[14] Meselson, M., F. W. Stahl, and J. Vinograd, these PROCEEDINGS, **43**, 581 (1957).
[15] Szybalski, W., *Experientia*, **16**, 164 (1960).
[16] Zubay, G., *Nature*, **182**, 1290 (1958); Stent, G., *Adv. in Virus Res.*, **5**, 95 (1958); Sinsheimer, R. L., *J. Mol. Biol.*, **1**, 218 (1959); Rich, A., *Ann. N. Y. Acad. Sci.*, **81**, 709 (1959).
[17] Zalokar, M., *Exp. Cell Res.*, **19**, 559 (1960).

---

*SYNTHETIC POLYNUCLEOTIDES AND THE AMINO ACID CODE, II**

BY JOSEPH F. SPEYER, PETER LENGYEL, CARLOS BASILIO,† AND SEVERO OCHOA

DEPARTMENT OF BIOCHEMISTRY, NEW YORK UNIVERSITY SCHOOL OF MEDICINE

Communicated November 21, 1961

Previous studies[1] on the incorporation of amino acids into acid-insoluble products by an *Escherchia coli* system, in the presence of synthetic polyribonucleotides, have been continued. This paper presents additional results with previously used and new copolymers containing uridylic[2] and guanylic acid (UG); uridylic, adenylic, and cytidylic acid (UAC); uridylic, cytidylic, and guanylic acid (UCG); and uridylic, adenylic, and guanylic acid (UAG). Assuming a triplet code, code letters (although in an as yet unspecified sequence) can now be assigned to the following eleven amino acids: cysteine, histidine, isoleucine, leucine, lysine, phenylalanine, proline, serine, threonine, tyrosine, and valine.

Preparations and Methods.—These were as previously described.[1] The following new polymers were prepared with polynucleotide phosphorylase. Poly UC (3:1), from a mixture of UDP and CDP in molar ratio 3:1; poly UG (5:1), from UDP and GDP in molar ratio 5:1; poly UAC, from UDP, ADP, and CDP in molar ratio 6:1:1; poly UCG and poly UAG, from mixtures of the appropriate nucleoside 5′-diphosphates in molar ratio of 6 of UDP to 1 of each of the other two. Their sedimentation coefficients ($s_{20,w}$) were as follows. UC (3:1), 11.9; UG, 3.4; UAC, 2.4; UCG, 5.0; UAG, 4.8 S.

Results.—Experiments with poly UG: In the previous work, the following amino acids were found to be incorporated into acid-insoluble products with use of various polynucleotides: phenylalanine with poly U (cf. also footnote 3); phenylalanine, leucine, serine, and proline with poly UC; and phenylalanine, isoleucine, and tyrosine with poly UA.

As shown in Table 1, poly UG promoted, as expected, the incorporation of phenylalanine and added two more amino acids, valine and cysteine, to the list. In contrast to poly UA, poly UG had no effect on the incorporation of isoleucine and tyrosine and, contrary to poly UC, it had no influence on the incorporation of serine and proline (cf. previous paper[1] and Table 2). Poly U, poly UC, and poly UA had no effect on the incorporation of valine and cysteine.

As further shown in Table 1, poly UC (3:1) promoted the incorporation of serine and leucine to about the same extent (an average of 0.75 mμmole/mg ribosomal protein). The incorporation ratio phenylalanine/serine (or leucine) was 2.2/0.75 = 2.9. On the assumption of a triplet code, the ratio of UUU to UUC (or UCU, or CUU) triplets in this polymer would be 3. This result is in good agreement with the results reported in this and the previous paper with poly UC (5:1).

