[16] Yčas, M., and W. S. Vincent, *Exp. Cell Research*, **21**, 513 (1960).
[17] Zalokar, M., *Exp. Cell Research*, **19**, 559 (1960).
[18] Jacob, F., and J. Monod, *J. Mol. Biol.*, **3**, 318 (1961).
[19] Bishop, J., G. Favelukes, R. Schweet, and E. Russell, *Nature*, **191**, 1365 (1961).

---

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

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

DEPARTMENT OF BIOCHEMISTRY, NEW YORK UNIVERSITY SCHOOL OF MEDICINE

Communicated December 20, 1961

Previous work[1, 2] on the incorporation of amino acids into acid-insoluble products by a cell-free *Escherichia coli* system in the presence of synthetic polyribonucleotides led to the assignment of triplet nucleotide code letters (of as yet unknown sequence) for eleven amino acids. Since out of 64 triplets in RNA only six, corresponding to A + C + G permutations,[3] were not represented in the polymers tested, it seemed that code letters for more than eleven amino acids should be present among the 58 triplets covered by those polymers and that some of them might stimulate the incorporation of additional amino acids if it were possible to increase the over-all activity of the system.

It may be remembered that, of the various polymers tried, poly U accounted for one amino acid (phenylalanine), poly UC and UA for seven (isoleucine, leucine, lysine, proline, serine, threonine, and tyrosine), and poly UG and UAC for three (cysteine, histidine, and valine), and that poly UCG and UAG did not stimulate the incorporation of additional amino acids. Moreover, the activity of the system with poly UG, UCG, and UAG, as reflected by phenylalanine incorporation, was rather low. Accordingly, these polymers were selected for further study.

It has now been possible to boost the activity of the incorporation system with poly UG or poly UCG eight- to ten-fold by increasing the amount of (*a*) polymer, (*b*) transfer RNA, and (*c*) cold amino acids. At the same time, the sensitivity of the assay was augmented by increasing the specific radioactivity of the C^{14}-amino acid tested. In this way, three more amino acids, arginine, glycine, and tryptophan, have been added to the list. Their code letters are 1U 1C 1G, 1U 2G, and 1U 2G, respectively.

Preparations and Methods.—DL-asparagine-2,3-C^{14} was obtained from the New England Nuclear Corporation, Boston, Mass. All other preparations were as previously described.[1, 2] Poly UG (5:1) and poly UCG (6:1:1) were from the same batch of polymers used in the experiments of the preceding paper.[2] The experimental procedure followed closely that already described[1] except that the mixture containing ribosomes and supernatant (each with 5 mg protein) was preincubated for 30 instead of 15 min at 37° with no cold amino acids present and that the final reaction mixture contained transfer RNA in saturating amounts (1.3 to 2.5 mg, as assayed for phenylalanine incorporation in the presence of poly U);

poly UG or poly UCG, 80 instead of 20 μg previously employed; and a mixture of 19 cold L-amino acids (each 0.044 instead of 0.009 μmoles previously present).

Results.—Table 1 shows the results of one out of two closely agreeing experiments. It may be seen that, under the new experimental conditions, the activity of the system both with poly UG and poly UCG was much higher than previously. Under the present conditions, poly UCG stimulated the incorporation of arginine and both poly UG and UCG stimulated that of glycine and tryptophan. Subtracting the "blanks," i.e. the incorporation values in the absence of added polymer, the following incorporation ratios were obtained: Phenylalanine/arginine with UCG, 10.4/0.35 = 30; phenylalanine/glycine, with UG 13.2/0.55 = 24, with UCG 10.4/0.26 = 40; phenylalanine/tryptophan, with UG 13.2/0.67 = 20, with UCG 10.4/0.43 = 24. Since the UUU/1U 2G frequency ratio in poly UG (5:1) is 25 and the UUU/1U 1C 1G frequency ratio in poly UCG (6:1:1) is 36, the results justify the code letter assignments shown in the last column of Table 1.

