15 March 1963, Volume 139, Number 3559 SCIENCE

On Biochemical Variability and Innovation

Seymour S. Cohen

Many biochemists are concentrating their thought and effort on the similarities that appear in the study of cellular biochemistry and have tended to overlook or to discount that which is different and novel. Kety (*1*) has also pointed to the relation of certain biochemical orientations and assumptions to the slow rate of development of some fields, of which neurochemistry is an example. In this consideration of the general physiology of specialized cells, I wish to explore that which might appear different and novel in specializations of function and to discuss some aspects of the origin of metabolic and physiological function as they are related to whatever we know about the associated molecular structure.

Several factual and historical circumstances have served to establish the notion that specialization of biochemical function originates in cells already well endowed to perform many functions. Thus, we are asked to believe that a so-called primitive cell a priori possessed the molecular basis of any given specialization. First, among multicellular organisms specialized cells do arise from the undifferentiated zygote which possesses all the potentialities later expressed in its progeny of differentiated cells. Second, the term "the unity of biochemistry" (*1*) has been both widely adopted and frequently accepted to signify that there are broad and detailed similarities of composition and metabolic patterns among cells that have been placed both low and high on the phylogenetic totem pole. Third, many of the important advances in the past two decades in our knowledge of the relation of genetic mechanisms to biochemical function have been achieved with microbial techniques in which mutational change has been accompanied by a loss of function. These data, supplementing those on the nutritional requirements of parasites and some other microorganisms (*2*), have led to the proposition that acquisition of biosynthetic function was primarily a very early step in our biochemical origin (*3*). As a consequence of this development of facts and hypothesis, the opinion has been expressed rather broadly that the biochemical aspect of evolution has been monophyletic and has conformed to the simple pattern shown in Fig. 1.

The formulation is not objectionable as far as it goes. However, the possibility of a relatively recent gain of structure and function is not considered in this scheme, and it is precisely this possibility which I wish to consider. The need to define the biochemical nature of specialization has led in recent years to the accumulation of data pointing to the existence of molecular diversity within a single organism and among organisms. As new metabolites and enzymes continue to appear, one asks whether they are merely variations of known forms and if so how many variations of the old forms are possible. One also asks if the new metabolites and enzymes may not have arisen independently of the old metabolites and enzymes, and if so how this could have occurred. A growing interest has thereby developed in the definition of origin as well as in the definition of function. The advances of the past decade have clarified genetic problems to the degree that many problems of evolution, phylogeny, and embryology are now ripe for study by biochemists.

Advances in both the chemotherapy of infectious disease and agents of selective toxicity have emphasized that biochemical differences do exist between host and parasite, qualitative and quantitative differences which make chemotherapy feasible. In addition, the isolation and characterization of the antibiotics have revealed an astonishing new world of apparent biochemical freaks. Little attention has been centered on the origin and place of these compounds in the biochemical world. As a result of the observed

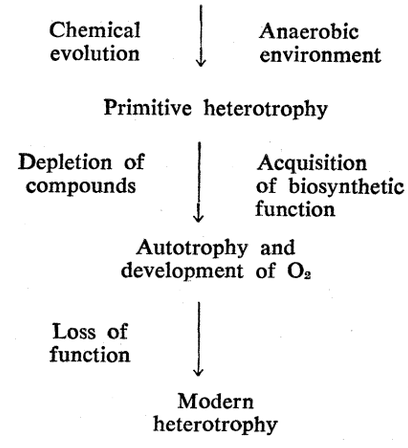

Fig. 1. Simple pattern of monophyletic evolution.

The author is Charles Hayden American Cancer Society professor of biochemistry at the University of Pennsylvania School of Medicine, Philadelphia. This paper was prepared in the Library of the Marine Biological Laboratory, Woods Hole, Massachusetts, and was presented at a symposium on the General Physiology of Specialized Cells held by the Society of General Physiologists at Corvallis, Ore., in August 1962. It will be included in the book, *Physiology of Cell Specialization*, A. Tyler and D. Mazia, Eds. (McGraw-Hill, New York, scheduled for publication in September).

toxicities on important (and fashionable) metabolic systems, the extraordinary nature of the inhibitors themselves is slowly entering the consciousness of biochemists.

Finally, investigators occasionally run into the unexpected; this can have a sobering effect, and it may produce some contemplation on the assumptions used in their work. For my part I have been fortunate enough to have had this experience several times, of which the most recent occurred when the DNA of the T-even bacteriophages was found to contain a new and unique pyrimidine and an unexpected but common sugar (*4*).

Common and Uncommon Nucleotides

I would like to think now about some of the major supports of the concepts of the "unity of biochemistry." These are the ubiquity of ATP, the nucleic acids, and the 20 amino acids in our proteins. Let us begin with the nucleotides and nucleic acids.

Can a replication, coding, and information-transfer system be conceived which does not contain nucleic acid? I have never heard of an effort to construct such a system of organic molecules, starting from the functional requirements rather than from the known components of cells. We know that if HCN and NH_3 are heated together, adenine is formed in high yield; this suggests that this base would be an expected product of chemical evolution (5). Furthermore, if adenine is made abundantly in the presence of pentose and polyphosphates, we can expect the formation of some adenylate (*6*). An abundance of adenine and adenylate would lead to polynucleotides containing adenine; these structures are in turn complementary to uracil and thymine nucleotides. The latter can be selected for and may be organized into a paired strand leading to the formation of double-stranded molecules (*6*). Although these chemical syntheses have been observed in the laboratory, it is not unreasonable to ask if other possibilities exist and might be encountered in genetic systems.

As it happens, a number of curious naturally occurring analogues of nucleotides are known; their structures are presented in Fig. 2. These are by no means all of the known naturally occurring analogues of adenosine. Some of the compounds such as the α-ribosyl benzimidazole (ribazole) in vitamin B_{12} are functional in normal cells, and some, such as ribosyl purine (nebularine), psicofuranine, cordycepin, and so forth, are inhibitory substances. Even so, these are all biosynthetic products. In comparison with adenosine, modifications are possible in each ring of the purine, in the ring substituents, and in the various properties of the sugar, its composition, steric relation to the base, and so forth. The nature of some of these modifications, as in ribazole or tubercidin, is such that the compounds could not have arisen merely as minor aberrancies of the biosynthesis of adenine. Their existence suggests biosynthetic mechanisms that are almost totally unexplored. What are these mechanisms, how might they also have originated, are they ubiquitous or sparsely distributed, are they early or late innovations? What conditions facilitated their survival and expression in some organisms and not in others? Some of these compounds can be phosphorylated but do not become part of a nucleic acid. Some, like tubercidin, in growing organisms do become constituents of nucleic acid after conversion to the appropriate phosphorylated derivative. The analogues which can be phosphorylated have not yet been observed (or really sought for) in the nucleic acids of the cells which produce them; indeed, they are quite inhibitory under conditions of incorporation. This is not surprising since most of these compounds were isolated initially as antibiotics.

The fact that these substances are made by cells raises the question of whether polymers similar in function and biosynthetic origin to the nucleic acids might be constructed of some suitable heterocyclic compounds other than adenine derivatives, if the heterocyclic compounds were produced under conditions in which adenine was either not produced or not competitive. In short, we must ask if this is the only form, as well as the best, of all possible biochemical worlds or will we ever find a cell in which the energy currency and nucleic acid are not constructed of adenosine triphosphate.