TABLE 1

AMINO ACID INCORPORATION IN *E. coli* SYSTEM WITH VARIOUS POLYNUCLEOTIDES*

Amino acid	Polynucleotide† None	U	UC (5:1)	UC (3:1)	UA (5:1)	UG (5:1)
Phenylalanine	0.06	18.0	5.0	2.2	3.4	1.7
Serine	0.04	...	1.3	0.9	...	0.02
Tyrosine	0.02	...	...	...	...	0.03
Leucine	0.06	0.2?	0.8	0.6	0.4?	0.3?
Isoleucine	0.02	...	...	...	...	0.03
Proline	0.01	...	...	...	...	0.01
Valine	0.02	0.02	0.01	0.02	0.04	[0.3]
Cysteine	0.01	0.05	0.01	0.01	0.01	[0.4]

* mμmoles/mg ribosomal protein. Values for new amino acids incorporated (i.e. valine and cysteine) are enclosed in rectangles. 19 amino acids were tested individually in the case of poly UG; those not listed gave negative results. The values for valine and cysteine with poly UG are averages of duplicate runs.
† Amount as in the previous paper[1] except for poly UC (5:1), 0.03 μmole (as mononucleotide).

Experiments with poly UAC, UCG, and UAG: The results of these experiments are shown in Table 2. Since poly UAC caused definite incorporation of three amino acids not previously reported, i.e. lysine, threonine, and histidine, and we had observed some stimulation of the incorporation of lysine and threonine by poly UA and UC, respectively, in earlier experiments, the effect of these polymers was reinvestigated. The results are also shown in Table 2. Clearly lysine incorporation was stimulated by poly UA and threonine incorporation by poly UC. The incorporation of histidine was stimulated by poly UAC and by no other polymer. There was no stimulation of the incorporation of new amino acids by poly UCG or UAG. However, contrary to poly UAC, these polymers had low activity for phenylalanine. The reason for this difference is unknown.

TABLE 2

AMINO ACID INCORPORATION IN *E. coli* SYSTEM WITH VARIOUS POLYNUCLEOTIDES*

Amino acid	Polynucleotide† None	UC (5:1)	UA (5:1)	UAC	UCG	UAG
Phenylalanine	0.07	**7.0**	**3.0**	**7.5**	**1.0**	**1.1**
Serine	0.10	...	...	**2.0**	**0.3**	0.1
Tyrosine	0.03	...	...	**1.9**	0.06	**0.25**
Leucine	0.09	...	...	**3.3**	**0.4**	0.4?
Isoleucine	0.04	...	...	**1.2**	0.1?	**0.4**
Proline	0.04	**0.4**	...	**0.2**	**0.1**	0.04
Lysine	0.04	0.04	[0.2]	**0.2**	0.04	**0.1**
Threonine	0.10	[0.5]	0.1	**0.8**	0.15	0.15
Valine	0.03	...	...	0.07	**0.2**	**0.25**
Cysteine	0.10	...	...	0.1	**0.5**	**0.45**
Histidine	0.05	0.05	0.05	[0.25]	0.03	0.02

* mμmoles/mg ribosomal protein. Amino acids not included in this table gave negative results. Values for new amino acids incorporated (i.e. lysine, threonine, and histidine) are enclosed in rectangles. Boldface type is used to underline polynucleotide stimulation of the incorporation of a given amino acid. Most values are averages either of duplicate runs or of two separate experiments.
† Amount as in the previous paper.[1]

Each of the new copolymers (UAC, UCG, UAG) promoted incorporation of several amino acids in the range expected from its nucleotide composition in the light of experience with poly UC, UA, and UG. Thus, poly UAC stimulated the incorporation of phenylalanine, serine, tyrosine, leucine, isoleucine, proline, lysine, threonine, and histidine; poly UCG that of phenyalanine, serine, leucine, proline,

valine, and cysteine; and poly UAG that of phenylalanine, tyrosine, isoleucine, lysine, valine, and cysteine.

The leucine and isoleucine problem: It was seen in the previous paper[1] that all three polymers U, UC, and UA stimulated to a varying extent the incorporation of leucine and isoleucine. However, highest incorporation of leucine was promoted by poly UC, whereas poly UA caused the highest incorporation of isoleucine. If one takes maximal stimulation as the meaningful result, and on the basis of the incorporation ratio phenylalanine/leucine and phenylalanine/isoleucine with poly UC (5:1) and UA (5:1), respectively (which was close to 5 in either case), the triplet code letter for leucine would contain 2U and 1C and that for isoleucine 2U and 1A.

