Problems in Protein Biosynthesis

PETER LENGYEL

From the Department of Molecular Biophysics, Yale University, New Haven

ABSTRACT Outline of the steps in protein synthesis. Nature of the genetic code. The use of synthetic oligo- and polynucleotides in deciphering the code. Structure of the code: relatedness of synonym codons. The wobble hypothesis. Chain initiation and *N*-formyl-methionine. Chain termination and nonsense codons. Mistakes in translation: ambiguity in vitro. Suppressor mutations resulting in ambiguity. Limitations in the universality of the code. Attempts to determine the particular codons used by a species. Mechanisms of suppression, caused by (*a*) abnormal aminoacyl-tRNA, (*b*) ribosomal malfunction. Effect of streptomycin. The problem of "reading" a nucleic acid template. Different ribosomal mutants and DNA polymerase mutants might cause different mistakes. The possibility of involvement of allosteric proteins in template reading.

Barely 21 years after the great discovery of Avery, MacLeod, and McCarty, indicating that nucleic acids might carry hereditary information (1), we are confident that we understand the broad outlines of the chemical steps through which the information encoded in the nucleotide sequence of DNA is expressed in the form of protein. In the time available I can give only a short and necessarily superficial draft of some concepts of protein synthesis and try to point out a selected few of the gaps in our knowledge.

Steps in Protein Synthesis. The Nature of the Genetic Code

The first step in protein synthesis (Fig. 1) is the transcription of DNA into messenger RNA (2) complementary to one of the two DNA strands (3, 4). This is catalyzed by RNA polymerase (5–7). The second step will be the topic of this discussion and we will return to it subsequently. The third step is the folding of the peptide chain into a 3-dimensional biologically active form. We have come to believe that the 3-dimensional structure is determined by the amino acid sequence. This belief is based on the following observations: enzymes can be denatured, that is, their 3-dimensional structure destroyed, leading to loss of their biological activity. It was demonstrated in several cases that under proper conditions denatured enzymes regain their activity and 3-dimensional structure (8). This indicates that the 3-dimensional structure of the active protein (including its disulfide bonds, if any) represents the

most stable conformation of its polypeptide chain under physiological conditions.

In the second step, called translation, the sequence of four kinds of nucleotides in mRNA serves as the template for ordering the sequence of the 20 kinds of amino acids into a polypeptide chain.

The code used in this translation is quite simple. Since 20 kinds of amino acids have to be specified by 4 kinds of nucleotides, at least 3 nucleotides are needed to specify one amino acid. A nucleotide sequence specifying an amino acid is called a "codon." Each nucleotide participates in one codon only. This conclusion follows from the finding that a point mutation, i.e. the replacement of a single nucleotide, leads to the replacement of a single amino acid (9, 10). The order of mutations in the genetic map is the same as the order of amino acid replacements caused by these mutations in the peptide chain: the genetic map and the peptide are colinear (11, 12).

Dr. Hogness discussed in this symposium experiments correlating length of DNA fragments with distance on the genetic map. His results show that

$$\text{DNA} \xrightarrow{\text{transcription}} \text{RNA} \xrightarrow{\text{translation}} \underset{\text{chain}}{\text{polypeptide}} \xrightarrow{\text{folding}} \text{active protein}$$

FIGURE 1. Steps in protein synthesis.

the order of genes (on the genetic map) is the same as the order of polynucleotide blocks chemically "incarnating" them (in DNA); i.e., colinearity between the genetic map and DNA (see also reference 13).

That a codon consists of 3 nucleotides was first shown by genetic experiments with proflavin mutants (14, 15). Proflavin is a mutagen which can cause insertion into or deletion from DNA of a nucleotide during DNA replication. Crick et al. recombined sets of properly selected proflavin mutants of a single gene. While insertion (or deletion) of 1, 2, 4, or 5 nucleotides from a functional gene always led to an inactive gene product, insertion (or deletion) of 3 or 6 nucleotides led to an active gene product. Presumably, the gene into which 3 nucleotides are inserted, causes the formation of a biologically active protein one amino acid longer than the wild-type protein. If 1 or 2 nucleotides are inserted (or deleted), the triplet reading frame will be shifted into an incorrect position and will specify a different peptide from the insertion point to the end of the gene. These results indicate that codons are nucleotide triplets and that the sequence of nucleotides is translated triplet by triplet from a fixed starting point (14, 15).

How do the codons actually specify amino acids? It was first believed that there might be a stereochemical fit between the side chain of an amino acid and the group of nucleotides which specifies the amino acid (16) (see also reference 17). Crick argued, however, that such a fit was unlikely. He sug-

gested that since nucleotides have specific hydrogen-bonding sites, they should be recognized by these and predicted the existence of adaptor molecules as the tools of this recognition process (18).

The adaptors, subsequently discovered, turned out to be a kind of RNA, consisting of about 80 nucleotides, and were called transfer RNA (19). One of the most remarkable achievements in biology announced this year was the complete nucleotide sequence determination of alanyl-tRNA from yeast (20). Each adaptor is specific for one amino acid and is believed to have a portion

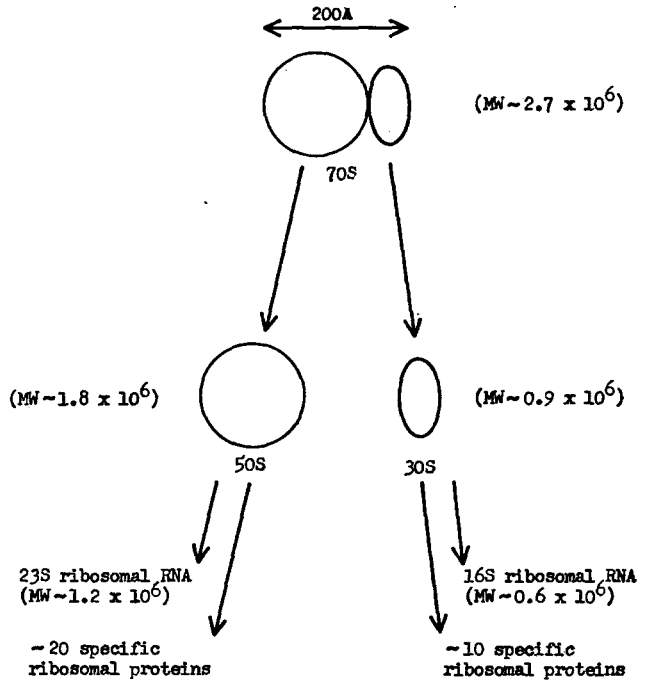

FIGURE 2. Diagrammatic representation of the structure of the *E. coli* 70S ribosome (23).

consisting of a nucleotide sequence complementary to the codon specific for the amino acid carried by the adaptor. This hypothetical sequence is the anticodon.

The nucleotide sequence of yeast alanyl-tRNA permits formation of a structure with several loops connected by double-stranded regions (20). Identification of the hypothetical anticodon, determination of the 3-dimensional structure of tRNA, and assignment of functions to different portions of tRNA are still outstanding.

The enzymes attaching amino acids to tRNA are called aminoacyl-tRNA synthetases (21). The reaction they catalyze is the following:

$$\text{amino acid}_1 + \text{ATP} + \text{enzyme}_1 \rightleftharpoons \text{aminoacyl}_1\text{—AMP—enzyme}_1 + \text{PP}$$
$$\text{aminoacyl}_1\text{—AMP—enzyme}_1 + \text{tRNA}_1 \rightleftharpoons \text{aminoacyl}_1\text{—tRNA}_1 + \text{enzyme}_1 + \text{AMP}$$

In the above equations, enzyme_1 stands for aminoacyl-tRNA synthetase specific for amino acid_1.

Each aminoacyl-tRNA synthetase appears to be specific for one amino acid. In line with the adaptor hypothesis, it was demonstrated that after an amino acid was attached to tRNA the coding specificity resided solely in the

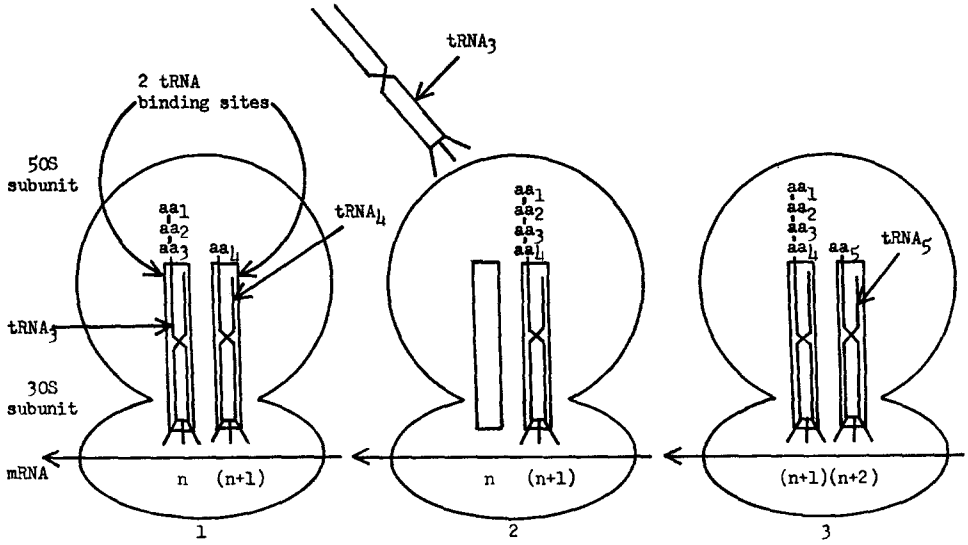

FIGURE 3. Schematic representation of the stepwise growth of the peptide chain on the ribosome.

