hypothesis

Expected from *Random* Polymerization		*Found* from *Nonrandom* Polymerization
αUαUY	YαUU	
αUγUY	YγUU	
γUαUY	YαUG	
γUγUY	YγUG	
αUGY	YαUY	
γUGY	YγUY	
αUYU	PαUY	
γUYU	PγUY	
αUYG	PGY	*PGY*
γUYG	PYU	
αUYY	PYG	*PYG*
γUYY	PYY	
GαUY	YGU	
GγUY	YGG	
GGY	YGY	
GYU	YYU	
GYG	YYG	
GYY	YYY	

Note: The dominant fraction obtained from the thermal copolymerization of glutamic acid, glycine, and tyrosine proved to be an equimolar complex of pyroglutamylglycyltyrosine and pyroglutamyltyrosylglycine (Nakashima et al., 1977; Hartmann et al., 1981) U = glutamic acid residue, Y = tyrosine residue, G = glycine residue, P = N-pyroglutamyl from unfractionated peptides, each had but a single amino acid in each terminal residue. In the N-terminal position, the only amino acid was pyroglutamic acid. Three individual peptides could, of course, have no more than three C-termini in total. Three of the six conceptual possibilities were thus found. Each C-terminus was singular: glycine, alanine, or leucine. While multiple C-terminal types are found analytically in unfractionated proteinoids (Harada and Fox, 1975), the singularity suggests three single peptides. Proline and phenylalanine were totally absent from the C-terminal and N-terminal positions. Glutamic acid, which in the pyro form is known as an N→C polymerization initiator (Fox 1980), is totally absent from all three C-termini

easily by warming glutamic acid, is an excellent N→C polymerization initiator (Melius and Sheng, 1975; Fox, 1980). The fact that N-pyroglutamic acid starts one or more, usually more, amino acid sequences results in induced internal arrangements in those sequences.

The main aspects of what was been learned about mechanism are that N-pyroglutamic acid is a polymerization initiator, that aspartic acid and glutamic acid are elongators, and tyrosine is a terminator. Since it is increasingly clear that protein alone could have served as a bridge from prebiotic matter to the first animate systems, the question of how that bridging protein received, or developed, its own internal information becomes extremely fundamental to understanding the emergence of life and its initial information content.

In pursuit of the answer to this question, the P.I. spent April to September 1988 in the Protein Department of Director Dr. Gerhard Braunitzer at the Max Planck Institute for Biochemistry, Martinsried, Germany. Braunitzer (1984) (who died on 27 May, 1989) had determined the primary sequences in haemoglobins from over 200 species. Braunitzer's deep interest in the history of science led him to surround himself with those whom he identified with the "ursprung" (origin) of his studies. For example, he made labs available for Pehr Edman of the Edman Sequenator. These labs were used by Edman until his death, and then by Dr. Agnes Henschen, widow and scientific colleague of both Edman and Braunitzer until the death of the latter. In his invitation to the P.I.–author of this paper he was acknowledging, as he told visitors, pioneering work on sequence determination (Fox, 1945, et seq.). In the course of this association, Braunitzer became further interested in the evolutionary origins of protein, having already aided work on the inheritance of the self-ordering mechanism by Ivano and Förtsch (*Origins Life* 17: 35–49, 1986), in which his suggestions and support are acknowledged.

Braunitzer and the P.I., with Dr. Peter Rücknagel, established or confirmed that:

1. The thermal polycondensation of amino acids is highly nonrandom.
2. The initial steps of thermal protein synthesis reproducibly yield a few component macromolecules.
3. The synthesis of larger molecules evidently involves a kind of "springboard" mechanism in which small and quite large peptides are more favoured than intermediate-sized ones.
4. Pyroglutamic acid is a frequent feature of thermal proteins, and also of a number of modern proteins.
5. Use of Melius' techniques (Melius and Sheng, 1975) to convert pyroglutamic acid residues to glutamic acid residues has made sequential studies a possibility through at least six residues (incomplete experiments in Martinsried).
6. As a step in evolution, the patterns obtained suggest diversity within near-unity rather than unity within diversity.

6 Future Studies of Sequence

1. Survey of modern N-pyroglutamyl proteins and peptides. Approximately 125 were identified in August 1988 at MIPS (Martinsried Institute for Protein Sequences).
2. Compositional values of amino acid types in the 2nd, 3rd, 4th, and 5th positions should be calculated.
3. Glutamic acid to be heated with various combinations suggested in the preceding evaluation, and under varied conditions. While the laboratory

model suggested the computer work that has been done, the computer studies may now lead to ongoing laboratory models.