The reason for stimulation of isoleucine incorporation by poly UC has since been explained as due to contamination with leucine of our C^{14}-isoleucine sample. Addition of an excess of highly purified cold leucine to the incubation mixture completely eliminated the incorporation of radioactivity from C^{14}-isoleucine caused by poly UC but had no effect on that brought about by poly UA.

We have not as yet been able to explain the overlapping results with leucine which, if due to contamination, should be caused by multiple contamination with several amino acids. Experiments with highly purified C^{14}-leucine would provide an answer to this question. As seen in Tables 1 and 2, all of the polymers investigated promoted the incorporation of this amino acid to a greater or lesser degree. In order to indicate that only the stimulation of leucine incorporation by U- and C-containing polymers is considered meaningful, question marks have been placed next to the leucine values obtained with other polymers. The small stimulation of isoleucine incorporation by poly UCG, shown with a question mark in Table 2, is now explained.

Code letter assignments: A summary of the amino acids incorporated into acid-insoluble products by the *E. coli* system, in the presence of each of the polyribonucleotides so far investigated, is given in Table 3. With use of the quantitative data in Tables 1 and 2 of this paper and Table 4 of the previous one,[1] and assuming a triplet code for the sake of simplicity, code letter assignments can be made for each of eleven amino acids. They are listed in Table 4. Except for phenylalanine, the nucleotide sequence of the code letters is still unknown.

TABLE 3

SUMMARY OF AMINO ACIDS INCORPORATED BY *E. coli* SYSTEM WITH VARIOUS POLYNUCLEOTIDES

Amino acid	Polynucleotide U	UC	UA	UG	UAC	UCG	UAG
Cysteine	..	..	..	+	..	+	+
Histidine	..	..	..	..	+	..	..
Isoleucine	..	..	+	..	+	..	+
Leucine*	..	+	..	..	+	+	..
Lysine	..	..	+	..	+	..	+
Phenylalanine	+	+	+	+	+	+	+
Proline	..	+	..	..	+	+	..
Serine	..	+	..	..	+	+	..
Threonine	..	+	..	..	+	..	..
Tyrosine	..	..	+	..	+	..	+
Valine	..	..	..	+	..	+	+

* As noted in the text, leucine incorporation was stimulated to a greater or lesser extent by all of the above polymers. Although this overlapping is still unexplained, it is assumed for the present that only stimulation of the incorporation of this amino acid by U- and C-containing polymers is meaningful (see text).

TABLE 4

TRIPLET CODE LETTERS FOR AMINO ACIDS*

Amino acid	Code letter†
Cysteine	2U 1G
Histidine	1U 1A 1C
Isoleucine	2U 1A
Leucine	2U 1C
Lysine	1U 2A
Phenylalanine	UUU
Proline	1U 2C
Serine	2U 1C
Threonine	1U 2C
Tyrosine	2U 1A
Valine	2U 1G

* From data for *E. coli* system from this and the previous[1] paper.
† Sequence unknown except for phenylalanine.

The assignments of leucine and isoleucine have been discussed in the preceding section. Those for serine, tyrosine, and proline are supported by evidence presented in the previous paper[1] and by the additional data in Tables 1 and 2 of the present paper. Of these, only the assignment for proline calls for further comment.

As shown in the previous paper, proline incorporation was very weakly stimulated by poly C but, in marked contrast to this, effectively promoted by poly UC (5:1). If maximal stimulation is taken as the meaningful result, the code letter for proline would contain both U and C. The phenylalanine/proline incorporation ratio with poly UC (5:1) was 7/0.6 or about 12. In the present work (Table 2), a ratio of 7/0.4, or about 18, was obtained. Thus, the phenylalanine/proline ratio was well above the calculated value of 5 for UUU/2U 1C although below that of 25 for UUU/1U2C. This result suggests 1U 2C as the triplet code letter for proline and this assignment is further supported by amino acid replacement data to be discussed below. The small stimulation by poly C might be a consequence of contamination with UDP of the CDP used for preparation of this polymer.