1. The COOH terminal amino acid (aa_3) of the peptide is attached to $tRNA_3$. The peptidyl-tRNA is bound to the ribosome. aa_4tRNA_4 specified by $(n + 1)$ codon is bound to the ribosome through hydrogen bonding between the codon and the anticodon.

2. aa_3 forms a peptide linkage through its COOH group with the NH_2 group of aa_4. $tRNA_3$ is released.

3. The ribosome moves 3 nucleotides along the mRNA. aa_5tRNA_5 specified by codon $(n + 2)$ is bound to the ribosome.

tRNA component. When cysteinyl-tRNA was converted to alanyl-tRNA, the incorporation into protein of this alanine attached to cysteine-specific tRNA was still promoted by a codon specific for cysteine (22).

Polymerization of the aminoacyl moieties of the aminoacyl-tRNA's into a peptide occurs on the surface of the ribosome. Ribosomes are complex structures (Fig. 2), and at this time we cannot assign functions to their many components (23). Ribosomes are attached to mRNA through their 30S subunit (24). The mRNA-30S ribosome subunit complex can specifically bind

aminoacyl-tRNA (25, 26). For example, a poly U-30S ribosome subunit complex binds phenylalanyl-tRNA. A single codon, a triplet, is sufficient to direct specific binding of aminoacyl-tRNA to the ribosome (27).

The growing polypeptide chain, which is bound to tRNA through its terminal carboxyl group, seems to be attached to the 50S subunit (28). This attachment is not known to require mRNA. The growth of the peptide chain starts with the amino terminal amino acid and continues stepwise, amino acid by amino acid, toward the carboxy terminal amino acid (29, 30).

Fig. 3 illustrates aminoacyl-tRNA, specified by the forthcoming codon, in process of being attached to the mRNA ribosome complex. A peptide bond is formed between the carboxyl group of the carboxy terminal amino acid and the amino group of the new amino acid attached to tRNA. This step requires participation of at least two transfer enzymes (31–33), and GTP. One GTP molecule is split into GDP and orthophosphate for each amino acid attached to the peptide chain (34). Several ribosomes moving along and translating the same mRNA at one time form a so-called polyribosome (35, 36). The half-life of mRNA varies from a few minutes (certain mRNA's in bacteria, 37) to essentially stable mRNA in reticulocytes specifying hemoglobin (38).

Evidence has been provided that tRNA and ribosomal RNA are synthesized on the DNA template (39–41).

The Deciphering of the Code

The actual deciphering of the code was based on the following conceptual insights and experimental achievements.

For some time ribosomal RNA itself was assumed to be the template in protein synthesis. Jacob and Monod (2) pointed out that this could not be the case, since while ribosomal RNA in bacteria was metabolically stable (42), synthesis of inducible enzymes (i.e. new proteins) rapidly started or ceased on addition, respectively removal, of inducer to (or from) the bacterial culture. This and other considerations prompted them to predict the existence of mRNA carrying the information from DNA to the ribosomes. They suggested that ribosomes are nonspecific, could participate in the synthesis of any protein, and that it was the function of mRNA to attach to the ribosome and program it for the synthesis of a particular protein.

A cell-free amino acid–incorporating system was developed along the path laid down by the pioneer work of Zamecnik and Keller (43).

Grunberg-Manago and Ochoa discovered polynucleotide phosphorylase, an enzyme which can polymerize any kind of ribonucleoside diphosphate, singly or mixed, into long polyribonucleotide chains (44).

Reaping the harvest of all the above findings, Nirenberg and Matthaei were the first to use synthetic polynucleotides as messengers in a cell-free

amino acid–incorporating system. They found that a homopolynucleotide (poly U) promoted the formation of a homopolypeptide (polyphenylalanine) (45) (Table I). Lengyel, Speyer, and Ochoa established that copolynucleotides consisting of 2 or 3 kinds of nucleotides in a random sequence promoted the incorporation of several amino acids; i.e., they caused the formation of a copolypeptide (46). The kinds of amino acids incorporated into the copolypeptide depended upon the kinds of nucleotides present in the copolynucleotide used as messenger; furthermore, the ratio of amino acids in the co-

TABLE I

HOW SYNTHETIC OLIGO- AND POLYNUCLEOTIDES HAVE BEEN USED IN THE IN VITRO SYSTEM TO DECIPHER THE CODE

Poly U		Polyphe
Poly A		Polylys
Poly AC (5:1)	Promoted	$\begin{pmatrix}\text{Lys-thr-asn-gln-his-pro} \\ 100\ 26\ 24\ 24\ 6\ 7\end{pmatrix}$*
Poly AC (1:5)	formation	$\begin{pmatrix}\text{Pro-thr-his-asn-gln} \\ 100\ 21\ 23\ 5\ 5\end{pmatrix}$*
Poly $(UC)_n$	of	$(\text{Ser-leu})_n$
Poly $(AAG)_n$		Polyarg, polygln, polylys
UUU	Promoted attachment	Phe-tRNA
UCG	to ribosomes of	Ala-tRNA

* The numbers indicate relative amounts of amino acids incorporated into a random copolypeptide. For references see the text.

polypeptide product depended upon the base ratio of the copolynucleotide. The study with random copolynucleotides as messengers led to the elucidation of the nucleotide composition of about 50 codons (47, 48).

The sequence of nucleotides within codons has been (and is being) established mainly by two methods: Nirenberg and his colleagues have been using their elegant technique promoting the specific binding of aminoacyl-tRNA to ribosomes by nucleotide triplets (27, 49) (see also reference 50). Khorana and his collaborators developed in the last few years the chemical and enzymatic methods for synthesizing chemically defined polynucleotides with repeating oligonucleotide sequences and have been using the latter as messengers in cell-free amino acid–incorporating systems (51). They found, for example, that poly $(UC)_n$ (consisting of a repeating UC sequence) pro-

moted the formation of a polypeptide consisting of an alternating sequence of leucine and serine residues (52). Poly $(AAG)_n$ promoted the formation of 3 homopolypeptides: polyarginine, polyglutamic acid, and polylysine (53). These results strongly support the triplet nature of the code. The great significance of Khorana's studies with chemically defined messengers is partly due to the shortcomings of the method of aminoacyl-tRNA binding to the ribosome triplet complex: there are certain triplets which for unknown reasons

TABLE II
THE GENETIC CODE

1st nucleotide of codon	2nd nucleotide of codon				3rd nucleotide of codon
	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu (?)	Ser	Nons*	(?)	A
	Leu	Ser	Nons‡	Try	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu (?)	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ileu	Thr	Asn	Ser (?)	U
	Ileu	Thr	Asn	Ser (?)	C
	Ileu (?)	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg (?)	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

* Also called ochre codon.
‡ Also called amber codon.

do not promote aminoacyl-tRNA binding, though some of them, when occurring in polymers, do promote amino acid incorporation. The binding promoted by some other triplets is not very far above the background level. Finally, some triplets promote binding of several aminoacyl-tRNA's (49, 50).

The Structure of the Code

The code, as it is known in the fall of 1965, is shown in Table II (49, 50). The question marks in the table indicate that the experimental basis for the triplet assignment is weak or nonexistent. The pattern emerging is the following: most amino acids are represented by several triplets; i.e., the code is highly degenerate. Codons representing the same amino acid are called synonyms.

Synonym codons usually differ only in the third nucleotide. In this third position of the codon, U is equivalent to C (e.g. both UUU and UUC code for the same amino acid: phenylalanine) in all cases so far; A is equivalent to G in most cases; and all 4 nucleotides are equivalent in several cases. This pattern could be due to evolutionary causes. If at an earlier stage the codons were doublets coding for 15 amino acids only, and later the code expanded into its present triplet form specifying 20 amino acids, the nucleotides in the first and second position of a triplet specifying one of the 15 amino acids might

TABLE III

THE WOBBLE HYPOTHESIS

The most likely recognition set between the third nucleotide of the codon and a nucleotide of the anticodon (56).