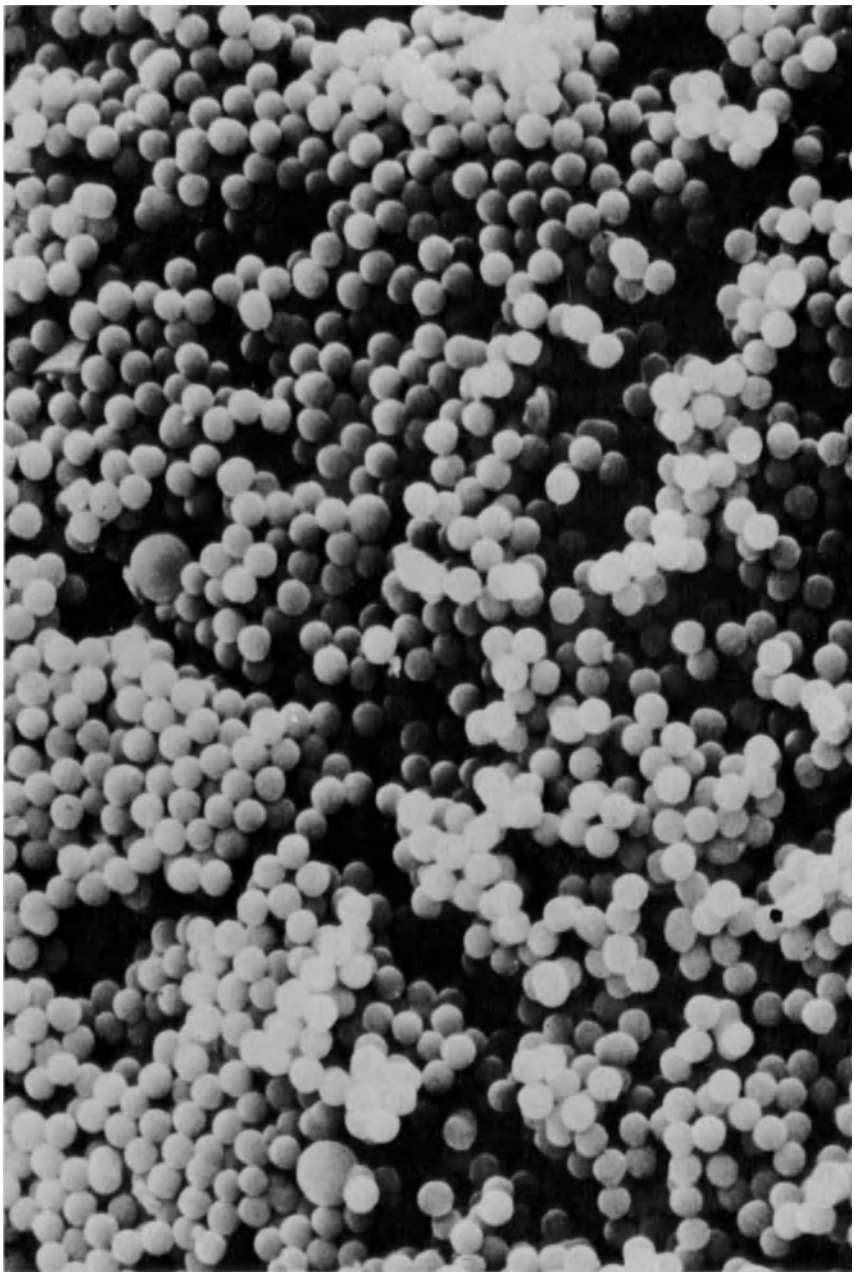
4. Relating HPLC analyses of polyamino acids to oscilloscope patterns of proteinoid microspheres assembled from such polymers. This advance converts molecular electronics from a primarily theoretical exercise to an experimental avenue.
5. Spinoffs are: potential applications to computers more versatile than the present generation, clarification of the protein-cell-body-mind relationships, and therapy that has already begun to appear.

7 Self-Organization of Thermal Proteins to Cells

The spontaneous formation of sequences of amino acids starting with the amino acids themselves is a kind of self-organization (plus bond formation). It is a unique process inasmuch as the properties, especially the reactions, of the amino acids are found in a family of molecules without parallel in nature. The twenty types of amino acid that we know best are sibling molecules. They all bear the family imprint of amino and acid groups; they are distinguished one from the other by their characteristic side chains. While the amino and acid groups participate in the formation of peptide bonds, the contributions of the side chains are responsible for the rate at which any one type of amino acid couples with another or with the growing peptide chain.

The terminology employed with amino acid sequences reflects the dominance of the analytic approach over the synthetic approach in science. Strictly speaking, the sequences determined analytically are amino acid residue sequences. The use of primary analysis on peptides and proteins has come to be spoken of as "sequencing" the peptides or proteins. The verb "sequencing" belongs more properly to synthesis from amino acids, however.

While the amino acids themselves function to yield a spontaneous ordering into peptide chains, the peptides formed manifest their attractions by organizing themselves into cells through intramacromolecular binding. This is a crucial step, discovered in this program (Fox et al., 1959). The attractions of one amino acid for another when they are largely fixed in molecules of larger size can then be expressed by the pulling inward of parts of the peptide molecule. The tendency of polyamino acid molecules to organize themselves thus is believed to be at the heart of the original cell formation (Fox, 1985). This kind of expression of information has not received detailed study; it offers much promise (Fig. 5).



Proteinoid Microspheres

Fig. 5. Scanning electron micrograph of proteinoid microspheres. Note uniformity and huge number. These are nearly all of a narrow diameter size within a range of 1–3 microns. Illustration prepared by Mr. Steven Brooke of Steven Brooke Studios

8 The "Mind-Body" Product

In the beginning of this research, it was not visualized that the flowsheet would enlarge to the emergence of life and mind from, first, the study of the emergence of protein and, later, the study of the emergence of life from proteins. Once again, the *direction* of approach is found to be crucial. When one employs the standard approach of science, that of analysis, one cannot analyse life, and one cannot analyse mind, without dissecting them in a way that causes loss of their essential property of organization. The investigator can only experiment, somewhat as nature has done, to assemble putative components to yield, first the cell, and later, as suggested by that first exercise, the mind.