Some Aspects of Amino Acid and Protein Synthesis

The last decade has seen the partial unraveling of mechanisms of protein synthesis. Included is a genetic system in which the information for amino acid sequence contained in a polynucleotide is presumably transferred during synthesis to a messenger RNA which then attaches to ribosomes and acts as a template thereon for the assembly of 20 common L-amino acids borne to this site by adaptor S-RNA molecules. Even as described above for our natural nucleoside analogues,

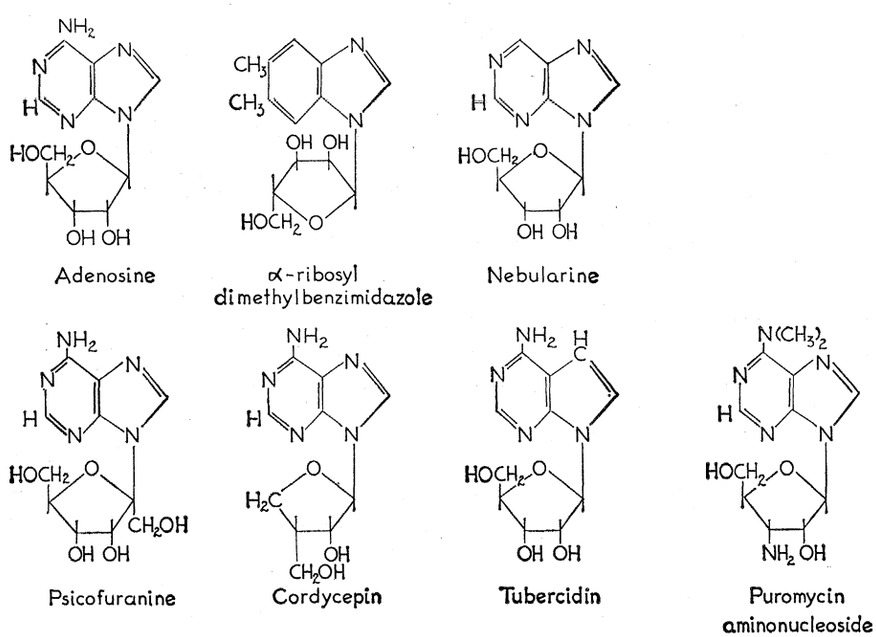

Fig. 2. Adenosine and some naturally occurring analogues.

there are some 60 other amino acids, D- as well as L-, most of which are found in the free state in plant extracts (7). Some of these are intermediates in the biosynthesis of the common amino acids and some are metabolites derived from the same amino acids. The role and origin of many of the "60-other" amino acids are obscure at present. Some of these, such as α-aminoadipic acid or *N*-methyl glycine (Fig. 3), are present in low concentration in plant proteins. We therefore ask if these uncommon amino acids are essential for function in the proteins when they are present. Are there activating enzymes for these amino acids, are they associated with specific S-RNA molecules, do they have a specific nucleotide sequence in DNA and messenger RNA, or is their presence the expression of a less than complete specificity for the more common amino acids? We also wish to know how the other uncommon amino acids of biosynthetic origin are excluded from polypeptides (if, indeed, they all are). It would evidently be of interest to know how the specific selection of 20 particular amino acids from approximately 80 occurred and whether the presence of these excluded amino acids in proteins would lead to proteins incapable of normal function, as happens when naturally occurring canavanine is incorporated into proteins, or whether these cannot be activated, coded for, or transported. The selection process may well be determined entirely competitively by relatively small differences in utilization at one of many metabolic steps. Thus, *Escherichia coli* can make selenomethionine from selenite when the concentrations of sulfate are low (*8*). However, although a methionine-requiring auxotroph of *E. coli* is able to utilize selenomethionine for exponential growth, it does so at less than the growth rate attained on normal methionine (*9*).

The existence of biosynthetic mechanisms for at least 80 amino acids raises numerous problems concerning the origin of the protein-synthesizing mechanisms. Before touching on this matter, it should be pointed out that for several common amino acids a number of different biosynthetic mechanisms exist. Indeed, the distribution of such pathways for the synthesis of a particular amino acid is of some phylogenetic interest. For example, two mutually exclusive mechanisms are known for the biosynthesis of lysine in which α-aminoadipic acid is the precursor for one pathway and α, ϵ-diaminopimelic acid is a precursor for another, as presented in Table 1. We can note parenthetically that α-aminoadipic acid is found in plants which do not use this amino acid for the synthesis of lysine. Different paths of biosynthesis are also known for ornithine, a precursor of ubiquitous arginine (*10*). In short, the mere presence of one of the inviolable 20 does not of itself provide full information concerning the details of its origin and of its predictable behavior within an organism. Contemporary organisms are the products of selection and result not only from many kinds of biochemical choice but they have frequently derived selective advantages by choosing several of the available possibilities.

In making this point it should be recorded that different enzymes, subject to different kinds of genetic and metabolic control in the same cell, such as *Escherichia coli*, participate at specific steps in the synthesis of several amino acids (*11*). Different enzymes for the same reaction are known in several instances when this reaction has two distinctly different physiological roles, as in catabolism and biosynthesis. At least two enzymes for a particular reaction in the biosynthesis of an amino acid are also known when one intermediate substance at a branch point is a precursor for two different amino acids; an example of this is the case of the two aspartokinases that function in the biosynthesis of lysine, threonine, and methionine. In these cases the differing enzymes which act similarly early in the biosynthetic sequence are affected differently by end-product inhibition and repression by the completed amino acids.

In an analysis of end-product inhibition of aspartyl transcarbamylase (*12*), the enzyme has several distinct sites of activity, of which that for the inhibitor could be inactivated without inactivating the catalytic function. Several instances of such double-headed systems are known; possibly the evolution of such enzymes by the acquisition of the inhibitable site occurred relatively late, at a time when the development of such control mechanisms as end-product inhibition conferred a

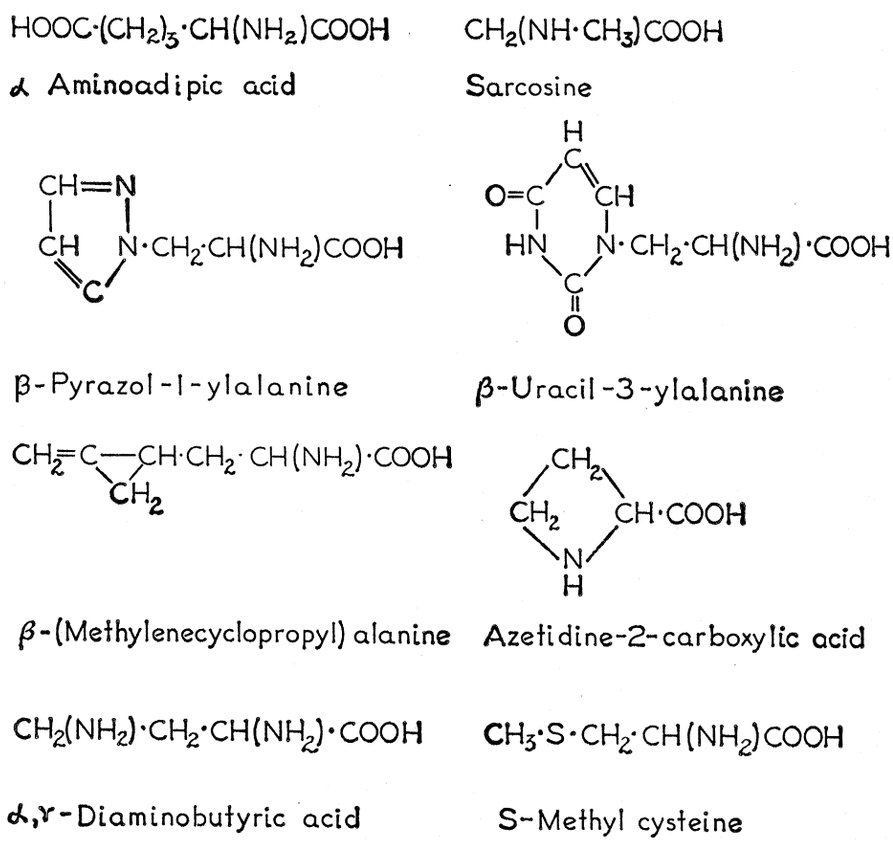

Fig. 3. Some naturally occurring amino acids.