Nucleotide of anticodon	Nucleotide recognized on codon (in third position)
U	A G
C	G
A	U
G	U C
I	U C A

still be the same as they were in the doublet code. Addition of the third nucleotide split the specificity of some of these doublets into two in order to be able to specify 5 more amino acids (54). Obviously, it is difficult to provide evidence for such an evolutionary hypothesis.

On the other hand, the particular relationship of synonym codons might be due to a feature of the translation mechanism. Evidence in favor of this view is accumulating. Highly purified alanyl-tRNA from yeast is bound to ribosomes in the presence of all 4 alanine codons having G and C in the first, respectively second position (49). Most of the phenylalanyl-tRNA from *E. coli* is bound to ribosomes with either UUC or UUU triplets (55). The above findings might indicate that base complementarity in the standard sense might not be an absolute requirement for pairing between the codon and the hypothetical anticodon. Crick suggests that while the standard base pairs may be used rather strictly in the first two positions of the triplet, there may be some "wobble;" i.e., more relaxed complementarity requirements in the pairing

of the third nucleotides (56). Among several alternatives offered, he prefers a pattern in which the recognition set shown in Table III exists between the third nucleotides of codon and anticodon. The suggestion of I as a third nucleotide of the anticodon, capable of recognizing U, C, or A (or even G) is based on the finding of Holley that I occurs in alanyl-tRNA adjacent to G and C and could well be part of the anticodon (20). The rules of "wobbling" between the third nucleotides might well be the cause for the relationship pattern between synonym codons.

TABLE IV

DETERMINATION OF THE NUCLEOTIDE SEQUENCE OF A NONSENSE CODON FROM AMINO ACID REPLACEMENTS RESULTING FROM MUTATIONS OF THE NONSENSE CODON (62)

Nonsense codon	Related codons		
	AAG (lys)	UCG (ser)	UAC (tyr)
UAG	CAG (gln)	UUG (leu)	UAU (tyr)
	GAG (glu)	UGG (try)	UAA (nonsense)

Initiation and Termination of the Peptide Chain

One mRNA molecule might serve as a template for the synthesis of several kinds of polypeptides; i.e., might be polycistronic (57, 58). In consequence of this, there is a need for a signal (or signals) in mRNA to indicate where the synthesis of the peptide chain should be finished and the synthesis of the next polypeptide started. It is not obvious whether such signal(s) are required for the translation of the actual beginning and of the end of the mRNA chain too.

Some codons do not seem to specify any amino acid in certain *E. coli* strains. They are referred to as nonsense codons. If such a nonsense codon is generated in a cistron (59, 60), the mRNA of the cistron will be translated into a peptide only from its 5′ end to the nonsense triplet; i.e., nonsense codons cause chain termination in peptide synthesis (12). The nucleotide sequences of two such nonsense triplets were deciphered as UAG and UAA in a study with specific mutagens (61). Another study in which different amino acid substitutions were analyzed occurring at the same position of the peptide chain and resulting from the reversion of the same nonsense mutation led to the same conclusions (62, 63). As illustrated in Table IV, the codon assignment for all the substituting amino acids can be related by single base changes only to a UAG nucleotide sequence.

Most of the peptide product formed in the cell-free *E. coli* system with natural mRNA (like phage RNA or endogenous mRNA) is free polypeptide (64). In contrast to this, many synthetic polynucleotides promote the formation mostly of peptidyl-tRNA still attached to the ribosomes. With synthetic polynucleotides containing both U and A in proper ratios, more than 50% of the peptide product is in the form of free polypeptide (64, 65).

The amino group of the amino terminal amino acid of many proteins is acetylated. However, the need for the acetyl residue in peptide chain initiation was made unlikely by the finding of an enzyme acetylating several finished proteins having different amino terminal amino acids (66). Detailed studies could not establish the existence of either acetyl amino acid activation or acetate transfer to aminoacyl-tRNA in *E. coli* extracts (67).

47% of the amino terminal amino acids of the protein from cell-free *E. coli* extract is methionine, 26% alanine, and 26% serine. On the other hand, methionine constitutes only 2.5% of the amino acid in the protein of this extract (68). It was of great interest, therefore, when formylmethionyl-tRNA was found in *E. coli* extracts (69). It was established that formylation occurs at the methionyl-tRNA level and that while at least 2 methionyl-tRNA fractions exist in *E. coli*, at least one of these cannot be formylated. Poly UG was found to promote incorporation of formylmethionyl-tRNA into the amino terminal position of peptides (70).

The coat protein of f2 phage has alanine as its amino terminal and serine as its second amino acid. It was discovered that the coat protein synthesized in the cell-free *E. coli* system with f2 RNA as messenger starts with *N*-formylmethionine, and alanine is the second, serine the third amino acid (71, 72). This raises the intriguing possibility that *N*-formylmethionine might be the compulsory chain-initiating amino acid in *E. coli* in vivo. That formylmethionine is not found as amino terminal of all *E. coli* peptides might be due to subsequent removal of formate, or formylmethionine (or even more amino acid residues) from these proteins. If formylmethionine or any other amino acid with a blocked alpha amino group is the chain initiator, there is no necessity to have special signals for chain termination. Since formylmethionine cannot form a peptide bond through its blocked amino group, wherever it is coded for in a polycistronic mRNA, it might serve at the same time as chain terminator and chain initiator. In vitro chains can be initiated without formylmethionine, since, e.g. phenyllactyl-tRNA (synthesized by deaminating phenylalanyl-tRNA) can serve as the source of the amino terminal moiety of the peptide chain, and phenyllactic acid has no amino group to which another amino acid could have been originally linked (73).

The direction of translation of mRNA (15, 74–76) is the same as the direction of its synthesis; i.e., from the 5′ end to the 3′ end (77, 78). This being so, it is possible that translation might be started while the mRNA is being built.

Mathematicians and theoreticians had speculated about a very fancy and complicated code structure (see reference 79). It turned out to be simpler than one dared to hope for. Not only was the simplest messenger a homopolynucleotide, not only did a single codon, a triplet, suffice to promote the binding of aminoacyl-tRNA to the ribosome, but recently it was reported that the shortest possible messenger, consisting of only two codons, promoted peptide formation. A_3U_3 caused the formation of some lysylphenylalanine in a cell-free system (75).

MISTAKES IN PROTEIN SYNTHESIS. AMBIGUITY AND SUPPRESSION

Under the heading of mistakes in protein synthesis, we will consider cases in which the same codon is translated in more than one way. It was observed that poly U promotes in vitro the incorporation into peptide not only of phenylalanine, but also of smaller amounts of leucine, isoleucine, serine, and tyrosine (80–82). This phenomenon, called ambiguity, was first thought to be an artifact of the cell-free system. Since similar phenomena were noticed in vivo, it was realized that ambiguity was an inherent characteristic of the protein-synthesizing machinery of the cell. A mutation which restores the function of a gene inactivated by a previous mutation is called a suppressor mutation. There is a class of suppressor mutations which cause ambiguity in translation (83, 84). A clear example of this was described for the case of *E. coli* tryptophan synthetase A (85). There is a glycine residue in a certain position of the peptide chain of the wild-type enzyme. As a result of a point mutation, the same glycine is replaced by arginine; this altered enzyme is inactive, therefore, the cell becomes tryptophan-requiring. A suppressor mutation occurring in a distant part of the genome (outside of the structural gene for this enzyme) causes the cell to produce two kinds of tryptophan synthetase A: 10% identical with the wild-type enzyme (with glycine), and 90% in the inactive mutant form (with arginine). In other words, as a result of suppressor mutation, the triplet, which specifies only arginine in the strain without suppressor activity, will be mistranslated into glycine in 10 cases out of a 100 in the strain with an active suppressor. The suppressed mutant is able to grow without added tryptophan. This phenomenon is called missense suppression.

We discussed earlier that in certain strains of *E. coli* the UAG and UAA triplets do not specify any amino acid, but behave as signals for chain termination. In other *E. coli* strains, harboring particular suppressor mutations, these nonsense triplets are translated to a varying extent into different amino acids. The kind of amino acid specified and the efficiency of translation depend on the suppressor mutation (86–89) (Table V). It is unlikely that the UAG triplet should be the real chain-terminating signal in vivo, since strains in which in 65 cases out of a 100 serine is inserted whenever the UAG triplet

is expressed (61, 90), grow with a normal generation time. It is possible, however, that UAG (or UAA) might be starting portions of the real chain-terminating signal. If that was so, one might expect to find that the carboxy terminal amino acid of proteins was more frequently serine in a strain having an efficient nonsense suppressor inserting serine than in strains lacking this suppressor. It is of interest that suppression itself can be ambiguous, since the same suppressor translates two different triplets.