The "mind-body" problem (Popper and Eccles, 1977; Uttall, 1977) should, accordingly, be referred to as the *body-mind* problem. The "mind-brain" problem, likewise and for the same reason, should be referred to as the *brain-mind* problem. This understanding has been recognised since the early 1970s (Young, 1984); see Purves (1990).

Much as the dicovery that cells will assemble from wetted thermal protein, the properties of the units of the brain were found in the artificial neurons, which are simply in experimentation special properties of the artificial cells.

9 Artificial Neurons

Approximately thirty years of research have gone into cataloguing and understanding the emergence of cells. These are properly referred to, of course, as artificial cells. But they are not simply models. They represent retracement of the

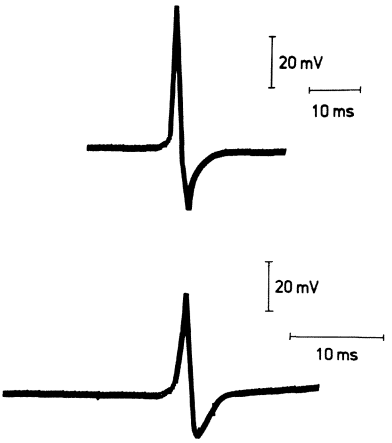


Fig. 6. Action potential resembling that of neuron. (Upper) Spiking in crayfish stretch receptor neuron. (Lower) Spiking in microsphere of 2:2:1-proteinoid. (By Dr. A. Przybylski.)

evolutionary pathway by which more and more modern cells arose. This is reassuring because we know in some respects quite thoroughly what it is that eventuated from evolution.

As indicated, the properties of the artificial cells have been rather thoroughly catalogued. The electrical properties, coupled with the connectional properties

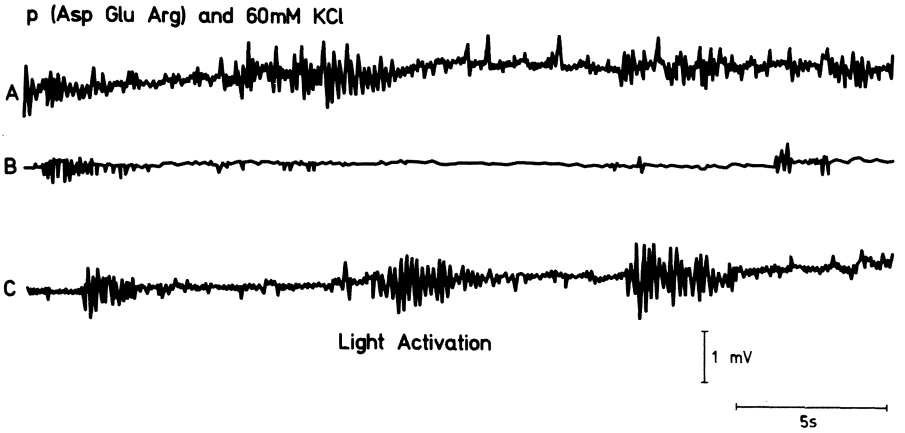


Fig. 7 A–C. Electrical activity begins with illumination (10 lux) at **A**. It dies off when light is extinguished (**B**), and it begins again with reillumination (100 lux) at **C**

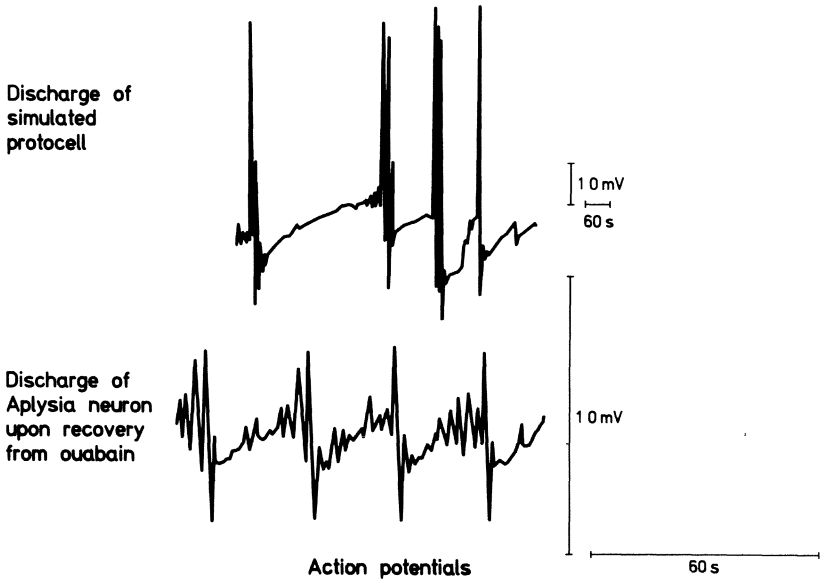


Fig. 8. Pattern of spontaneous electrical discharges of the 2:2:1 proteinoid-lecithin cell and of *Aplysia* neuron

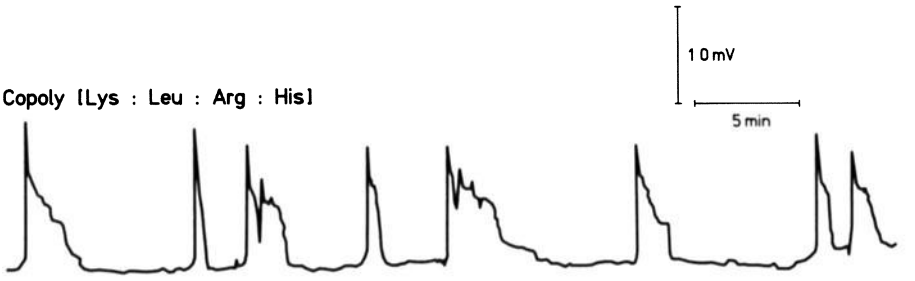


Fig. 9. Pattern of spontaneous electrical discharges of the proteinoid cell made of copoly (lys, leu, arg, his)