selective advantage. Indeed, many such variations are becoming known; some of the most interesting include the reversible association and dissociation of protein subunits under the influence of various reactants of low molecular weight, such as the dissociation of glutamic dehydrogenase to alanine dehydrogenase which results from interaction with steroid hormones, and the association promoted by NAD (*13*) or ADP (*14*). The study of the relations of the two dehydrogenases in bacteria where sterols are lacking might elucidate a process of evolution, since the control of these enzymatic activities by control of aggregation of protein subunits might be more characteristic of higher organisms.

In addition, the phenomenon of multiple forms of the same catalytic protein has recently attracted many workers engaged in studies of higher organisms. Much effort has been devoted to the several lactic dehydrogenases in vertebrates in which the distribution of the two very different enzymes in relatively aerobic heart muscle and relatively anaerobic skeletal muscle is related to their different behavior toward lactate and pyruvate. Using apparently very different (by the criteria of serological cross-reactivity) protein subunits from the two types (*15*), the animals have found it possible to hybridize the subunits to adapt to particular functional requirements. Again, it seems reasonable to suppose that the development of such hybridized functional protein represents a relatively late evolutionary event.

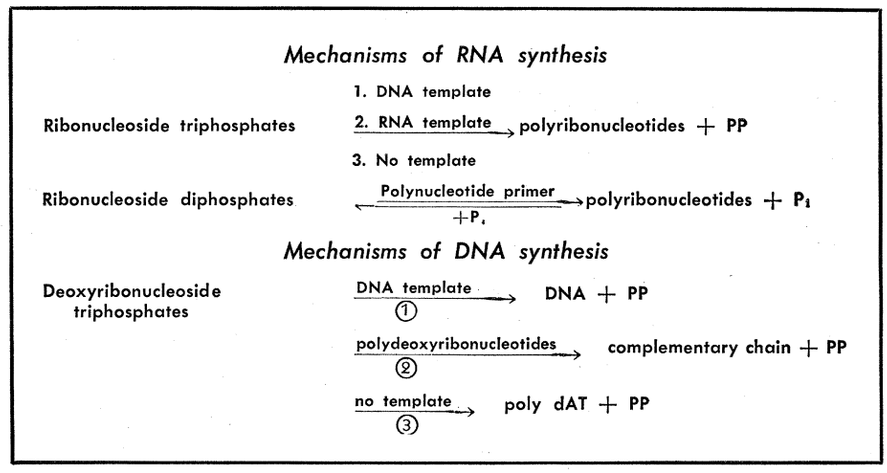

Fig. 4. Enzymatic mechanisms of polynucleotide synthesis.

Evolution of the Nucleic Acids

In the literature there is evidence that 3′-, 5′-phosphodiester bonds present in polynucleotides can be synthesized by numerous pathways, as shown in Fig. 4. There are many different ways of making a polyribonucleotide, including DNA-dependent RNA synthesis, RNA-dependent RNA synthesis, and nucleic acid-independent RNA synthesis, all of which require ribonucleoside triphosphates, as well as a route involving polynucleotide phosphorylase, which requires nucleoside diphosphates. Polynucleotide phosphorylase, which is being used widely to make the polynucleotides to stimulate the incorporation of amino acid in testing coding mechanisms, appears to have a limited distribution. This enzyme probably fulfills the role of turning off protein synthesis by destroying messenger RNA in microorganisms (*16*).

That there are multiple paths for RNA synthesis indicates the existence of at least two genetic systems, one based on the DNA in chromosomes and another based on RNA, as in the case of some RNA viruses (Fig. 5). Among the latter group the RNA of tobacco mosaic virus may be both a genetic determinant and a messenger in coding directly for protein synthesis (*17*). One should consider the possibility that such a system which combines several functions may be more primitive than the DNA system found in cells. Of course, we are not certain that there are not independent RNA systems in some cells, as posed by the problem of chloroplast inheritance. Furthermore, since deoxynucleotides are derived from ribonucleotides, the DNA polymerase and DNA-dependent RNA polymerases might be relatively late evolutionary developments from the RNA-dependent RNA polymerase. There is also the possibility that a primitive cell might contain RNA alone; it must then be asked whether a chloroplast, which has no known aerobic mechanisms, could not be such a simple cell that might grow and multiply in an appropriately complete organic environment. It has long been postulated that the plant viruses, all of which appear to contain RNA, are derived from variegated plastids. These viruses would then be degenerate forms of our hypothesized primitive cell, forms that conserve the essential qualities of inheritance and duplicability. In the search for intermediates between a hypothetical primitive cell and the crystalline plant viruses, the characterization of the difficultly characterizable plant viruses such as the amorphous sack-like potato yellow-dwarf virus (*18*) takes on a new significance. Conceivably the protein of an RNA virus may itself be an RNA-dependent RNA polymerase, or at the very least virus infection may induce the formation of such an enzyme.

Genetic systems containing DNA have had an extensive evolution. However, knowledge of the comparative biochemistry of chromosomes, division, and mitosis is virtually nonexistent, although we would imagine that the respective molecular structures and mechanisms would show differences that reflect the extensive evolution of the genetic mechanisms. There is accumulating evidence that the mechanisms of genetic control vary for different metabolic systems within one

Table 1. Partial distribution, phylogenetic, of the paths of biosynthesis of lysine, according to Vogel (*45*).

2, 6-Diaminopimelic acid path	2-Aminoadipic acid path
Bacteria	
Pseudomonads	
Eubacteria	
Actinomycetes	
Lower fungi	
Hyphochytriales	Chytrids
Saprolegniales	Blastocladiales
Leptomitales	Mucorales
Higher fungi	
	Ascomycetes
	Basidiomycetes
Green organisms	
Green algae	Euglenids
Ferns	
Flowering plants	

organism, and from organism to organism; this leads naturally to hypotheses of the evolutionary origin of the genetic apparatus. Of particular interest is the theoretical discussion of Monod and Jacob (*19*) on possible models interrelating a complex genetic apparatus and the control of metabolism, growth, and differentiation. Recent evidence indicates that structural genes control polypeptide sequence and regulator genes control other genes, even as described earlier by McClintock (*20*). The development of regulator genes and related systems may be thought to be later evolutionary events which have made possible the complexities of differentiation in advanced cells and multicellular organisms.