Limitations in the Universality of the Code

It is a widely accepted conclusion that the code is universal in Nature. Some findings which support this conclusion are: the alkaline phosphatase structural gene of *E. coli*, if introduced into *Serratia marcescens*, causes the formation of the

TABLE V
SUPPRESSION OF NONSENSE CODONS (86–89)

Suppressor	Nonsense codon	
	UAG	UAA
Su –	Chain termination	Chain termination
Su 1	Ser	Chain termination
Su 2	Gln	Chain termination
Su 3	Tyr	Chain termination
Su 4	Neutral amino acid	Neutral amino acid
Su 5	Basic amino acid	Basic amino acid

E. coli alkaline phosphatase in its new host (91). Satellite tobacco necrosis virus RNA, used as messenger in a cell-free *E. coli* system, promotes the formation of a protein which seems to be the same as its normal coat protein formed in the tobacco plant (92); vaccinia virus is formed in *Bacillus subtilis* cells infected with vaccinia DNA (93). The fact, however, that the same triplet is translated in different ways in different strains of the same species sets limits to the universality of the code.

Attempts to Establish Which Codons Occur in a Species

The codons established in studying synthetic messenger-promoted amino acid incorporation and triplet-promoted aminoacyl-tRNA binding to the ribosomes in the cell-free system of a given species are the codons which this species is able to translate. It does not follow, however, that all these codons occur in its mRNA's (94). The availability of suppressors for nonsense mutations makes it possible to attempt the solution of this problem. To determine which codons do actually occur in a species, we have to study mutational amino acid replacements and establish which of the possible codons fit the amino acid replacement pattern. Alkaline phosphatase of *E. coli* contains 40

glutamine and/or glutamic acid and 25 lysine residues per subunit (95). According to the list of codons (Table II) lysine can be coded by AAA and AAG, glutamic acid can be coded by GAA and GAG. To find out whether all these codons do occur in the alkaline phosphatase structural gene of *E. coli*, a large number of nonsense codons were generated in this gene (62, 63, 96). Fifteen of these nonsense codons (each at a different site) were identified as UAG and 15 were identified as UAA by their response to different suppressor genes. The electrophoretic mobilities of the enzyme protein of the suppressed nonsense mutants, having a neutral amino acid specified by the nonsense condon, were compared with that of the wild-type enzyme. The mobility of all the proteins containing a UAG codon (suppressed by the insertion of a neutral amino acid) was the same as that of the wild-type enzyme. Thus, it was concluded that the original triplets, from which the UAG triplet arose by mutation, specified only neutral amino acids, and none of them specified lysine or glutamic acid. On the other hand, similar experiments led to the conclusion that about half of the UAA codons arose from triplets specifying lysine or glutamic acid in the wild type. These studies suggest that the AAG and the GAG triplets occur rarely (or not at all) in this gene; the primary lysine codon appears to be AAA, and the glutamic acid codon GAA. The UAG nonsense triplet was, however, found to revert in single step mutations to triplets specifying lysine or glutamic acid. This result shows that *E. coli* can translate the AAG and GAG codons. It would be premature to jump to general conclusions concerning the actually used codon set of *E. coli*.

How could one synonym codon be preferred to the other if both are recognized by the same tRNA? The answer to this question is probably indicated in the observation that the binding of a specific tRNA to the ribosome is more stimulated by one synonym codon than by the other synonyms (49).

How do the Mistakes in Translation Arise?

What is known about the mechanism of mistake formation? Which components of the protein-synthesizing machinery are to be blamed? The possible mechanisms can be grouped into two classes: in one class the binding between codon and anticodon is correct; i.e., complementary in the classical sense between the first, respectively second nucleotide pair, and complementary with the permitted wobble between the third nucleotide pair. (It should not be forgotten that the anticodon is still a hypothetical entity.) In this class the faulty component is aminoacyl-tRNA, more precisely, the amino acid is attached to a tRNA with an anticodon not specific for the amino acid. This might result, for example, from a mutational change in the anticodon of the tRNA itself, or a faulty activating enzyme loading an amino acid on the wrong tRNA, or a modification of the aminoacyl-tRNA by some other faulty enzyme, etc.

In the light of the wobble hypothesis (56) (Table III) a mutation in the

third position of the anticodon replacing a C (which only pairs with G) by a U (which pairs both with A and U) might lead to an ambiguous suppressor. In this context, it is not without interest that while there are suppressors specific for UAG alone, suppressors for UAA also suppress UAG (61, 89). This might turn out to fit the wobble hypothesis, according to which in the third position of a codon G can be specified separately from A; however, A is specified together with G (or with U and C).

The experimental proof for a tRNA being the culprit in the case of the suppressor inserting serine in response to the UAG triplet was provided recently (97, 98). It was established that the RNA from an R17 phage mutant, containing a nonsense codon, does not promote the formation of its coat protein in a cell-free system derived from an *E. coli* strain not having the suppressor. However, when serine-specific tRNA from a strain containing the suppressor was added to the above inactive system, the same RNA from the mutant phage became a good messenger for coat protein formation. As discussed earlier, it is not clear whether the suppressor mutation affected the serine-specific tRNA directly or indirectly.

A second class of mistake mechanisms involves mismatching in the codon-anticodon recognition; i.e., binding to a codon ribosome complex of an aminoacyl-tRNA with an anticodon not complementary to the codon in question (82). Actually, it was shown that all the isoleucyl-tRNA of *E. coli* could be bound to a poly U ribosome complex when the tRNA concentration was sufficiently low (99). The extent of this kind of mistake in the in vitro system decreases with increasing tRNA concentration (99, 100). This indicates that the tRNA, the anticodon of which is complementary to the codon promoting its attachment to the ribosome, is a better competitor for binding than the tRNA, the anticodon of which is only partially complementary to the codon.

It is most unlikely that this kind of base-pairing mistake could be attributed to faulty aminoacyl-tRNA or activating enzyme. As long as these mistakes were only found in the in vitro system it was possible to blame the biochemist for selecting poor conditions for his experiments or damaging components of the cell-free extract. It became obvious, however, that the same kinds of mistakes do occur in vivo too. It turned out that the source of such mistakes in vivo was a malfunction of the ribosomes (101). This unexpected finding came to light in a somewhat roundabout way.

The Role of Ribosomes. The Effect of Streptomycin on Translation

First it was established that ribosomes are the (or at least a) site of action of streptomycin. This was predicted (102) and then verified by the following experiment: poly U–promoted phenylalanine incorporation in vitro was inhibited by streptomycin if the ribosomes in the system were taken from a streptomycin-sensitive strain. It was not inhibited if they were taken from a

streptomycin-resistant strain. It made no difference whether the source of the supernatant was a sensitive or a resistant strain (103, 104). Subsequently, the site of action was further narrowed to the 30S subunit of the ribosome (105), and it was also shown that the genetic marker determining streptomycin sensitivity or resistance was cotransducible (certainly closely linked, if not identical) with a gene specifying a protein of the 30S subunit (106).

Second, streptomycin was found to act like a suppressor mutation. Actually, it was found that certain arginine-requiring bacterial mutants lacking one of the enzymes (ornithine transcarbamylase) of the arginine pathway grew without arginine if streptomycin was added to the medium. It was shown that under the latter conditions some ornithine transcarbamylase was formed in the cell (107). Certain kinds of suppression are due to mistakes in translation, and it was established that streptomycin increases the ambiguity of the cell-free system (101), causing mismatching in the codon-anticodon recognition process (99). Streptomycin is an aminoglycoside antibiotic. It was found that other members of this class also cause mistakes in translation (101). There is even some specificity in the action of the different aminoglycoside antibiotics: streptomycin stimulates poly U–dependent incorporation of isoleucine, but not of tyrosine; neomycin stimulates that of tyrosine (108). Neomycin causes not only mistakes in the recognition of the purine, pyrimidine bases, but even of the sugar component. In its presence, denatured (single-stranded) DNA acts as a messenger (109).

The fact that streptomycin, known to act on the ribosome, causes binding of the wrong aminoacyl-tRNA to the mRNA ribosome complex and increases the mistake level in translation indicates that it might be a function of the ribosome to permit only binding of the codon with the matching anticodon carrying tRNA. Ribosomes, which are distorted (for example, by streptomycin), however, permit binding to each other of a nonmatching codon-anticodon pair (99, 101). The above conclusion was reinforced when it was established that "miscoding" and, thus, the presumed distortion of the ribosomes, does not have to be caused by an external agent. Actually it was found testing the poly U–promoted isoleucine and serine incorporation in the absence of streptomycin, that streptomycin-resistant ribosomes were only one-tenth as ambiguous in vitro as streptomycin-sensitive ribosomes (110). Thus, in essence, ribosomes with a different structure differ in their accuracy of translation.

The opposite approach, which attempted to show that cells harboring a suppressor mutation might have ribosomes with a different structure from that of ribosomes of the wild-type cell, was reported to be successful. The disc-electrophoretic pattern of ribosomes from a strain containing a suppressor (arisen by a single step mutation which suppresses a nonsense mutation in the beta galactosidase gene) was claimed to lack one protein band which was present in the strain not having the suppressor (111).