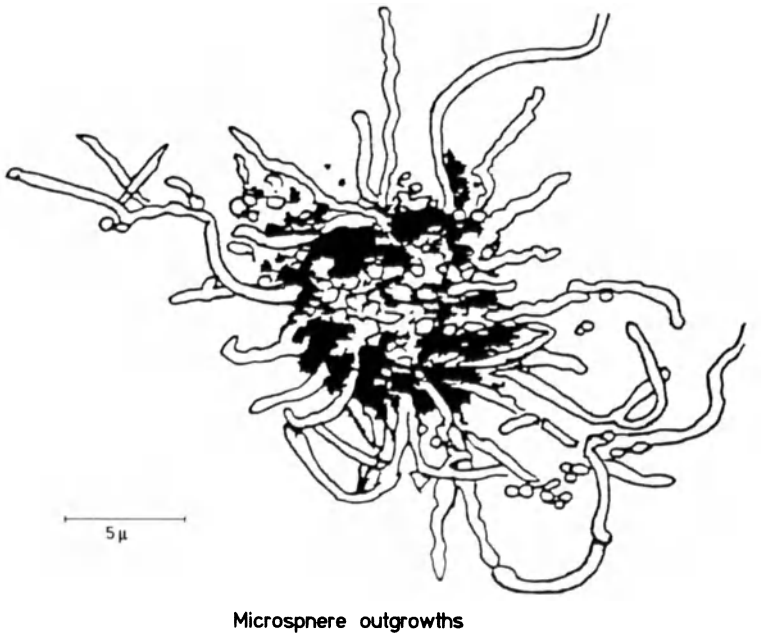


Fig. 10. A microsphere made from leucine, proline-rich proteinoid projects outgrowths resembling those from neurons

of some of these cells (Figs. 6 to 10) has allowed them to be regarded by some experts in the sense of artificial neurons.

This perspective allows us to engage in the possibility of constructing, or engineering, an artificial brain through the artificial neuron and the self-organizing properties that were employed by evolutionary processes.

The starting point is informed thermal protein. Thermal protein has yielded cells in one step. As a result of studying the properties of these cells we realise that the details indicate that a large proportion of them, or all, are excitable. The physiologist Howland (1973) has characterised all cells as excitable. This agrees smoothly with what has been found in the work here. In fact, we can with much

assurance state that the electrical behaviour is due to the protein portion of the cellular membrane and that such can be modified by an appropriate lipid.

The mind-body resolution is then one of thermal protein yielding cells, and the thermal protein also simultaneously yielding excitable cells. When the protein is of an appropriate compositional type the cells are neurons. *Starting thus with informed proteins, the appearance on the Earth of thermal proteins led quite directly into neurons, which then (experiments yet incomplete) evolved by self-association into brains.* This picture is of course greatly simplifying and greatly unifying, even as it omits innumerable details.

In the final section we will review some of the principal functions of proteinoid microspheres that permit them to be referred to as artificial neurons. It is especially simplifying that both cells in general and the excitable type that function as neurons in living systems are so closely derived from the same source. Crucial to that recognition was the finding that *informed* proteins could arise so easily and that *informed* cells could arise in turn by self-organization. Also crucial to this judgement is the finding of other cytofunctions.

10 The Evolutionary Direction

In many evolutionary studies the direction of the flow of events is largely overlooked. Science is predominantly analytic and approaches problems from the outside (e.g., into the cell). Evolution is however synthetic and in its flow of events moves from the inside out. For example, we know of protein by dismantling the cell and by charaterising the extracted materials. In order to learn about an evolution from protein to cell we had to synthesize protein from its junior precursors. That, however, is the evolutionary direction, opposite to the analytical direction.

From studies in the evolutionary direction, we could ascertain what we could not learn from analysis. Thereby it became possible to understand how first cells arose by self-organization of the thermal protein, whereas analytical studies of cells merely destroy the evolved organization.

Before this research began, the question could be asked: did the type of information processing influence later stages of evolution? When amino acid self-sequencing was suggested as the original source of bioinformation, it became possible to begin to recognise how that source influenced subsequent developments. In 1986, using data from the base in Martinsried, Ivanov and Förtsch explained that amino acid sequences in *modern* proteins are the consequence of self-sequencing. Also explained thereby were the results of Lipmann with polypeptide antibiotics, in which specific amino acids are added to the C-terminus one at a time.

The organization of thermal proteins into cellular structures is not limited to self-organization into cellular structures such has have served empirically as

models for the protoorganism. The thermal proteins, especially those rich in basic amino acids, e.g., lysine, organize cooperatively with other compounds, such as lipids and polynucleotides. The kind of datum obtained has provided understanding of how the first and later ribosomes came into existence (Fox and Dose, 1977, pp 232–240; Waehneldt and Fox, 1968). This area may be especially valuable and ready for future study of information processing.

The parallelisms between natural protein-polynucleotide complexes and the artificial ones examined in this project also deserve further study. In one example, lysine-rich proteinoid has been treated with, on the one hand, DNA, and on the other hand with RNA. The one containing DNA is fibrous whereas the one containing RNA in this controlled experiment is globular (Fox and Dose, 1977, p. 234) (Fig. 5). In natural systems, too, DNA is known to yield fibrous complexes, whereas RNA yields globular ones.