The 1U 2C assignment for proline was already proposed in the previous paper without completely excluding CCC. It was conceivable that a poly C-mediated synthesis of polyproline might be largely missed if this polypeptide were fairly soluble in trichloroacetic acid. Although there was no support for this view in the literature,[4] it was decided to check our results by an independent method. For this purpose, phenylalanine and proline labeled with C^{14} in the carboxyl group were used. Following incubation as previously described,[1] the reaction mixtures were acidified with formic acid and treated with ninhydrin. Suitable aliquots were then plated and evaporated to dryness under an infrared lamp. Under these conditions, the carboxyl group of free amino acids is lost as CO_2 and any radioactivity remaining in the residue is a measure of peptide bond formation. When applied to incubation mixtures with phenylalanine-1-C^{14} and poly U, this method gave results in good agreement with those obtained by the usual measurement of radioactivity in acid-insoluble products, indicating extensive formation of polyphenylalanine, but no formation of polyproline was detected in parallel experiments with proline-1-C^{14} and poly C. We feel that these results definitely exclude CCC as the code letter for proline.

The assignments for valine and cysteine are based on the data of Table 1. Of all the polymers in this table, only poly UG (5:1) caused significant stimulation of the incorporation of these amino acids. Since the polymer promoted approximately equal incorporation of valine and cysteine one may take 0.35 mμmole /mg ribosomal protein as the average value for each. This gives a phenylalanine/valine (or cysteine) incorporation ratio of 1.7/0.35 = 4.85 in good agreement with the calculated value of 5 for UUU/2U 1G. Equally straightforward is the assignment for

histidine, since, as seen in Table 2, only poly AUC promoted the incorporation of this amino acid. The incorporation ratio phenylalanine/histidine was 7.5/0.25 = 30 (or 7.5/0.2 = 37.5, when corrected for the blank), in good agreement with the calculated frequency ratio UUU/UAC = 36 for a polymer prepared from UDP, ADP, and CDP in molar ratio 6:1:1.

The incorporation of lysine was significantly stimulated by poly UA (5:1) with a phenylalanine/lysine ratio of 3/0.2 = 15 (Table 2). This value, although lower than the calculated value of 25 for the UUU/1U 2A frequency ratio in this polymer, is well above the theoretical value of 5 for the UU/2U 1A frequency ratio. The assignment of a 1U 2A code letter to lysine appears therefore justified. As regards threonine, its incorporation (Table 2) was stimulated by poly UC (5:1) to approximately the same extent as that of proline, hence the assignment 1U 2C. Stimulation of the incorporation of this amino acid by poly UAC was abnormally high (expected phenylalanine/threonine ratio 36; found, about 10) and may have been due to contamination of our C^{14}-threonine sample with other amino acids.

Discussion.—It was pointed out in the previous paper that the replacement of a proline by a leucine residue in a nitrous acid mutant (No. 171) of tobacco mosaic virus, studied by Tsugita and Fraenkel-Conrat,[5] is in line with the proposed code letters, 1U 2C and 2U 1C respectively, for these amino acids. This replacement would be explained by conversion of one C of the proline code letter to one U. Another replacement observed in the same mutant,[5] that of threonine by serine, can now be explained in the same way. In a recent review, Tsugita[6] reported an interesting new replacement in a nitrous acid mutant (No. 167), that of serine by phenylalanine, which is in line with the code letters 2U 1C and UUU, respectively. One other replacement (nitrous acid mutant No. 167) is that of glutamine by valine.[6] Since our code letter for valine is 2U 1G. this replacement leads to the prediction that 1U 1C 1G is the code letter for glutamine. We did not detect stimulation of glutamine incorporation by poly UCG. However, as already pointed out, both this polymer and poly UAG brought about a relatively small incorporation of phenylalanine. Since the calculated UUU/1U 1C 1G frequency ratio for poly UCG is 36, the expected incorporation of glutamine promoted by this polymer would be 1/36 = 0.03, i.e. of the order of magnitude of the blank values in the absence of added polymer.