Different Ribosomal Mutants Might Cause Different Translation Mistakes

The most intriguing aspect of this story is that the different ribosomal mutants do not just increase the mistake level, but they behave like allele-specific suppressors. Different ribosomal mutants cause (at least predominantly) different mistakes in the presence of streptomycin.

This conclusion was based on the following experiments (110): the *E. coli* strain used was streptomycin-sensitive and had a mutation in the ornithine transcarbamylase gene which was suppressible by streptomycin. A number of streptomycin-resistant mutants originating from this strain were isolated. These mutants were divided into: (*a*) one set in which the mutation in ornithine transcarbamylase was still suppressible by streptomycin, and one set (*b*) which lost the suppressibility. In a subsequent step, further auxotrophic mutants were selected from strains of each set. On testing these new strains it turned out that there occurred streptomycin-suppressible auxotrophic mutations in both sets, thus, even in (*b*), in which the original ornithine transcarbamylase mutation was not suppressible by streptomycin any more. This indicated that the mutant ribosomes, which mediated the effect of streptomycin, behaved like allele-specific suppressors. It was also established by in vitro studies that ribosomes from different strains are different in their response to the misreading action of streptomycin on the incorporation of particular amino acids. For example, in the wild-type (streptomycin-sensitive) strain streptomycin causes a 4.1-fold increase in poly U–promoted isoleucine incorporation, and an 8.4-fold increase in poly U–promoted serine incorporation. In one streptomycin-resistant strain the corresponding numbers are 12 and 2.6, in another streptomycin-resistant strain both are around 1 (110).

The Problem of Template Reading

The discussion of all these experiments concerning the function of ribosomes in translation served to emphasize the conclusion that while "normal" ribosomes might contribute to the accuracy of translation, differently distorted "bad" ribosomes (by either action of an external agent or a mutation) cause different, more or less specific mistakes in translation. The function of ribosomes (performed in conjunction with the transfer enzymes) is in essence analogous to the function of RNA polymerase and DNA polymerase. All three are template-translating machines with nucleic acids serving as the template. If normal ribosomes assure the accuracy of translation and bad ribosomes make mistakes, could not bad DNA polymerase make mistakes in DNA replication? Attempting to test this possibility, one has to take into consideration that the percentage of mistakes caused by a bad DNA polymerase might still be at or below the limit which can be satisfactorily picked up by the usual

biochemical assays. With the much more sensitive genetic assays, however, this might be more feasible.

Testing thermosensitive mutants of T4 phage, which have the mutation in gene 43, the gene known to specify T4 DNA polymerase (112), Speyer found that some of these mutant phages had a 1000-fold increased mutation rate compared to that of the wild type (113) (see also reference 114). He found that one particular allele of mutant DNA polymerase increased the reversion rate of two amber nonsense mutants (of the r_{II} region) 1000-fold above the rate with the wild type. The same allele had no effect on the reversion rate of two ochre nonsense mutations. A second allele of DNA polymerase had no effect on the reversion rate of the two amber mutations, but increased the mutation rate of at least one of the two ochre mutations to amber 800-fold (115). Thus, it seems that like bad ribosomes, bad DNA polymerase can also cause specific mistakes.

It is likely that some RNA polymerase mutants might also make more mistakes in transcription than the wild-type enzyme.

Streptomycin, as discussed at length, seems to act by distorting the ribosomes and not by interaction with mRNA or tRNA. It is not impossible that there should be mutagens which exert their action through interaction with DNA polymerase and not with DNA. It is not without interest that streptomycin itself is a very efficient mutagen for nonchromosomal genes (116).

We are used to taking it for granted that in the read-out of a template nucleic acid the selection of each nucleotide being subsequently built into the new chain and each anticodon in protein synthesis—occurs through direct interaction with the appropriate nucleotide(s) of the template. This is a simple and wonderful working hypothesis. There are analogies to support it, it explains many findings, but still it is, as of now, a hypothesis only which would warrant any effort to prove it. The fact that both DNA polymerase mutants and ribosomal mutants might cause specific errors in template reading makes one wonder whether the process of selecting the incoming nucleotide or triplet is not mediated through allosteric protein; i.e., with no direct interaction whatsoever between template and incoming nucleotide (113, 117, 118). To cite just two facts which might be indications that such a possibility is not completely absurd: (*a*) Enzymes can recognize oligonucleotides. This can be seen, for example, in the high specificity of DNA methylating enzymes. (See the contribution of Gold (page 5) this Symposium.) (*b*) Oligonucleotide sequences can cause allosteric alterations of enzymes which affect the specificity of the latter. This conclusion is derived from the following experiment: isoleucyl-tRNA synthetase can form not only the isoleucyl-AMP enzyme complex, but also a valyl-AMP enzyme complex (119). Loftfield and Eigner (120) studied the quantitative aspects of these two reactions by

assaying for aminoacyl hydroxamate formation that occurs according to the following reaction sequence:

$$\text{Amino acid} + \text{ATP} + \text{enzyme} \rightleftharpoons \text{aminoacyl} - \text{AMP} - \text{enzyme} + \text{PP}$$

$$\text{Aminoacyl} - \text{AMP} - \text{enzyme} + \text{hydroxylamine} \rightarrow \text{aminoacylhydroxamate} + \text{AMP} + \text{enzyme}$$

They found with purified isoleucyl-tRNA synthetase that if isoleucine and valine were present simultaneously at 10^{-6} M in a reaction mixture, 45 moles of valyl hydroxamate were formed for every 1000 moles of isoleucyl hydroxamate. If they added 2 mg per ml of tRNA, only 5 moles of valyl hydroxamate were formed for every 1000 moles of isoleucyl hydroxamate. This about tenfold increase in specificity of the reaction resulted from the alteration of the K_m and V_{max} values caused by the tRNA for isoleucine (120) (see also reference 121). 2 mg of tRNA which was previously unloaded and periodate-treated (and possessed less than 0.3% of its original capacity to accept isoleucine), had almost the same effect on the hydroxamate-forming reaction as 1 mg of untreated tRNA. Consequently the tRNA must have increased the specificity of isoleucyl-tRNA synthetase through changing the conformation of the enzyme; i.e., acting as an allosteric effector.

After discussing how the mistakes in protein synthesis might arise, one should not forget about their possible consequence: the microheterogeneity of proteins (122). To what extent molecules of identical amino acid sequence, comprising the bulk of a particular protein, are contaminated with molecules similar to them, but not identical with them, is not quite clear at the moment, but it is likely that this contamination in general is below the level of detection of the routinely used methods in protein chemistry. On the other hand, the case of tryptophan synthetase in the missense suppressor–containing *E. coli* strain, in which 90% of the molecules of this enzyme differed from 10% of the molecules by one amino acid is certainly not isolated (85, 123).

In conclusion it seems that we are not far from having completely deciphered the genetic code, namely, the one used in mRNA to specify amino acids. The problem of what might be called the second genetic code; i.e., the signals by which aminoacyl-tRNA synthetases recognize their specific tRNA has been hardly touched. The time is ripe to attack this question.

ABBREVIATIONS

The capital letters A, U, C, G, and I are used for the residues of adenylic, uridylic, cytidylic, guanylic, and inosinic acids in oligo- or polynucleotide chains. AMP, adenosine-5′-monophosphate; ATP, adenosine-5′-triphosphate; GDP, guanosine-5′-diphosphate; GTP, guanosine-5′-triphosphate; PP, inorganic pyrophosphate; DNA, deoxyribonucleic acid; RNA, ribonucleic acid; mRNA, messenger RNA; tRNA, transfer

RNA; poly AC (5:1) stands for the random copolymer consisting of 5 parts of adenylate residues and 1 part of cytidylate residue; poly AC (1:5), for the random copolymer consisting of 1 part of adenylate and 5 parts of cytidylate residues; poly UG, for the random copolymer consisting of uridylate and guanylate residues; poly $(UC)_n$, for a copolymer with a repeating sequence of UC; poly $(AAG)_n$, for a copolymer with a repeating sequence of AAG.

I want to thank several colleagues who kindly permitted me to discuss or refer to their observations prior to publication.
This work was supported by Grant No. GM 13707-01 of the United States Public Health Service.