11 Biofunctions

The term "evolution of information processing" connotes in this paper "evolution of biofunctional information processing". In preceding sections evidence has been reviewed from heuristic experiments, for which the results showing the emergence of increasingly complex molecular structures are regarded as demonstrating the origination and subsequent emergence of biomacromolecular information. That the information in these molecular structures is biofunctional has been assumed largely on the basis that the kinds of compound known to be the repository of modern biological information have been produced in the laboratory under widespread geological conditions that exist now, and therefore plausibly existed earlier.

The argument has focussed on the view that the necessary kind of compound to contain sufficient information must be a polymer. The two plausible kinds of polymer are the variegated ones – nucleic acids and proteins. The experiments, and classical biochemistry, have indicated biofunctional molecules to be proteins while the nucleic acids serve as repositories, coded repositories, of information. The new research on origins has been initially all chemical. However, evidence has accumulated that a limited variety of appropriate chemical structures, not just exact duplicates of modern protein, could have arisen from the evolution of such informational molecules. The common feature for information appears to be in each case the exact spectrum of constellations of amino acid side chains in polymers. What has been learned, then, is that the kind of molecule that represents biofunctionality has been shown to be capable of spontaneous origin and also, in more recent research, that the biofunctions are themselves present in such polymers. By 1989, a long list of specific and general biological functions has been located in fairly full measure or, in a few cases, in root amounts in thermal proteins and their aggregation products. The first kind of evidence, the structural,

has been indirect whereas the second kind, that demonstrating biofunctionality, is direct. Both kinds of evidence are now abundant; the latter is reviewed in what follows.

Enzymatic. Shortly after the finding was reported that one could make polymers of α-amino acids in the laboratory by simple heating of α-amino acids (Fox and Middlebrook, 1954; Fox and Harada, 1958; Fox, 1960a), a number of colleagues asked for either samples or information so that they could test proteinoids for enzymatic activity. (The term thermal proteins was applied subsequently by Chemical Abstracts to replace the term proteinoid which they had used previously. The term thermal protein came into use in 1972.) What was found in a period of about ten years by various interested biochemists were representative activities of each of the major classes of enzymatic power as categorised by the International Union of Biochemistry. The table of enzymatic activities of thermal proteins has been reviewed a number of times (Rohlfing and Fox, 1969; Melius, 1982; Dose, 1984; Fox, 1980). The activities thus found are all weaker than those recorded for the modern counterpart enzymes, but this is explained on the basis that the power of an enzyme increased during evolution, probably with chiral purification of the monomers during evolution, an idea developed by W.A. Bonner, at Stanford University.

The enzymatic activities were found not only in the polymers but also in the microspheres assembled therefrom. Indeed, the activities are sometimes greater when the agents are in the microsphere form than when in aqueous solution. A few instances of coupled activities have been reported (Rohlfing and Fox, 1969; Fox, 1980) and have made their way into textbooks (Wessells and Hopson, 1988, pp. 443–444). These, then, are instances of the emergence of metabolic units, which are of more relevance than single enzymatic activities. Moreover, these examples include specificity in enzymatically active thermal proteins, e.g., lysine-rich proteinoid catalyzes the decarboxylation of oxaloacetic acid, whereas acidic proteinoid catalyzes the decarboxyalation of its reaction product, pyruvic acid. These two reactions already constitute a metabolic unit.

Classical cytofunctions. The activities that have been imputed for a long time by biologists to living systems are the properties of metabolism, growth and reproduction. These have been found in primitive form in proteinoid microspheres, as well as many other functions. (Fox, 1980, 1988). Except for the property of growth, each of these has been reproduced in textbooks by various authors. The property of growth is in some ways the most characteristic of living systems. In the experiments already published it can be seen that the growth of microspheres is quite exactly programmed. This then is a property of the protein molecules, without need to explain the control by invoking nucleic acids. The internal control is seen in the high uniformity of diameter of the microspheres.

Inasmuch as the self-sequencing of amino acids has been reported as a property of modern proteins by Ivanov and Förtsch (1986) the new data emphasizing internal control of cellular growth suggests looking for the controls of program-

med growth and their metabolic destruction in pursuit of understanding cancer processes.

Among other major properties that have been found are bioelectricity, membrane selectivity, ability to synthesize peptide and polynucleotide bonds, etc.

Other Cytofunctions. Yet other functions have been found in primitive form and related to specific thermal proteins and the intermediate proteinoid microspheres. They include intercellular recognition, hormonal activity (MSH), photochemical response, electrotaxis, compartmentalization (much influenced by calcium), osmotic behaviour, motility, attraction and avoidance, Brownian motion, conjugation, protocommunication between microspheres (a) via endoparticles and (b) via electric signals; antiaging, and protosocial protection.

Roots of Cerebrofunctions. In addition to enzymatic functions and classical cytofunctions, a group of what belongs under other cytofunctions is here however singled out for special attention. Any behaviour that appears to relate to cerebral activity at its evolutionary roots is potentially in a special position in EIP. Most, but not all of these, have been studied in proteinoid microspheres impaled by microelectrodes (Przybylski and Fox, 1986).