Among the bacterial viruses different molecular devices are invoked to store DNA strands; some viruses like ϕX-174 preserve a selected single strand, and others contain the more normal paired complementary strands. Furthermore, the evolution of these normal double-stranded DNA's among the viruses has developed considerably to form species of DNA very similar to that of the host in which it must multiply, as in lysogeny, or of DNA whose unusual composition is intimately related to the virulence of its parasitic relations, as in the T-even phage systems. The relations of these DNA templates to protein is most obscure in these forms, as indeed, are the relations of DNA and proteins, including enzymes, in a circular bacterial chromosome. For that matter, little enough is really known about the molecular structure of the chromosome in higher organisms.

Since only a very few investigators are examining different systems of synthesis of DNA and associated proteins in macromolecules or even chromosomes, there is at present a dearth of data on the variability of these systems. Although the basic mechanism in the polymerization to DNA seems to be similar in many cells, differences in DNA polymerases have nevertheless been observed. Thus, the DNA polymerase from thymus gland can utilize very small polydeoxynucleotide chains for complementary synthesis in contrast to the DNA polymerase from *Escherichia coli* which requires larger DNA templates. It is probably of significance in consideration of evolution that in the absence of a DNA template, the polymerization of dATP and dTTP can occur to form double-stranded poly dAT. Indeed, naturally occurring poly dAT has recently been discovered in a marine crab (*21*). Since the affinities of these particular strands are relatively low and the depurination of such strands by heat under very slightly acid conditions proceeds relatively easily and is mutagenic, a mechanism does seem to exist for the replacement of adenine by guanine and for the rapid evolution of DNA as we know it now.

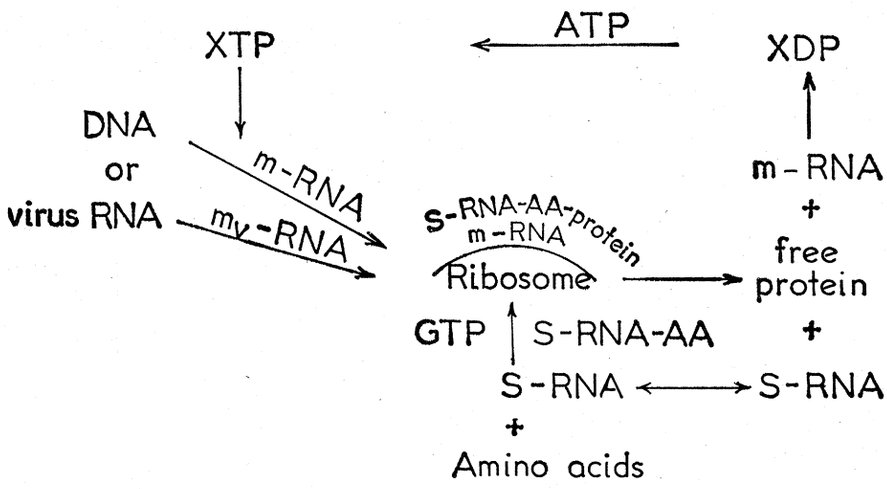

Fig. 5. A popular view of protein synthesis, showing different stores of genetic information.

It is relevant to ask why we have both DNA and RNA anyway, since an RNA system like that in the plant viruses can combine both functions. If RNA was a primary, and DNA a secondary, development in the process of evolution we can imagine that the need to arrest uncontrolled protein synthesis led to the development of RNA-degrading systems such as polynucleotide phosphorylase. In the face of such a system the development of a mechanism of preserving the pattern in nucleic acid conferred a selective advantage. Such mechanisms either protected the RNA by changing its conformation, or caused the addition of protecting components, or separated the storage function from the messenger function. Thus, the innovation of DNA may have been preserved because the structure protected the stored information from RNA-degrading enzymes and still permitted the formation of messenger RNA.

Variability of DNA and Proteins

Of considerable interest in the problem of the evolution of DNA are the data on the composition of DNA in organisms, data that may bear on phylogenetic relationships (*22*). Although higher cells (higher plants, invertebrates, vertebrates) have DNA whose contents of guanine (G) and cytosine (C) are in the range of 34 to 48 percent of the total bases, microbes have

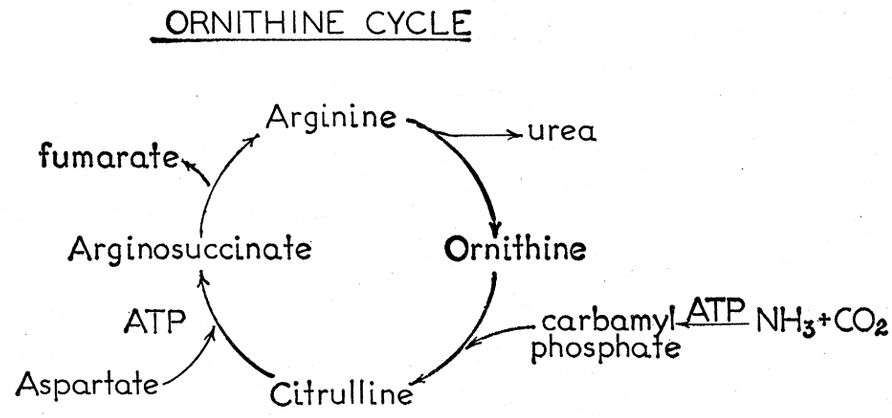

Fig. 6. The insertion of carbamyl phosphate into the ornithine cycle.

a far broader distribution, namely from 26 to 74 percent GC. Within any one organism the distribution of the G-C content of the DNA molecules is unimodal and the range is so narrow that if mean G-C contents of two organisms differ by 10 percent there are few DNA molecules of the same G-C content common to the two species. There have been several hypotheses to explain this startling result. These include: (i) The code translating polynucleotide sequence to amino acid sequence is not universal. The partial synthesis of the polypeptide sequence of tobacco mosaic virus on ribosomes from *E. coli* suggests that this hypothesis is incorrect. However, the proof of the universality of the code is far from complete at present. (ii) The degeneracy of the code, that is, the observed ability of different sets of nucleotides to control the insertion of a single amino acid into a polypeptide, may account for the wide variation in G-C contents of the DNA's from different organisms. It is not yet clear whether this can be true in a quantitative sense or is in fact so qualitatively. (iii) The DNA of organisms may contain a high proportion of nonsense sequences. The data are not complete on this possibility either; however, the characteristic narrow distribution even extends down to smaller regions of DNA molecules as revealed by fragmentation studies. (iv) The proteins and enzymes of organisms differ very considerably in amino acid composition and sequences within polypeptides. A few tests of this last hypothesis seem to support this postulate. Thus, the alkaline phosphatases of two bacteria, *Escherichia coli* and *Serratia marcescens*, organisms which differ markedly in DNA composition, are physiologically similar but have quite different peptide sequences (*23*). If a fragment of *E. coli* carrying a genetic determinant for alkaline phosphatase is transferred to *S. marcescens*, an alkaline phosphatase is produced in the latter which is like that of *E. coli* (*23*).

If this last result proves to be of general significance, the proteins and enzymes of comparable function in many types of microbial cells should vary widely in their composition. Briefly, then, the same catalytic and functional activity might presumably have evolved independently many times. This suggests the possibility of a much greater potentiality for the development of proteins and enzymes and for the production of functional entities than had previously been suspected. In this event, it might be asked whether the cells whose DNA and protein compositions differ considerably might not have arisen independently.