REFERENCES

1. AVERY, O. T., MACLEOD, C. M., and MCCARTY, M., Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from Pneumococcus Type III, *J. Exp. Med.*, 1944, **79**, 137.
2. JACOB, F., and MONOD, J., Genetic regulatory mechanisms in the synthesis of proteins, *J. Mol. Biol.*, 1961, **3**, 318.
3. MARMUR, J., GREENSPAN, C. M., PALECEK, E., KAHAN, F. M., LEVINE, J., and MANDEL, M., Specificity of the complementary RNA formed by *Bacillus subtilis* infected with bacteriophage SP8, *Cold Spring Harbor Symp. Quant. Biol.*, 1963, **28**, 191.
4. SPIEGELMAN, S., and HAYASHI, M., The present status of the transfer of genetic information and its control, *Cold Spring Harbor Symp. Quant. Biol.*, 1963, **28**, 161.
5. WEISS, S. B., and NAKAMOTO, T., The enzymatic synthesis of RNA: Nearest neighbor base frequencies, *Proc. Nat. Acad. Sc.*, 1961, **47**, 1400.
6. HURWITZ, J., BRESLER, A., and DIRINGER, R., The enzymic incorporation of ribonucleotides into polyribonucleotides and the effect of DNA, *Biochem. Biophysic. Research Commun.*, 1960, **3**, 15.
7. STEVENS, A., Incorporation of the adenine ribonucleotide into RNA by cell fractions from *E. coli* B, *Biochem. Biophysic. Research Commun.*, 1960, **3**, 92.
8. HABER, E., and ANFINSEN, C. B., Side-chain interactions governing the pairing of half-cystine residues in ribonuclease, *J. Biol. Chem.*, 1962, **237**, 1839.
9. WITTMANN, H. G., Comparison of the tryptic peptides of chemically induced and spontaneous mutants of TMV, *Virology*, 1960, **12**, 609.
10. TSUGITA, A., and FRAENKEL-CONRAT, H., The amino acid composition and C-terminal sequence of a chemically invoked mutant of TMV, *Proc. Nat. Acad. Sc.*, 1960, **46**, 636.
11. YANOFSKY, C., CARLTON, B. C., GUEST, J. R., HELINSKI, D. R., and HENNING, U., On the colinearity of gene structure and protein structure, *Proc. Natl. Acad. Sc.*, 1964, **51**, 266.
12. SARABHAI, A. S., STRETTON, A. O. W., BRENNER, S., and BOLLE, A., Co-linearity of the gene with the polypeptide chain, *Nature*, 1964, **201**, 13.
13. KAISER, A. D., The production of phage chromosome fragments and their capacity for genetic transfer, *J. Mol. Biol.*, 1962, **4**, 275.

14. CRICK, F. H. C., BARNETT, L., BRENNER, S., and WATTS-TOBIN, R. J., General nature of the genetic code for proteins, *Nature*, 1961, **192,** 1227.
15. TERZAGHI, E., OKADA, Y., STREISINGER, G., TSUGITA, A., INOUYE, M., and EMRICH, J., Properties of the genetic code deduced from the study of proflavine-induced mutations, *Science*, 1965, **150,** 387.
16. GAMOW, G., Possible relation between deoxyribonucleic acid and protein structures, *Nature*, 1954, **173,** 318.
17. WOESE, C. R., Order in the genetic code, *Proc. Nat. Acad. Sc.*, 1965, **54,** 71.
18. CRICK, F. H. C., On protein synthesis, *Symp. Soc. Exp. Biol.*, 1958, **12,** 138.
19. HOAGLAND, M. B., The present status of the adaptor hypothesis, *in* Structure and Function of Genetic Elements, *Brookhaven Symp. Biol.*, 1959, **12,** 40.
20. HOLLEY, R. W., APGAR, J., EVERETT, G. A., MADISON, J. T., MARQUISEE, M., MERRILL, S. H., PENSWICK, J. R., and ZAMIR, A., Structure of a ribonucleic acid, *Science*, 1965, **147,** 1462.
21. BERG, P., Specificity in protein synthesis, *Ann. Rev. Biochem.*, 1961, **30,** 293.
22. CHAPEVILLE, F., LIPMANN, F., VON EHRENSTEIN, G., WEISBLUM, B., RAY, W. J., JR., and BENZER, S., On the role of soluble ribonucleic acid in coding for amino acids, *Proc. Nat. Acad. Sc.*, 1962, **48,** 1086.
23. WATSON, J. D., The synthesis of proteins upon ribosomes, *in* Cinquantenaire de la Société de Chimie Biologique, Paris, 1965, 157.
24. OKAMOTO, T., and TAKANAMI, M., Interaction of ribosomes and some synthetic polyribonucleotides, *Biochim. et Biophysica Acta*, 1963, **68,** 325.
25. MATTHAEI, H., AMELUNXEN, F., ECKERT, K., and HELLER, G., 1964, as quoted in Suzuka, I., Kaji, H., and Kaji, A., *Biochem. and Biophysic. Research Commun.*, 1965, **21,** 187.
26. KAJI, H., SUZUKA, I., and KAJI, A., Binding of specific soluble ribonucleic acid to ribosomes. Binding of soluble ribonucleic acid to the template-30S subunits complex, *J. Biol. Chem.*, 1966, **241,** 1251.
27. NIRENBERG, M. W., AND LEDER, P., RNA codewords and protein synthesis. The effect of trinucleotides upon the binding of sRNA to ribosomes, *Science*, 1964, **145,** 1399.
28. CANNON, M., KRUG, R., and GILBERT, W., The binding of sRNA by *Escherichia coli* ribosomes, *J. Mol. Biol.*, 1963, **7,** 360.
29. BISHOP, J., LEAHY, J., and SCHWEET, R., Formation of the peptide chain of hemoglobin, *Proc. Nat. Acad. Sc.*, 1960, **46,** 1030.
30. DINTZIS, H. M., Assembly of the peptide chains of hemoglobin, *Proc. Nat. Acad. Sc.*, 1961, **47,** 247.
31. BISHOP, J. O., and SCHWEET, R. S., Participation of two enzyme fractions in the transfer of amino acids from aminoacyl-RNA to reticulocyte ribosomes, *Biochim. et Biophysica Acta*, 1961, **54,** 617.
32. FESSENDEN, J. M., and MOLDAVE, K., Evidence for two protein factors in the transfer of amino acids from soluble RNA to ribonucleoprotein particles, *Biochem. and Biophysic. Research Commun.*, 1961, **6,** 232.
33. ALLENDE, J. E., MONRO, R., NATHANS, D., and LIPMANN, F., Resolution in the *E. coli* ribosome sRNA system of the aminoacyl transfer factor into complementary fractions, *Fed. Proc.*, 1962, **21,** 414.

34. NISHIZUKA, Y., and LIPMANN, F., Comparison of guanosine triphosphate split and polypeptide synthesis with a purified *E. coli* system, *Proc. Nat. Acad. Sc.*, 1966, **55**, 212.
35. GILBERT, W., Polypeptide synthesis in *Escherichia coli*. I. Ribosomes and the active complex, *J. Mol. Biol.*, 1963, **6**, 374.
36. WARNER, J. R., KNOPF, P. M., and RICH, A., A multiple ribosomal structure in protein synthesis, *Proc. Nat. Acad. Sc.*, 1963, **49**, 122.
37. KEPES, A., Kinetics of induced enzyme synthesis. Determination of the mean life of galactosidase specific messenger RNA, *Biochim. et Biophysica Acta*, 1963, **76**, 293.
38. MARKS, P. A., BURKA, E. R., CONCONI, F. M., PERL, W., and RIFKIND, R. A., Polyribosome dissociation and formation in intact reticulocytes with conservation of messenger ribonucleic acid, *Proc. Nat. Acad. Sc.*, 1965, **53**, 1437.
39. GIACOMONI, D., and SPIEGELMAN, S., Origin and biologic individuality of the genetic dictionary, *Science*, 1962, **138**, 1328.
40. GOODMAN, H. M., and RICH, A., Formation of a DNA soluble RNA hybrid and its relation to the origin, evolution and degeneracy of soluble RNA, *Proc. Nat. Acad. Sc.*, 1962, **48**, 2101.
41. YANKOFSKY, S. A., and SPIEGELMAN, S., The identification of the ribosomal RNA cistron by sequence complementarity. II. Saturation of and competitive interaction at the RNA cistron, *Proc. Nat. Acad. Sc.*, 1962, **48**, 1466.
42. DAVERN, C. I., and MESELSON, M., The molecular conservation of ribonucleic acid during bacterial growth, *J. Mol. Biol.*, 1960, **2**, 153.
43. ZAMECNIK, P. C., and KELLER, E. B., Relation between phosphate energy donors and incorporation of labeled amino acids into proteins, *J. Biol. Chem.*, 1954, **209**, 337.
44. GRUNBERG-MANAGO, M., and OCHOA, S., Enzymatic synthesis and breakdown of polynucleotides: polynucleotide phosphorylase, *J. Am. Chem. Soc.*, 1955, **77**, 3165.
45. NIRENBERG, M. W., and MATTHAEI, J. H., The dependence of cell free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides, *Proc. Nat. Acad. Sc.*, 1961, **47**, 1588.
46. LENGYEL, P., SPEYER, J. F., and OCHOA, S., Synthetic polynucleotides and the amino acid code, *Proc. Nat. Acad. Sc.*, 1961, **47**, 1936.
47. LENGYEL, P., SPEYER, J. F., BASILIO, C., WAHBA, A. J., GARDNER, R. S., SZER, W., and OCHOA, S., Studies on the amino acid code: the ambiguity in translation, *in* Nucleic Acids, Structure, Biosynthesis, and Function, New Delhi, Council of Scientific and Industrial Research, 1965, 309.
48. NIRENBERG, M. W., JONES, O. W., LEDER, P., CLARK, B. F. C., SLY, W. S., and PESTKA, S., On the coding of genetic information, *Cold Spring Harbor Symp. Quant. Biol.*, 1963, **28**, 549.
49. BRIMACOMBE, R., TRUPIN, J., NIRENBERG, M., LEDER, P., BERNFIELD, M., and JAOUNI, T., RNA codewords and protein synthesis. VIII. Nucleotide sequences of synonym codons for arginine, valine, cysteine and alanine, *Proc. Nat. Acad. Sci.*, 1965, **54**, 954.
50. SÖLL, D., OHTSUKA, E., JONES, D. S., LOHRMANN, R., HAYATSU, H., NISHIMURA,

S., and KHORANA, H. G., Studies on polynucleotides. XLIX. Stimulation of the binding of aminoacyl-sRNA's to ribosomes by ribotrinucleotides and a survey of codon assignments for 20 amino acids, *Proc. Nat. Acad. Sci.*, 1965, **54,** 1378.