Among the non-electrophysiological attributes are extension of outgrowths and prolongation of life of true neurons when appropriate "primordial polymers" are added to cultures (Fox et al., 1987; Hefti, 1991), induction of enhanced memory when similar polymers are injected into mice (Fox and Flood, unpublished), and formation of junctions between microspheres (Fox et al., 1988). Indications of oscillograms characteristic of the composition of the polymers is then more promising (Przybylski and Fox, 1986). In this connection it should be recalled that the evidence is that the composition determines the sequence.

12 Bidirectional Transfer of Information

A principal conceptual resistance to the inferences from the experimental findings that indicate that the original source of biological information was protein is undoubtedly the mindset induced by the Watson–Crick Central Dogma, as touched on earlier. Crick's (1958) exact words were "the transfer of information from nucleic acid to nucleic acid, or from nucleic acid to protein may be possible, but transfer from protein to protein, or from protein to nucleic acid is impossible."

In his esteemed book on Origins, Shapiro (1986, pp. 290–291) points out that Shapiro interviewed Crick to learn that Crick had meant Central Hypothesis rather than Central Dogma. Shapiro also quotes Crick (1970) who said that the Central Dogma "was intended to apply only to present-day organisms and not to events in the remote past, such as the origin of life or the origin of the code". This judgment is compatible with our inference that the information can flow in

Table 2. (turbidities)

Polyribonucleotide	Lys-Rich Proteinoid	Arg-Rich Proteinoid
Poly C	++++	0
Poly U	+	+
Poly A	0	+
Poly G	0	+++
Poly I	0	++++

either direction (experimental) but only in one direction when the mechanism arose in organisms. We see also that, in an analytical perspective, composition can be inferred from sequence, while in a synthetic perspective from experiments, *composition determines sequence.*

The experiments of Yuki and Fox (1969) are relevant for the prebiotic relationships. Those experiments with models of prebiotic nucleic acids and prebiotic proteins indicated that information could be transferred either from polynucleotides to polyamino acids or from polyamino acids (thermal proteins) to polynucleotides in noncellular systems. Since thermal proteins are models for prebiotic proteins (Crick's "remote past") the Central Dogma does not apply. Shapiro favours proteins as the first informational macromolecules; this may be especially noteworthy inasmuch as Shapiro is himself a nucleic acid chemist and is among those who have been unable to rationalize the unsupported assumption that nucleic acids arose first. The results of Yuki's experiments are in Table 2.

The possibility that information could have been transferred in a kind of "reverse translation" from proteins has been suggested by a number of authors apart from these experiments. Some evidence exists of the possibility of such a process (Fox and Nakashima, 1984) and it is supported by investigators such as de Duve (1988), Shapiro (1986) and Fox (Fox and Dose, 1972). This is however an area in which more comprehensive experimentation should be carried out.

13 Informational Electrical Responses

Communication between microspheres is recognised as occurring in a fast mechanism and a slow one, as in modern systems. The slow one in the laboratory involves the transfer of informational proteinoid endoparticles from one spherule to another. This observation was made in 1970 (Hsu et al., 1971) and it is believed to have preceded the finding of transfer from modern cell to modern cell, as described by the Western Reserve workers (Lasek et al., 1977). Its identification was preceded by the finding that macrospheres avidly form junctions with each other and the junctions are hollow (Hsu et al., 1971). The proteinoid "endoparticles" are informational, since proteinoids are known to react selectively with

other proteinoids (Hsu and Fox, 1976), to react selectively with enzyme substrates (Fox and Dose, 1977), and to react selectively with polynucleotides (Yuki and Fox, 1969).

The fast communication in the laboratory is electrical. It was sought because of the earlier finding of the slow kind of communication (endoparticle transfer), because of the finding that the proteinoid membranes have electrical properties (Ishima et al., 1981; Przybylski and Fox, 1986) and because of the urging of the late physiologist, H. Burr Steinbach. To seek such properties first required awarenesses that had come along in the development of the research based on the evidence for proteins-first. One of the resultant inferences was the realization that the artificial protein was no more heterogeneous than unfractionated protein from most natural sources. The other change in mindset came from the finding that the electrical properties of the membrane are centered in the protein portion rather than the lipid, much as Nachmansohn had been insisting on in a rather lonely emphasis about 1970 (Nachmansohn, 1970).

Relatively specific readouts have been recorded for specific thermal proteins which were then assembled into microspheres and impaled with microelectodes (Przybylski and Fox, 1986). This is the one known research route that at present offers promise of explaining how specific information contained in neuronal systems as involved in learning can be produced and monitored (Schmitt, 1962; Fox, 1988). What is needed here is correlation of amino acid composition with electrical readout and with a cellular property such as photoreactivity (perhaps with Dr. Gordon Tollin), or with supramicrosphere morphology correlated with amino acid composition (this work has begun with Dr. Aristotel Pappelis). This kind of correlative, collaborative research is believed by the writer to be an avenue of outstanding promise. It includes the possibility of computerizing creative thinking, and of retracing the construction of a truly artificial brain.

The ways in which neurons interact in networks can be studied in these retracement models due to the fact that the proteinoid microspheres readily, and sometimes avidly, interact to form clusters. They also form junctions readily (Hsu et al., 1971).

A further supporting spinoff is found in the observation that administration of selected members of the thermal protein group enhance memory in mice. The first experiments (with Professor Franz Hefti) were set up on the basis that the model protoorganism, when appropriately constituted chemically, extends outgrowths. In the light of the connectionist theory of Hebb (1949), the outgrowths were hypothesized as being also connecting units and they were therefore tested in mouse brain by Dr. James Flood for memory effects. Such polymers as exhibited the outgrowth effects in the model were found significantly to promote memory-enhancement when injected into mice intracerebroventricularly (Fox and Flood, 1991). This is to be viewed in conjunction with the finding of Hefti that similar polymers stimulate growth of true axons and furthermore promote extended life in the neurons (Fox et al., 1987; Hefti et al., 1991).