Variability in Polypeptide Biosynthesis

There is as yet too little information on the process of assembly of polypeptide on ribosomes to know whether there are significant variations of this particular process. However, we must ask if the system of protein synthesis presented in Fig. 5 is the only mechanism of protein synthesis. The rate of polypeptide synthesis systems of this type is of the order of one thousandth of that observed in intact cells and it is only reasonably cautious to ask if this discrepancy arises from technical difficulties alone or if there may not be something important which is being missed, such as a different process. Systems of amino acid activation other than those forming intermediate adenylates and S-RNA complexes are known; for example, whereas the system in Fig. 5 uses a pyrophosphate (PP) exchange:

$$\text{Amino acid} + \text{ATP} \text{ enzymes} \rightleftharpoons \text{PP} + \text{enzyme} \cdot \text{adenylate} \cdot \text{amino acid}$$

numerous important peptides such as glutathione may be formed by way of a reaction liberating orthophosphate (Pi):

$$\text{Glutamate} + \text{cysteine} + \text{ATP} \rightarrow \gamma\text{-glutamyl cysteine} + \text{ADP} + \text{P}_i$$

$$\gamma\text{-Glutamyl cysteine} + \text{glycine} + \text{ATP} \rightarrow \text{glutathione} + \text{ADP} + \text{P}_i$$

Similarly the D-alanyl-D-alanine present in the mucopeptide of the bacterial cell wall is made (*24*) by the reaction

$$2\ \text{D-alanine} + \text{ATP} \rightarrow \text{D-alanyl-D-alanine} + \text{ADP} + \text{P}_i$$

In neither case does RNA appear to be an intermediate in peptide formation in these systems.

Amino acids have been incorporated into acid-precipitable protein by systems which have no apparent pyrophosphate exchange mechanism (*25*). Continuing study of such a bacterial system (*26*) has revealed the formation of peptides by specific reactions of all of the major nucleoside triphosphates according to the reaction:

$$\text{Ribonucleoside triphosphate} + \text{amino acids} \rightarrow \text{nucleoside diphosphate} + \text{P}_i + \text{peptides}.$$

In this system glycine is activated maximally by ATP, glutamate by GTP, leucine by UTP and phenylalanine by CTP. This system is also possibly related to one in which the energy-dependent degradation of certain proteins leads to the formation of nucleotide-bound peptides in mammalian mitochondria (*27*). Nucleotide-bound peptides also participate in yeast extracts in orthophosphate exchange reactions with all of the ribonucleoside triphosphates (*28*), reactions which are possibly the converse of those studied by Beljansky (*26*).

The discoveries of the DNA polymerase, and the structure of DNA, S-RNA, and messenger RNA, and the composite systems leading to protein synthesis are unquestionably among the great achievements of modern biology and presage important discoveries ahead. However, from the preceding discussion we see that the evidence that this is the first and only significant route for the phenotypic expression of the genetic apparatus is not compelling; even if the data are made compelling, we would still wish to know in what directions and to what extent evolution has modified the numerous components of the schema. The data on these questions are far less complete than we would like. Apparent clarification of one important biochemical path frequently occurs to the neglect of another (*29*).

Improvement of Metabolic Function

Granick (*30*) has suggested a reasonable schema at least for the improvement of function, if not for its total acquisition. In discussion of the evolution of porphyrin biosynthesis, even as postulated by Horowitz (*3*), it was thought that compounds arising in chemical evolution probably already possessed the qualities necessary for biochemical function. In addition to the acquisition of the ability to synthesize these primitive compounds, Granick proposed that evolution included "the progressive elaboration of mechanisms for carrying out these functions in a more efficient manner." Thus the precursors to the formation of chlorophyll might also have been capable of photosynthetic activity and he exemplified the concept by noting that the precursor to chlorophyll *a*, protochlorophyll, is converted to the fully functional pigment by a photochemical reduction. Granick further proposed that the biochemist might approach problems of evolution with the idea

that "biosynthesis recapitulates biogenesis." Today it is possible to extend Granick's limited list of improved function with numerous others.

Thus pantetheine, phosphopantetheine, and dephosphocoenzyme A, precursors in the synthesis of coenzyme A, are active in many of the reactions in which the contemporary coenzyme A is functional. Also, the precursor adenosine-5′-phosphosulfate (*31*) is active in sulfate reduction in the anaerobe, *Desulfovibrio desulfuricans* (thought to be a more primitive process), whereas 3′-phosphoadenosine 5′-phosphosulfate is the active intermediate in sulfate reduction in yeast and *E. coli.*

In another instance NADP is derived from the phosphorylation of NAD, implying the primacy of the NAD systems and a later evolution of NADP- and NADPH-requiring systems. Both NADP and NADPH have been implicated in many specialized functions of differentiated cells. Even as suggested by Granick, NAD is active to a lesser extent with many NADP-requiring enzymes. Although NADP is most frequently required for the hexose monophosphate shunt, an occasional microorganism, like the anaerobic *Leuconostoc mesenteroides,* uses NAD in this path. This may be a vestige of a primitive character and there might be other potentially primitive characters that survive in this organism. In possible explanation of these NAD-requiring systems in anaerobes, it has been postulated that the present role of NADP was established with the advent of molecular oxygen; this then led to the need for reduced components such as NADPH which could be used for reductive biosynthesis under conditions in which the oxidative reoxidation of NADH was adapted to supply energy. However, the apparent role of NADP in anaerobic photosynthetic phosphorylation suggests that NADP arose and was active in this function even before the accumulation of molecular oxygen.

On True Innovation

I do not know of a systematic attempt to construct a consistent schema of biochemical evolution in terms of our present rather well-advanced knowledge of intermediary metabolism. Perhaps this is a good time to embark on such a theoretical project, one that hopefully might reveal highly significant gaps and breaks in the paths, breaks which could suggest the origins of true innovation. Goldfine and Bloch in a most important paper (*32*), pointed out that analysis of biosynthetic mechanisms requiring oxygen revealed important breaks in the record of biochemical evolution. Thus, such oxygen-requiring mechanisms essential to the synthesis of certain unsaturated fatty acids, sterols, nicotinic acid, and tyrosine are found in all higher cells but cannot be found in bacteria. Indeed, sterols have not even been shown to be present in bacteria. Bacteria synthesize nicotinic acid and tyrosine by routes quite different from those used (and presumably more recently acquired) by higher cells for which molecular oxygen was available. Furthermore these paths appear to be mutually exclusive, and the relation of the acquisition of the aerobic oxygen mechanisms to the loss of the anaerobic mechanisms, if indeed bacteria are descended from an evolutionary precursor of the higher cells, remains an enigma.

Interestingly, oxygen was the only oxidant detected in the enzymatic conversion of coproporphyrinogen to protoporphyrin in animal cells (*33*). It is not known how cytochromes or chlorophylls are made in obligate anaerobes or which came first in evolution, the chlorophyll or the molecular oxygen. Also, a variety of quinones appear to require molecular oxygen for their biosynthesis and they have been implicated in oxidative phosphorylation, a clear-cut case of biosynthesis antedating function. Many of these compounds appear to participate in oxidative phosphorylation in microorganisms that lack other oxygen-requiring mechanisms. Thus, was the synthesis of quinone one of the earliest functions acquired in the presence of molecular oxygen or are there other mechanisms of biosynthesis of these compounds?