51. KHORANA, H. G., JACOB, T. M., MOON, M. W., NARANG, S. A., and OHTSUKA, E., Studies on polynucleotides. XLII. The synthesis of deoxyribopolynucleotides containing repeating nucleotide sequences. Introduction and general considerations, *J. Am. Chem. Soc.*, 1965, **87,** 2954.
52. NISHIMURA, S., JONES, D. S., and KHORANA, H. G., Studies on polynucleotides. XLVIII. The *in vitro* synthesis of a co-polypeptide containing two amino acids in alternating sequence dependent upon a DNA-like polymer containing two nucleotides in alternating sequence, *J. Mol. Biol.*, 1965, **13,** 302.
53. NISHIMURA, S., JONES, D. S., OHTSUKA, E., HAYATSU, H., JACOB, T. M., and KHORANA, H. G., Studies on polynucleotides. XLVII. The *in vitro* synthesis of homopeptides as directed by a ribopolynucleotide containing a repeating trinucleotide sequence. New codon sequences for lysine, glutamic acid and arginine, *J. Mol. Biol.*, 1965, **13,** 283.
54. JUKES, T. H., Coding triplets and their possible evolutionary implications, *Biochem. and Biophysic. Research Commun*, 1965, **19,** 391.
55. BERNFIELD, M. R., and NIRENBERG, M. W., RNA codewords and protein synthesis. The nucleotide sequences of multiple codewords for phenylalanine, serine, leucine and proline, *Science*, 1965, **147,** 479.
56. CRICK, F. H. C., Codon-anticodon pairing. The wobble hypothesis, Information Exchange Group No. 7, No. 14, 1965.
57. AMES, B. N., and MARTIN, R. G., Biochemical aspects of genetics: the operon, *Ann. Rev. Biochem.*, 1964, **33,** 235.
58. KIHO, Y., and RICH, A., A polycistronic messenger RNA associated with beta galactosidase induction, *Proc. Nat. Acad. Sc.*, 1965, **54,** 1751.
59. BENZER, S., and CHAMPE, S. P., A change from nonsense to sense in the genetic code, *Proc. Nat. Acad. Sc.*, 1962, **48,** 1114.
60. GAREN, A., and SIDDIQI, O., Suppression of mutations in the alkaline phosphatase structural cistron of *E. coli*, *Proc. Nat. Acad. Sc.*, 1962, **48,** 1121.
61. BRENNER, S., STRETTON, A. O. W., and KAPLAN, S., The genetic code: the "nonsense" triplets for chain termination and their suppression, *Nature*, 1965, **206,** 994.
62. WEIGERT, M. G., and GAREN, A., Base composition of nonsense codons in *E. coli*. Evidence from amino acid substitutions at a tryptophan site in alkaline phosphatase, *Nature*, 1965, **206,** 992.
63. WEIGERT, M. G., LANKA, E., and GAREN, A., Base composition of nonsense codons in *E. coli*. II. The N_2 nonsense codon, in press.
64. GANOZA, M. C., and NAKAMOTO, T., Studies on the mechanism of polypeptide chain termination in cell free extracts of *E. coli*, *Proc. Nat. Acad. Sc.*, 1966, **55,** 162.
65. TAKANAMI, M., and YAN, Y., The release of polypeptide chains from ribosomes in cell free amino acid incorporating systems by specific combinations

of bases in synthetic polyribonucleotides, *Proc. Nat. Acad. Sc.*, 1965, **54,** 1450.

66. MARCHIS-MOUREN, G., and LIPMANN, F., On the mechanism of acetylation of fetal and chicken hemoglobins, *Proc. Nat. Acad. Sc.*, 1965, **53,** 1147.
67. HAENNI, A. L., and CHAPEVILLE, F., The behaviour of acetylphenylalanine-soluble ribonucleic acid in polyphenylalanine synthesis, *Biochim. et Biophysica Acta*, 1966, **114,** 135.
68. WALLER, J. P., The NH_2-terminal residues of the proteins from cell free extracts of *E. coli*, *J. Mol. Biol.*, 1963, **7,** 483.
69. MARCKER, K., and SANGER, F., *N*-formyl methionyl-sRNA, *J. Mol. Biol.*, 1964, **8,** 835.
70. CLARK, B. F. C., and MARCKER, K. A., Coding response of *N*-formyl methionyl-sRNA to UUG, *Nature*, 1965, **207,** 1038.
71. WEBSTER, R. E., ENGELHARDT, D. L., and ZINDER, N. D., *In vitro* protein synthesis. Chain initiation. *Proc. Nat. Acad. Sc.*, 1966, **55,** 155.
72. ADAMS, J. M., and CAPECCHI, M. R., *N*-formylmethionyl-sRNA as the initiator of protein synthesis, *Proc. Nat. Acad. Sc.*, 1966, **55,** 147.
73. HERVÉ, G., and CHAPEVILLE, F., Incorporation of phenyllactic acid in terminal position during polyphenylalanine biosynthesis, *J. Mol. Biol.*, 1965, **13,** 757.
74. SALAS, M., SMITH, M. A., STANLEY, W. M., JR., WAHBA, A. J., and OCHOA, S., Direction of reading of the genetic message, *J. Biol. Chem.*, 1965, **240,** 3988.
75. THACH, R. E., CECERE, M. A., SUNDARARAJAN, T. A., and DOTY, P., The polarity of messenger translation in protein synthesis, *Proc. Nat. Acad. Sc.*, 1965, **54,** 1167.
76. SMITH, M. A., SALAS, M., STANLEY, W. M., JR., WAHBA, A. J., and OCHOA, S., Direction of reading the genetic message. II, *Proc. Nat. Acad. Sc.*, 1966, **55,** 141.
77. BREMER, H., KONRAD, M. W., GAINES, K., and STENT, G. S., Direction of chain growth in enzymic RNA synthesis, *J. Mol. Biol.*, 1965, **13,** 540.
78. MAITRA, U., and HURWITZ, J., The role of DNA in RNA synthesis. IX. Nucleoside triphosphate termini in RNA polymerase products, *Proc. Nat. Acad. Sc.*, 1965, **54,** 815.
79. WOESE, C., The genetic code, *Internat. Council Scient. Unions Rev.*, 1963, **5,** 210.
80. MATTHAEI, J. H., JONES, O. W., MARTIN, R. G., and NIRENBERG, M. W., Characteristics and composition of RNA coding units, *Proc. Nat. Acad. Sc.*, 1962, **48,** 666.
81. BRETSCHER, M. S., and GRUNBERG-MANAGO, M., Polyribonucleotide-directed protein synthesis using an *E. coli* cell free system, *Nature*, 1962, **195,** 283.
82. SZER, W., and OCHOA, S., Complexing ability and coding properties of synthetic polynucleotides, *J. Mol. Biol.*, 1964, **8,** 823.
83. YANOFSKY, C., and ST. LAWRENCE, P., Gene action, *Ann. Rev. Microbiol.*, 1960, **14,** 311.
84. BENZER, S., and CHAMPE, S. P., Ambivalent rII mutants of phage T_4, *Proc. Nat. Acad. Sc.*, 1961, **47,** 1025.
85. BRODY, S., and YANOFSKY, C., Suppressor gene alteration of protein primary structure, *Proc. Nat. Acad. Sc.*, 1963, **50,** 9.