The results suggest the broadest kind of evidence for both the vitality and the informational capacity of the continuously evolving thermal protein system.

Acknowledgements. Gratitude is expressed to a number of agencies that have supported the research, especially the National Aeronautics and Space Administration, the National Foundation for Cancer Research, and others cited in individual papers. Guidance in thinking and assistance in formulating concepts of the evolution of biological information are due to Professor Dr. Klaus Haefner and the Bertelsmann Foundation. Special thanks are recorded to Julie K Rowell for major help in composing this paper.

References

Braunitzer G (1984) Sequencing of hemoglobins. In: Brussels Hemoglobin Symposium 1983, Shnek AG, Paul C (eds) Univ Bruselles, Belgium, pp 341–353

Brooke S, Fox SW (1977) Compartmentalization in proteinoid microspheres, BioSystems **9**, 1 p 22

Cech TR (1983) RNA splicing. Three themes with variations, Cell **34**, pp 713–716

Crick FHC (1985) Protein synthesis, Symp Soc Exptl Biol **12**, pp 6–30

Crick F (1970) Central dogma of molecular biology, Nature **227**, pp 561–563

Crick F (1981) Life Itself: Its Origin and Nature. Simon and Schuster, New York

de Duve C (1988) Prebiotic syntheses and the mechanism of early chemical evolution. In: The Roots of Modern Biochemistry, Kleinkauf H, von Döhren H, Jaenicke L (eds) de Gruyter Co, Berlin, pp 881–894

von Döhren H, Kleinkauf H (1988) Research on nonribosomal systems: biosynthesis of peptide antibiotics. In: The Roots of Modern Biochemistry, Kleinkauf H, von Döhren H, Jaenicke L (eds) de Gruyter Co, Berlin, pp 355–367

Dose K (1984) Self-instructed condensation of amino acids and the origin of biological information, Int'l Quantum Chem, Quant Biol Symp **11**, pp 91–101

Eigen M (1971a) Self-organization of matter and the evolution of biological macromolecules, Naturwissenschaften **58**, pp 465–523

Eigen M (1971b) Molecular self-organization and the early stages of evolution, Quart Rev Biophys **4**, pp 149–212

Eigen M (1986) The physics of molecular evolution. In: The Molecular Evolution of Life, Baltscheffsky H, Jörnvall H, Rigler E (eds), Cambridge University Press, Cambridge, England, p 25

Fox SW (1945) Terminal amino acids in peptides and proteins. Advances Protein Chem **2**, pp 155–177

Fox SW (1956) Evolution of protein molecules and thermal synthesis of biochemical substances, Amer Scientist **44**, pp 347–359

Fox SW (1960a) How did life begin? Science **123**, pp 200–208

Fox SW (1960b) Self-organizing phenomena and the first life, Yearbook Soc Genl Systems Res **5**, pp 57–60

Fox SW (1977) Bioorganic chemistry and the emergence of the first cell. In: van Tamelen EE, (ed). Bioorganic Chemistry 3, Academic Press, New York, pp 21–32

Fox SW (1980a) The origins of behavior in macromolecules and protocells, Comp Biochem Physiol **67B**, pp 423–436

Fox SW (1980b) Life from an orderly cosmos, Naturwissenschaften **67**, pp 576–581

Fox SW (1981) Copolyamino acid fractionation and protobiochemistry, J Chromatogr **215**, pp 115–120

Fox SW (1985) Protobiological self-organization. In: Structure and Motion: Membranes, Nucleic Acids & Proteins, Clementi E, Corongiu E, Sarma RH (eds) Adenine Press, Guilderland, New York, pp 101–114

Fox SW (1988) The Emergence of Life, Darwinian Evolution from the Inside, Basic Books, New York

Fox SW, Dose K (1972, 1977) Molecular Evolution and the Origin of Life, rev edition, Marcel Dekker, New York