The fact that there are these differences in biosynthetic activity between bacteria and higher cells unequivocally exemplifies true acquisition of function after the development of an atmosphere rich in free (molecular) oxygen. Are there other and particularly more recently developed acquisitions of function that can be found?

Table 2. Partial survey of biochemical diversity among human cells (adapted from Florkin, *46*).

Cell of origin	Biosynthetic features
	Ectoderm
Epidermis	Keratin and cholesterol
Sebaceous gland	Lipids, saturated alcohols, hydrocarbons
Adenohypophysis	
Eosinophils	Growth hormone, lactogenic hormone
Basophils	Thyrotrophic hormone, ACTH, luteinizing hormone, follicle stimulating hormone
Nerve cells	Complex lipids, glutamine, acetyl choline
Modified nerve cells	
Adrenal medulla	Noradrenaline, adrenaline
Intestinal argentaffin	Hydroxytryptamine
Melanocytes	Melanin
Neurohypophysis	Vasopressin, oxytocin
	Endoderm
Salivary gland	Ptyalin
Parietal cells	HCl secretion
Peptic cells	Pepsinogen
Exocrine pancreas	Lipase, glycosidases, endopeptidases, nucleases
Endocrine pancreas	Glucagon, insulin
Intestinal mucosa	Exopeptidases
Liver	Glucose 6-phosphatase, acetoacetate, β-OH-butyrate, ornithine cycle, bile acids, glucuronates, bilirubin, carbamyl phosphate
Thyroid	Thyroxine, thyroid hormones
	Mesoderm
Muscle	Myosin
Connective tissue	
Histiocytes	Hemoglobin degradation
Fibrocytes	Collagen (hydroxyproline, hydroxylysine), elastin
Cartilage	Mucoproteins
Osteoblasts	Bone salts
Red cells	Hemoglobin
Adrenal cortex, gonads	Steroid hormones

Biochemical Innovation in Higher Organisms

In Table 2 are summarized a fairly large number of the biochemical specializations of some of the differentiated cells in man. Let us explore the possibility that many of these examples of specialized function and biosynthesis have appeared at relatively late stages in the process of evolution.

As can be seen in the table, the bio-

Table 3. Patterns of cellular organization, after Stanier and van Niel (*45*).

Eukaryotic	Procaryotic (bacteria and blue-green algae)
Nuclear and genetic features	
Nuclear membrane	No nuclear membrane
Mitosis	No mitotic mechanism
Multiple chromosomes	Single circular chromosome
Reciprocal recombination	Nonreciprocal recombination
Cytoplasmic features	
Membranous organelles, mitochondria, chloroplasts	No homologous membrane-surrounded structures
Separation of respiratory and photosynthetic organelles	Protoplast membrane as respiratory center Lamellae and vesicles in blue-green algae
Biosynthesis in stroma around chlorophyll in chloroplast	Chromatophores in bacteria Functional linkage-respiratory enzymes in chromatophores Biosynthesis in contiguous cytoplasm
Structures for cell movement	
Bundle of fibrils (9 around 2 central)	One fibril generally
Derived from basal body within cell	Not enclosed within cytoplasmic membrane
Structure of wall	
Sialic acid in mucoprotein in echinoderms and higher cells	Muramic acid in mucopeptide Presence of diaminopimelic acid (common)

synthesis of the thyroid hormone is a function of cells derived from endoderm. The iodinated thyroid hormones appear first in the Protochordata among the Urochordates and Cephalochordates (*34*), concentrated in a diffuse tissue in the first instance and in a special organ in the second; in both instances they are derived from the primitive pharynx. It is also well-known that iodotyrosine and other halogenated amino acids exist in the scleroproteins of numerous marine invertebrates, such as sponges. The spotty distribution of iodinated tyrosine derivatives in invertebrates and vertebrates poses the questions whether the existence of a mechanism for the synthesis of iodotyrosine in some invertebrates has any relation to the development of thyroid function, whether this iodination reaction was conserved for such a role in higher forms, or whether the two reactivities arose independently. A synthetic mechanism for such an unusual compound was probably not functionally significant in metabolic control in high organisms until a cell had evolved capable of forming the chemically more complex thyroxine and other thyroid hormones. Nevertheless, it would evidently be important to know whether the capacity for iodination survived the evolutionary steps immediately preceding the Protochordates and, conversely, whether the enzymatic functions which existed before the completion of the thyroid hormones can be detected in the lower forms.

Similar questions may be posed concerning the origin of the hemoglobins, whose role in oxygen transport is an important characteristic of the vertebrates. It has long been known that ferrous hemeproteins classified as hemoglobins have an apparently random distribution in many lower forms, including protozoa (*35*) and fungi (*36*). The significance in the process of evolution of this distribution is quite mysterious. Again, one might well imagine that the synthesis of such a protein might be discovered independently by many cells or that this capability was acquired very early by our hypothetical cellular precursor and was evoked and modified for use at the appropriate developmental stage in the appropriate ecological niche. There are some fish living in very cold water that do not have erythrocytes and hemoglobin (*37*) and it would be interesting to know whether these organisms have truly lost the capacity for erythrocyte and hemoglobin biosynthesis or whether they are not expressing this potentiality.

For vertebrates it is clearly not enough to be able to make hemoglobin. A functional blood system that contains hemoglobin also requires mechanisms for handling iron in large amounts, converting it to the ferrous form, absorbing it, transporting it (with a special transferrin), and storing and conserving it (with a special ferritin). Thus, the development of functional hemoglobin among the vertebrates seems to have included the acquisitions of several specialized proteins (*38*). It is here appropriate to ask whether these proteins or their precursors cannot be detected in lower forms also.

In contemplating the bone system of the vertebrates we would not expect to be concerned with the presence of bone in lower forms. However, the characteristic bone salt, hydroxyapatite, is formed in some unicellular flagellates (*39*). Indeed, its presence in a ciliate appears to be correlated with the performance of muscular work in burrowing (*40*). Once more we ask if the breadth of the evolutionary gap between protozoa and the vertebrates does not support the notion of independent discovery, as it applies to systems of ossification, but conclude that our point of view on this will require a far greater body of information on the mechanisms of bone growth. We suspect that even if the mechanisms of formation of hydroxyapatite are similar in the two groups vitamin D and parathyroid hormone, among other mechanisms, will not be required for the lower form, and probably do signify metabolic acquisition in the higher organisms.

Thus, special compounds often thought characteristic of (and functionally important for) specialized cells in higher animals, like iodotyrosine in thyroxine, hemoglobin, and hydroxyapatite are also known among lower forms. At first sight, this might be considered to be a demonstration of the validity of the theory of the "unity of biochemistry" and the assumed evolutionary implications of the theory. However, the data on these compounds are seen to be sufficiently scant so that little can be postulated about the evolutionary origin of these substances and the interrelationships of the organisms possessing the same substances. Not only are we confronted with a plethora of questions requiring experimentation and new data, but inspection of the requirements for synthesis and function of such specialized substances in the higher organisms suggests that these organisms have acquired new biosynthetic activities probably not present in the lower forms.

The Ubiquity of Carbamyl Phosphate

Few comparisons have been made of the enzymatic mechanisms for the formation of the specialized compounds listed in Table 2 which may exist in both higher and lower forms. Such analyses have been made, however, for the enzymes catalysing synthesis of carbamyl phosphate in many microorganisms and in animals (*41*).