86. WEIGERT, M. G., LANKA, E., and GAREN, A., Amino acid substitutions resulting from suppression of nonsense mutations. II. Glutamine insertion by the Su-2 gene, tyrosine insertion by the Su-3 gene, *J. Mol. Biol.*, 1965, **14,** 522.
87. STRETTON, A. O. W., and BRENNER, S., Molecular consequences of the amber mutation and its suppression, *J. Mol. Biol.*, 1965, **12,** 456.
88. NOTANI, G. W., ENGELHARDT, D. L., KONIGSBERG, W., and ZINDER, N. D., Suppression of a coat protein mutant of the bacteriophage f_2, *J. Mol. Biol.*, 1965, **12,** 439.
89. GALLUCCI, E., and GAREN, A., Suppressor genes for nonsense mutations. II. The Su-4 and Su-5 suppressor genes of *E. coli*, *J. Mol. Biol.*, 1966, **15,** 193.
90. GAREN, A., GAREN, S., and WILHELM, R. C., Suppressor genes for nonsense mutations. 1. The Su-1, Su-2, Su-3 genes of *E. coli*, *J. Mol. Biol.*, 1965, **14,** 167.
91. SIGNER, E. R., Gene expression in foreign cytoplasm, *J. Mol. Biol.*, 1965, **12,** 1.
92. CLARK, J. M., JR., CHANG, A. Y., SPIEGELMAN, S., and REICHMAN, M. E., The *in vitro* translation of a monocistronic message, *Proc. Nat. Acad. Sc.*, 1965, **54,** 1193.
93. ABEL, P., Evidence for the universality of the genetic code, *Cold Spring Harbor Symp. Quant. Biol.*, 1964, **29,** 185.
94. PROTASS, J. J., SPEYER, J. F., and LENGYEL, P., Amino acid code in *Alcaligenes faecalis*, *Science*, 1964, **143,** 1174.
95. ROTHMAN, F., and BYRNE, R., Fingerprint analysis of alkaline phosphatase of *Escherichia coli* K_{12}, *J. Mol. Biol.*, 1963, **6,** 330.
96. GAREN, A., personal communication.
97. CAPECCHI, M. R., and GUSSIN, G. N., Suppression *in vitro*. The identification of a serine-sRNA as a "nonsense" suppressor, *Science*, 1965, **149,** 417.
98. ENGELHARDT, D. L., WEBSTER, R. E., WILHELM, R. C., and ZINDER, N. D., *In vitro* studies on the mechanism of suppression of a nonsense mutation, *Proc. Nat. Acad. Sci.*, 1965, **54,** 1791.
99. PESTKA, S., MARSHALL, R., and NIRENBERG, M. W., RNA code words and protein synthesis. V. Effect of streptomycin on the formation of ribosome-sRNA complexes, *Proc. Nat. Acad. Sc.*, 1965, **53,** 639.
100. GRUNBERG-MANAGO, M., and DONDON, J., Influence of pH and s-RNA concentration on coding ambiguities, *Biochem. and Biophysic. Research Commun.*, 1965, **18,** 517.
101. DAVIES, J., GILBERT, W., and GORINI, L., Streptomycin, suppression and the code, *Proc. Nat. Acad. Sc.*, 1964, **51,** 883.
102. SPOTTS, C. R., and STANIER, R. Y., Mechanism of streptomycin action on bacteria: A unitary hypothesis, *Nature*, 1961, **192,** 633.
103. SPEYER, J. F., LENGYEL, P., and BASILIO, C., Ribosomal localization of streptomycin sensitivity, *Proc. Nat. Acad. Sc.*, 1962, **48,** 684.
104. FLAKS, J. G., COX, E. C., WITTING, M. L., and WHITE, J. R., Polypeptide synthesis with ribosomes from streptomycin resistant and dependent *E. coli*, *Biochem. Biophysic. Research Commun.*, 1962, **7,** 390.

105. DAVIES, J. E., Studies on the ribosomes of streptomycin sensitive and resistant strains of *E. coli*, *Proc. Nat. Acad. Sc.*, 1964, **51**, 659.
106. LEBOY, P. S., COX, E. C., and FLAKS, J. G., The chromosomal site specifying a ribosomal protein in *Escherichia coli*, *Proc. Nat. Acad. Sc.*, 1964, **52**, 1367.
107. GORINI, L., and KATAJA, E., Phenotypic repair by streptomycin of defective genotypes in *E. coli*, *Proc. Nat. Acad. Sc.*, 1964, **51**, 487.
108. DAVIES, J., GORINI, L., and DAVIS, B. D., Misreading of RNA code words induced by aminoglycoside antibiotics, *J. Mol. Pharmacol.*, 1965, **1**, 93.
109. MCCARTHY, B. J., and HOLLAND, J. J., Denatured DNA as a direct template for *in vitro* protein synthesis, *Proc. Nat. Acad. Sc.*, 1965, **54**, 880.
110. ANDERSON, W. F., GORINI, L., and BRECKENRIDGE, L., Role of ribosomes in streptomycin activated suppression, *Proc. Nat. Acad. Sc.*, 1965, **54**, 1076.
111. REID, P. J., ORIAS, E., and GARTNER, T. K., Ribosomal protein differences in a strain of *E. coli* carrying a suppressor of an ochre mutation, *Biochem. and Biophysic. Research Commun.*, 1965, **21**, 66.
112. DE WAARD, A., PAUL, A. V., and LEHMAN, I. R., The structural gene for deoxyribonucleic acid polymerase in bacteriophages T_4 and T_5, *Proc. Nat. Acad. Sc.*, 1965, **54**, 1241.
113. SPEYER, J. F., Mutagenic DNA polymerase, *Biochem. and Biophysic. Research Commun.*, 1965, **21**, 6.
114. KOCH, A. L., and MILLER, C., A mechanism for keeping mutations in check, *J. Theoret. Biol.*, 1965, **8**, 71.
115. SPEYER, J. F., personal communication.
116. SAGER, R., Streptomycin as a mutagen for nonchromosomal genes, *Proc. Nat. Acad. Sc.*, 1962, **48**, 2018.
117. LOFTFIELD, R. B., The frequency of errors in protein biosynthesis, *Biochem. J.*, 1963, **89**, 82.
118. BERG, P., and CHAMBERLIN, M., On the transcription of DNA sequences by RNA polymerase, *in* Cinquantenaire de la Société de Chimie Biologique, Paris, 1965, 185.
119. BERG, P., BERGMANN, F. H., OFENGAND, E. J., and DIECKMANN, M., The enzymic synthesis of aminoacyl derivatives of ribonucleic acid. I. The mechanism of leucyl-, valyl-, isoleucyl-, and methionylribonucleic acid formation, *J. Biol. Chem.*, 1961, **236**, 1726.
120. LOFTFIELD, R. B., and EIGNER, E. A., A soluble ribonucleic acid induced increase in the specificity of *Escherichia coli* isoleucine-activating enzyme, *J. Biol. Chem.*, 1965, **240**, P.C. 1482.
121. NORRIS, A. T., and BERG, P., Mechanism of aminoacyl RNA synthesis: Studies with isolated aminoacyl adenylate complexes of isoleucyl RNA synthetase, *Proc. Nat. Acad. Sc.*, 1964, **52**, 330.
122. COLVIN, J. R., SMITH, D. B., and COOK, W. H., The microheterogeneity of proteins, *Chem. Rev.*, 1954, **54**, 687.
123. RIFKIN, D. B., RIFKIN, M. R., and KONIGSBERG, W., The presence of two major hemoglobin components in an inbred strain of mice, *Proc. Nat. Acad. Sc.*, 1966, **55**, 586.

Discussion

Dr. Luria: Dr. Lengyel, you told us many wonderful things in this extremely illuminating summary, bringing together the most modern approaches.

It is no great task to read off coded molecules. Whatever the details of DNA replication, whether viral RNA is copied with the intermediate of a *minus* strain or not, however protein is read off RNA—it is clear that all these processes belong to the same kind of topological approach. That is, once you have a linear code, you use it, and you make the best you can of it. But one of the things that must have happened in evolution is that all these macromolecules must at some point have decided that they had better make a cell or perish. This has really been a very major step. In order to make a cell you need more than macromolecules with coded information. You must also have a selective continuity. By this I mean that once you have done something that works, you have to continue to do it even though the whole information for doing is not written anywhere.

In the same way that biologic evolution is separate from cultural evolution, I think that template-reading in evolution is separate from the steps in evolution that must have taken place when these coded macromolecules had to learn to make things for which they did not have all the knowledge written inside themselves. Typical of these are the two things we're going to be talking about this afternoon. First, how do you make a complex polysaccharide in which you have exact sequences and branches of repeat units of different types? There is probably nowhere written in the DNA or the RNA the formula for a polysaccharide. You have to try your luck and, if you have luck, then you must have a way to learn the trick. You have to have cultural information in order to continue to make a complex product whose blueprint is not written out.

Second, even if you can make a polysaccharide and, presumably, other products of that level of complexity, you then have to make a cell. In order to make a cell, you have to put all these things in order. In order to put the things in order, you have to have a layout. You have to pass from the linear information of the macromolecules to the 2- and 3-dimensional specificity that is found in the organization of a cell.

So the last two papers of the symposium are deliberately devoted to finding out how all these wonderful macromolecules have gotten together in order to make cells.