Fox SW, Flood J (1990) Effect of thermal proteins on memory processing for T-maze footshock avoidance training, Submitted
Fox SW, Harada K (1958) Thermal copolymerization of amino acids to a product resembling protein, Science, **128**, p 1214
Fox SW, Harada K (1960) The thermal copolymerization of amino acids common to protein, J Amer Chem Soc **82**, pp 3745–3751
Fox SW, Middlebrook M (1954) Anhydrocopolymerization of amino acids under primitive terrestrial conditions, Federation Proc **13**, p 211
Fox SW, Nakashima T (1984) Endogenously determined variants as precursors for natural selection. In: Individuality and Determinism, Fox SW (ed) Plenum Press, New York, pp 185–201
Fox SW, Harada K, Kendrick J (1959) Production of spherules from synthetic proteinoids and hot water, Science **129**, pp 1221–1223
Fox SW, Hefti F, Hartikka J, Junard E, Przybylski AT, Vaughan G (1987) Pharmacological activities in thermal proteins: relationships in molecular evolution, Intl J Quantum Chem Quantum Biol Symp **14**, pp 347–349
Harada K, Fox SW (1975) Characterization of functional groups of acidic thermal polymers of α-amino acids, BioSystems **7**, pp 213–221
Hartmann J, Brand MC, Dose K (1981) Formation of specific amino acid sequences during thermal polymerization of amino acids, BioSystems **13**, pp 141–147
Hebb DO (1949) The Organization of Behavior. John Wiley & Sons, New York
Hefti F, Junard EO, Knüsel B, Strauss WL, Przybylski A, Vaughan G, Fox SW (1991) Promotion of neuronal survival in vitro by thermal proteins and poly(dicarboxylic) amino acids, Brain Research **541**: 273–283
Howland JG (1973) Cell Physiology, Macmillan, New York, p 309
Hsu LL, Fox SW (1976) Interactions between diverse proteinoids and microspheres in simulation of primordial evolution, BioSystems **8**, pp 89–101
Hsu LL, Brooke S, Fox SW (1971) Conjugation of proteinoid microspheres: a model of primordial communication, Curr. Mod. Biol (now BioSystems) **4**, pp 12–25
Ishima Y, Przybylski AT, Fox SW (1981) Electrical membrane phenomena in spherules from proteinoid and lecithin, BioSystems **13**, pp 243–251
Ivanov OC, Förtsch B (1986) Universal regularities in protein primary structure: preference in bonding and periodicity, Origins Life, pp 35–49
Kleinkauf G, von Döhren H, Jaenicke L (eds) (1988) The Roots of Modern Biochemistry, de Gruyter, Berlin
Krampitz G, Fox SW (1969) The condensation of the adenylates of the amino acids common to proteins, Proc Natl Acad Sci USA **62**, pp 399–406
Lasek R, Grainer H, Barker JL (1977) The glia-neuron protein transfer hypothesis, J Cell Biol **74**, pp 501–523
Lehninger AL (1975) Biochemistry, 2nd edn, Worth and Co, New York
Lipmann F (1972) A mechanism for polypeptide synthesis on a protein template. In: Molecular Evolution, Prebiological and Biological, Rohlfing DL, Oparin AI (eds), Plenum Press, New York, pp 261–269
Lipmann F (1984) Pyrophosphate as a possible precursor of ATP. In: Molecular Evolution and Protobiology, Matsuno K, Dose K, Harada K, Rohlfing DL (eds), Plenum Press, New York, pp 133–135
Melius P, Sheng (1975) Thermal condensation of a mixture of six amino acids, Bioorg. Chem **4**, pp 385–391
Nachmansohn D (1970) Proteins in excitable membranes, Science **168**, pp 1059–1966
Nakashima T, Jungck JR, Fox SW, Lederer E, Das BC (1977) A test for randomness in peptides isolated from a thermal polyamino acid, Intl J Quantum Chem Quantum Biol Symp **4**, pp 65–72
Oparin AI, Rohlfing DL (1972) Molecular evolution and Sidney W Fox. In: Molecular Evolution: Prebiological and Biological, Rohlfing DL, Oparin AI (eds), Plenum Press, New York, pp 1–5
Popper KR, Eccles J (1977) The Self and its Brain, Springer, Berlin
Przyblyski AT, Fox SW (1986) Electrical phenomena in proteinoid cells. In: Bioelectrochemistry, Gutmann F, Keyzer H (eds), Plenum Press, New York, pp 377–396
Purves D (1990) A constructionist view of postnatal brain development, Society for Neuroscience Program, St Louis MO, Oct 28–Nov 2, Abstract **255**, p 119
Rohlfing DL, Fox SW (1969) Catalytic activities of thermal polyanhydro-α-amino acids, Advances Catal **20**, pp 373–418

Rosmus J, Deyl Z (1972) Chromatographic methods in the analysis of protein structure. Methods for identification of N-terminal amino acids in peptides and proteins, BJ Chromatogr **70**, pp 221–339

Schmitt FO (1956) Macromolecular interaction patterns in biological systems, Proc Amer Philos Soc **100**, pp 476–486

Schmitt FO (1962) Psychophysics considered at the molecular and submolecular levels. In: Horizons in Biochemistry, Kasha M, Pullman B (eds), Academic Press, New York, pp 437–457

Shapiro R (1986) Origins, Summit Books, New York

Tyagi S, Ponnamperuma C (1990a) A study of peptide synthesis by amino acyl nucleotide anhydrides in presence of complementary homopolynucleotides. In: Prebiological Self-Organization of Matter, Ponnamperuma C, Eirich FR (eds), Deepak Publishing, Hampton, VA, pp 197–210

Tyagi S, Ponnamperuma C (1990b) Nonrandomness in prebiotic peptide synthesis, J Mol Evolution **30**, pp 391–399

Uttall WR (1978) The Psychobiology of Mind, Lawrence Erlbaum Associates, Hillsdale, NJ

Waehneldt TV, Fox SW (1968) The binding of basic proteinoid with organismic or thermally synthesized polynucleotides, Biochem Biophys Acta **160**, pp 239–245

Waldrop MM (1989) Did life really start out in an RNA world? Science **246**, pp 1248–1249

Wessells NK, Hopson JL (1988) Biology, Random House, New York, pp 443–444

Wilson EO (1978) On Human Nature, Bantam Books, New York

Young RS (1984) Prebiological evolution: the constructionist approach to the origin of life. In: Molecular Evolution and Protobiology, Matsuno K, Dose K, Harada K, Rohlfing DL (eds), Plenum Press, New York

Yuki A, Fox SW (1969) Selective formation of particles by binding of pyrimidine polyribonucleotides or purine polyribonucleotides with lysine-rich or arginine-rich proteinoids, Biochem Biophys Res Commun **36**, pp 657–663