Numerous reactions that occur in the ornithine cycle (Fig. 6) in mammalian liver have been observed in the biosyn-

thesis of arginine and its degradation to urea in many microorganisms. At first sight there seems to be very little that is truly new in a biochemical sense in this cycle which is so important in equipping the vertebrate for emergence from an aqueous environment. The ammonia which must be detoxified enters the cycle at two points, one by forming carbamyl phosphate for the carbamylation of ornithine to citrulline and the other by forming aspartate essential to the conversion of citrulline to arginine. Enzymes which synthesize carbamyl phosphate have been found in many bacteria and fungi and are also present in the livers of ureotelic animals (*41*). Indeed, a broad distribution for such enzymes might be expected since, in addition to its role in the formation of arginine, carbamyl phosphate is an important precursor in the synthesis of pyrimidine by the sequence:

Carbamylphosphate + aspartate →
carbamyl aspartate → dihydroorotate
→ orotate → uracil

However, after looking at this enzyme more closely a number of important problems are posed. In the first place the enzyme which synthesizes carbamyl phosphate in the liver of all ureotelic animals utilizes a characteristic mechanism which is different from the synthetic reaction observed in other organisms. The enzymes of the livers of all ureotelic animals, termed carbamyl phosphate synthetase, are antigenically similar and quite different in this respect from the enzyme of microorganisms, suggesting that the development of this unique protein is correlated with the origin of ureotelism. The liver enzyme catalyses an irreversible reaction requiring 2 moles of ATP and the cofactor, acetyl glutamate:

$$NH_4^+ + HCO_3^- + 2ATP \xrightarrow{\text{acetyl glutamate}} NH_2COOPO_3H_2 + 2ADP + H_3PO_4$$

This is in contrast to the reversible reaction catalysed by the bacterial enzyme, carbamate kinase, which requires 1 mole of ATP and does not require acetyl glutamate:

$$NH_4^+ + HCO_3^- + ATP \leftrightharpoons NH_2COOPO_3H_2 + ADP + H_2O$$

It should be noted that other routes also exist for the formation of carbamyl phosphate (*41*), among which can be recorded a reaction occurring in extracts of the mushroom in which glutamine is a substrate instead of free ammonia (*42*). Despite this proliferation of known mechanisms it is still not known how carbamyl phosphate is made in or supplied to animal tissues other than ureotelic liver (*43*). In short, with respect to the synthesis of this important metabolite important in the transcarbamylations that produce essential building blocks of protein and nucleic acid biosynthesis at many phylogenetic levels, careful comparisons of the biosynthesis reveal significant variations of biosynthetic mechanisms. The exaggeration of synthesis of this substance in liver as required by the exigencies of ureotelism depended not merely on minor modification of a mechanism known in lower organisms but seems to have produced a new protein catalysing a new more demanding (irreversible) and complex reaction mechanism.

Variability Among Cells

Cells are functionally and morphologically different. The differences in their floor plans frequently seemed to accentuate their basic similarities, as if Henry Moore or Picasso had been at work. However, recently, as a result of the increasing knowledge of microbial structure and function, the question has been asked whether bacteria (and blue-green algae) are really built on the same master plan as the higher cells. Stanier and van Niel (*44*) have concluded that the structural differences among cellular forms are indeed so considerable that we must consider that we have two essentially different lines of cells. Table 3 outlines the major differences between the two cellular types, differences based on the presence or absence of the common bounded nuclear apparatus and mitosis present in higher cells, the presence or absence of a differentiated cytoplasm, and the chemical organization of the cell membrane and wall.

These workers have not ventured to suggest that the two cell types may have arisen independently. Indeed, such a conclusion might well be incorrect since the major differences could be thought of as a consequence of the evolution of the single limited membrane of the primitive cell. Thus, the cytoplasm of a higher cell may be looked upon as the culmination of the evolution and differentiation of the nuclear membrane, and the "protoplast membrane" of the bacteria may be thought of as the relatively undifferentiated nuclear membrane of the microbe. If this were true, these two cellular types might have a common progenitor after all. Regardless of the validity of this speculation, significant diversification has occurred to produce the eucaryotic and procaryotic cells of which only one type of organization, that of the former, resulted in the further development to differentiated multicellular systems.

Summary

Examples of variability and apparent innovation presented in this paper are but a portion of the many examples in the literature. Thus, the notion of the "unity of biochemistry" has been advanced in an overly simplified form and reflects a primitive stage in the development of the discipline. Cells contain many more compounds and biosynthetic mechanisms than we had suspected or list in our texts. Accordingly, either early chemical evolution was far more extensive than we have postulated or an evolution of biochemical synthesis and function took place which was more extensive than has been postulated; perhaps both have occurred. The chemical choices available from the environment have been considerable rather than limited and the cells have chosen, adapted, improved upon a limited number of these, and in turn have themselves been selected. In the case of the naturally occurring antibiotic substances, biosynthetic mechanisms for compounds which do not fit within the cells' own nucleic acids and proteins to advantage have nevertheless been preserved since they contribute to survival. On the other hand, some of these compounds such as α-ribazole, are fitted to other metabolic uses. The existence of the large number of uncommon relatives of the common components of the nucleic acids and proteins in turn implies an enormous untapped area of potential knowledge concerning their paths of biosyntheses and the genetic and physiological controls for these same paths. That these compounds exist perhaps also indicates an expanded material basis for a continuing biochemical evolution.

Even common metabolic components may arise in and be degraded by different metabolic paths, a result which strengthens our feeling about the underestimation of the potentials for variations in chemical and biochemical evolution. Such an underestimation or oversimplified view of the significance

and ubiquity of a single path has tended to and still does restrict the search for new paths and an evaluation of their functional role. Multiple paths for the biosynthesis of amino acids and fatty acids, polynucleotides, and polypeptides suggest the possibility of several lines of metabolic, genetic, and structural development; These lines have culminated in several kinds of cells, differing very markedly in basic design and chemical function.

Examples of biochemical diversity among organisms point to acquisition of biosynthetic function at numerous stages of evolution. Several novel mechanisms have arisen in response to the availability of the metabolite, molecular oxygen, itself produced by organisms. The exigencies of multicellular life resulting from advancing evolution have themselves culminated in acquisitions of a wide variety of new syntheses, biochemically specific substances, and activities which relate to some of the specialized functions of the higher forms.

We conclude from this analysis that it is invalid to assume that a biochemical exploration of specialized function will necessarily yield a mere modification of some well-recognized mechanisms. In our estimation the data of biochemistry are not only insufficient to warrant assumptions concerning the nature of substance and mechanism but suggest potentialities and capacity for great complexity in the evolution of function.