Summary.—An *E. coli* system, consisting of high-speed supernatant and ribosomes supplemented with transfer RNA, can incorporate several amino acids into acid-insoluble products in the presence of synthetic copolymers, prepared with polynucleotide phosphorylase, containing two or more kinds of nucleotides with a relatively high proportion of uridylic acid residues. Whereas poly U promotes the incorporation of phenylalanine only, poly UC stimulates that of phenylalanine, serine, leucine, proline, and threonine; poly UA that of phenylalanine, tyrosine, isoleucine, and lysine; poly UG that of phenylalanine, valine, and cysteine; poly UAC that of phenylalanine, serine, leucine, tyrosine, isoleucine, proline, threonine, lysine, and histidine; poly UCG that of phenylalanine, serine, leucine, proline, valine, and cysteine; and poly UAG that of phenylalanine, tyrosine, isoleucine, lysine, valine, and cysteine.

Overlapping results were obtained with leucine, whose incorporation was stimulated to a greater or lesser extent by all of the above polynucleotides. However,

maximal stimulation was caused in general by uridylic and cytidylic acid–containing copolymers. Stimulation by other polymers may be due to contamination with several amino acids of the C^{14}-leucine used and is assumed not to be meaningful.

From the observed incorporation of an amino acid relative to that of phenylalanine with a given polymer and the calculated frequency ratio of UUU triplets to other triplets in this polymer, triplet code letters (of as yet unknown nucleotide sequence except for UUU in the case of phenylalanine) have been assigned to eleven amino acids. The proposed code letters are in excellent agreement with amino acid replacement data in nitrous acid mutants of tobacco mosaic virus.

Note added in proof: The assignment of the following additional code letters will be shown in a subsequent article in these PROCEEDINGS: arginine 1U 1C 1G, glycine 1U 2G, and tryptophan 1U 2G. These were obtained by increasing the sensitivity of the testing method.

Our thanks are due to R. C. Warner and Jo-Yun Chen for the sedimentation cofficients of the synthetic polynucleotides, to Yoshito Kaziro for calling our attention to the review by Tsugita[6] and for the help with the Japanese text, and to Albert Lenny for skillful technical assistance. We are much indebted to W. H. Stein, The Rockefeller Institute, New York, for highly purified samples of L-isoleucine and L-phenylalanine.

* Aided by grants from the National Institute of Arthritis and Metabolic Diseases (grant A-1845) of the U.S. Public Health Service and from The Rockefeller Foundation.

† International Postdoctoral Fellow of the National Institutes of Health, U.S. Public Health Service. Permanent address: Instituto de Química Fisiológica y Patológica, Universidad de Chile, Santiago, Chile.

[1] Lengyel, P., J. F. Speyer, and S. Ochoa, these PROCEEDINGS, **47**, 1936 (1961).

[2] Abbreviations: The capital letters A, U, C, and G are used for the nucleotides adenylic, uridylic, cytidylic, and guanylic acid, respectively, or their corresponding residues in polynucleotide chains: ADP, UDP, CDP, and GDP, the 5′-diphosphates of adenosine, uridine, cytidine, and guanosine.

[3] Nirenberg, M. W., and J. H. Matthaei, these PROCEEDINGS, **47**, 1558 (1961).

[4] Berger, A., J. Kurtz, and E. Katchalski, *J. Am. Chem. Soc.*, **76**, 552 (1954).

[5] Tsugita, A., and H. Fraenkel-Conrat, these PROCEEDINGS, **46**, 636 (1960).

[6] Tsugita, A., *Protein, Nucleic Acid, Enzyme (Tokyo)*, **6**, 385 (1961).

INDUCTION OF ALTERED GLOBIN SYNTHESIS IN HUMAN IMMATURE ERYTHROCYTES INCUBATED WITH RIBONUCLEOPROTEIN, †*

BY AUSTIN S. WEISBERGER

DEPARTMENT OF MEDICINE, WESTERN RESERVE UNIVERSITY, SCHOOL OF MEDICINE, CLEVELAND, OHIO

Communicated by John H. Dingle, November 13, 1961

It is now generally assumed that information governing protein synthesis is present in deoxyribonucleic acid (DNA) and that this information is mediated by ribonucleic acid (RNA). It is currently postulated that information is recorded in the specific sequence of bases in DNA and RNA and that this arrangement of bases controls the sequential arrangement of amino acids during protein synthesis, thereby conferring specificity.[1, 2] Studies in the field of microbial genetics have demon-