References and Notes

1. S. Kety, *Science* **132**, 1861 (1960).
2. A. Lwoff, *L'Évolution physiologique. Étude des pertes de fonction chez les microorganisms. Actualités scientifiques et industrielles*, No. 970 (Hermann et Cie. Paris, 1944); B. C. J. G. Knight, *Harvey Lectures Series 43* (1947–8), p. 71.
3. N. H. Horowitz, *Proc. Natl. Acad. Sci. U.S.* **31**, 1953 (1945).
4. S. S. Cohen, *Advan. Virus Res.* **3**, 1 (1955).
5. J. Oro and A. P. Kimball, *Arch. Biochem. Biophys.* **96**, 293 (1960).
6. G. Schramm, H. Grotsch, W. Pollmann, *Angew. Chem. Intern. Ed. Eng.* **1**, 1 (1962).
7. L. Fowden, *Endeavour* **36**, 35 (1962).
8. T. Tuve and H. H. Williams, *J. Biol. Chem.* **236**, 597 (1961)
9. D. B. Cowie and G. N. Cohen, *Biochim. Biophys. Acta* **26**, 252 (1957).
10. H. J. Vogel, *Proc. Natl. Acad. Sci. U.S.* **39**, 578 (1953).
11. H. E. Umbarger, *Cold Spring Harbor Symp. Quant. Biol.* **26**, 301 (1961); E. R. Stadtman, G. N. Cohen, G. LeBras, H. de Robichon-Szulmajster, *ibid.* **26**, 319 (1961).
12. J. C. Gerhart and A. B. Pardee, *J. Biol. Chem.* **237**, 891 (1962).
13. Abbreviations: NAD and NAHD, nicotinamide adenine dinucleotide and its reduced form; NADP and NADPH, nicotinamide adenine nucleotide phosphate and its reduced form; ADP and dADP, the 5′ (pyro-) diphosphates of adenosine and deoxyadenosine respectively; ATP, dATP, and dTTP, the 5′ (pyro-) triphosphates of adenosine, deoxyadenosine, and thymidine respectively; poly dAT, a polydeoxyribonucleotide consisting of two complementary strands in each of which deoxyadenylate and thymidylate alternate and are the sole nucleotide components.
14. G. M. Tomkins and K. L. Yielding, *Cold Spring Harbor Symp. Quant. Biol.* **26**, 331 (1961).
15. R. D. Cahn, N. O. Kaplan, L. Levine, E. Swilling, *Science* **136**, 962 (1962).
16. M. Sekiguchi and S. S. Cohen, *J. Biol. Chem.* **238**, 349 (1963).
17. A. Tsugita, H. Fraenkel-Conrat, M. W. Nirenberg, J. H. Matthaei, *Proc. Natl. Acad. Sci.* **48**, 846 (1962).
18. M. K. Brakke, L. M. Black, R. W. G. Wyckoff, *Am. J. Botany* **38**, 332 (1951).
19. J. Monod and F. Jacob, *Cold Spring Harbor Symp. Quant. Biol.* **26**, 389 (1961).
20. B. McClintock, *ibid.* **21**, 197 (1956).
21. N. Sueoka and T. Cheng, *J. Mol. Biol.* **4**, 161 (1962).
22. N. Sueoka, *ibid.* **3**, 31 (1961).
23. E. R. Signer, A. Torriani, C. Levinthal, *Cold Spring Harbor Symp. Quant. Biol.* **26**, 31 (1961).
24. F. C. Neuhaus, *J. Biol. Chem.* **237**, 778 (1962).
25. M. Beljansky and S. Ochoa, *Proc. Natl. Acad. Sci. U.S.* **44**, 494 (1958).
26. M. Beljansky, *Comp. Rend.* **250**, 624 (1960).
27. N. W. Penn, *Biochim. Biophys. Acta* **37**, 55 (1960).
28. A. H. W. M. Schurrs, S. R. de Kloet, V. V. Koningsberger, *Biochem. Biophys. Res. Commun.* **3**, 500 (1960).
29. About 15 years ago I participated in the rediscovery of the long neglected hexose monophosphate shunt and helped to relate it to the origin of ribose. This work was done in the face of a world of biochemists committed, perhaps for the sake of simplicity, to the Embden-Meyerhof pathway. The plethora of alternative pathways of carbohydrate metabolism revealed in the recent past has developed our considerable respect for the evolutionary ingenuity of organisms, at least in these metabolic areas. There may be no possible analogy to protein synthesis at this point because there might be a selective advantage in minimizing the number of critical pathways for the biosynthesis of protein, just as there would be such an advantage in maximizing the ability to metabolize exogenous carbon sources to utilizable intermediates. However, insofar as broadening the latter capability tends to increase enzyme, that is proteins, the degree of synthesis and the presence of optimal equipment for survival must be balanced and compromised.

    The role of organic substrates in bacterial photosynthesis has recently been reexamined [R. Y. Stanier, *Bacteriol. Rev.* **25**, 1 (1961)]. The efforts to generalize the mechanisms of bacterial and green plant photosynthesis had led to a relatively simple scheme in which the organic substrates essential for purple sulfur bacteria and the purple nonsulfur bacteria were supposed to serve primarily as a reducing agent (even as does H_2S) to handle the oxidizing fragment generated in the photochemical cleavage of water. An early study of the stoichiometry of acetate utilization in certain bacterial photosyntheses was not accepted as a demonstration of a reductive photoassimilation of this carbon source because it did not fit the generalized theory. These recent studies of Stanier and his associates on the role of organic substrates in the photometabolism of bacteria [R. Y. Stanier, *Bacteriol. Rev.* **25**, 1 (1961)] show that these compounds do serve as a carbon source, as had been shown earlier by Gaffron and by Gest. Furthermore, the routes of the utilization of acetate in such photoassimilations are different in the purple sulfur or nonsulfur bacteria from those in the green sulfur bacteria. Thus development of our knowledge on the mechanisms of photosynthesis provides another interesting instance in which the predilection for simplicity has impeded the development of understanding.
30. S. Granick, *Harvey Lectures Series* **44** (1948–9), p. 220.
31. F. Egami, M. Ishimoto, M. Itahashi, 5th Intern. Congr. of Biochem. Symp. III (August 1961), in press.
32. H. Goldfine and K. Block, *Symp. Soc. Gen. Physiol.* (1961), in press.
33. S. Sano and S. Granick, *J. Biol. Chem.* **236**, 1173 (1961).
34. J. Roche, 5th Intern. Congr. of Biochem. Symp. III (August 1961), in press.
35. D. Keilin and J. F. Ryley, *Nature* **172**, 451 (1953).
36. D. Keilin and A. Tissières, *ibid.* **172**, 390 (1953).
37. J. T. Rudd, *ibid.* **173**, 848 (1954).
38. S. Granick, *Am. Naturalist* **87**, 65 (1953).
39. F. G. E. Pautard, *Biochim. Biophys. Acta* **28**, 514 (1958).
40. ———, *ibid.* **35**, 33 (1959).
41. P. P. Cohen and G. W. Brown, 5th Intern. Congr. Biochem. Symp. III (August 1961), in press; M. E. Jones, A. D. Anderson, C. Anderson, S. Hodes, *Arch. Biochem. Biophys.* **95**, 499 (1951); P. P. Cohen, *The Enzymes* (Academic Press, New York, 1962), vol. **6**, p. 477.
42. B. Levenberg, *J. Biol. Chem.* **237**, 2590 (1962).
43. M. D. Bowers and S. Grisolia, *Comp. Biochem. Physiol.* **5**, 1 (1962).
44. R. Y. Stanier and C. B. van Niehl, *Archiv. Mikrobiol.* **42**, 17 (1962).
45. H. J. Vogel, 5th Intern. Congr. of Biochem. Symp. III (August 1961), in press.
46. M. Florkin, *Unity and Diversity in Biochemistry* (Pergamon, New York, 1961).