TABLE 1

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

	Polynucleotide			
Amino acid	None	UG (5:1)	UCG (6:1:1)	Code letter†
Phenylalanine	0.18	13.4	10.6	UUU
Arginine	0.12	0.04	0.47	1U 1C 1G
Glycine	0.19	0.74	0.45	1U 2G
Tryptophan	0.03	0.70	0.46	1U 2G

* mμmoles/mg ribosomal protein. † Sequence unknown.

Negative results were obtained with alanine, aspartic acid, glutamic acid, asparagine, glutamine, and methionine. In the preceding paper,[2] we had ventured the prediction that the code letter for glutamine would be 1U 1C 1G. This was based on the replacement of glutamine by valine (code letter 2U 1G) reported by Tsugita[4] in a nitrous acid mutant of tobacco mosaic virus. However, no stimulation of glutamine incorporation was obtained in two separate experiments with poly UCG. In these as in previous experiments, glutamine gave high "blank" values. High blanks were also obtained with asparagine in the present experiments. Further investigation showed that the blank incorporation of glutamine-C^{14} is ribonuclease-insensitive and consequently cannot be due to incorporation of the amide into a polypeptide chain. The reason for our negative results with glutamine and poly UCG is unexplained. They might be due to the presence of inhibitory impurities or to deficiency of glutamine-activating enzyme and/or glutamine-specific transfer RNA in our system.

Conclusion.—Triplet code letter assignments, although of unknown sequence, are now available for fourteen amino acids: arginine, cysteine, glycine, histidine, isoleucine, leucine, lysine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, and valine. This leaves alanine, aspartic acid, asparagine, glutamic acid, glutamine, and methionine still to be accounted for.

Note added in proof: Evidence for the assignment of the following additional code letters will be given in a subsequent publication in these PROCEEDINGS: Alanine, 1U 1C 1G; asparagine, 1U 1A 1A; methionine, 1U 1A 1G; aspartic and glutamic, each probably 1U 1A 1G. There are some indications for degeneracy in the asparagine (1U 1A 1A and 1U 1A 1C) and threonine (1U 1C 1C and 1U 1A 1C) code letters.

* Aided by grants from the National Institute of Arthritis and Metabolic Diseases (grant A-1845) of the U.S. Public Health Service and 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] Speyer, J. F., P. Lengyel, C. Basilio, and S. Ochoa, these PROCEEDINGS, **48**, 63 (1962).

[3] Abbreviations as in the preceding paper.

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

---

*FURTHER SEQUENCES IN THE γ CHAIN OF HUMAN FETAL HEMOGLOBIN**

BY W. A. SCHROEDER, J. ROGER SHELTON, JOAN BALOG SHELTON, AND JEAN CORMICK

DIVISION OF CHEMISTRY AND CHEMICAL ENGINEERING,† CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated by Linus Pauling, December 26, 1961

Sequences which accounted for about 60 per cent of the γ chain of human fetal hemoglobin have previously been reported.[1] We report here a sequence of the γ chain in which only a few residues are not definitely placed.

Experimental.—The experimental procedure has followed and extended that previously described.[1] The most extensive information has been obtained from the peptides of tryptic, chymotryptic, and peptic hydrolysates of the γ chain itself and of a chymotryptic hydrolysate of oxidized γ chain. In the determination of sequence of the individual peptides, extensive use has been made of the Edman degradation in a modified form of the paper strip method.[2]

Results and Discussion.—The soluble tryptic peptides of the γ chain account for little more than 50 per cent of its 146 residues. The insoluble portion exhibits considerable heterogeneity. Much reliance, therefore, has been placed on chymotryptic and peptic peptides not only to provide the usual overlap between tryptic peptides but also to provide largely the sequence of the insoluble portion. The available evidence suggests the sequence of the γ chains as shown on page 285.

Inasmuch as the α chains of human adult and fetal hemoglobin probably are identical in sequence, the β and γ chains must play analogous roles in the two molecules. It is of interest to compare their primary structures and to discuss possible effects on secondary and tertiary structure.

The γ chain contains 146 residues, as does the β chain.[3] The difference in amino acid composition of the β and γ chains may now be listed; relative to the β chain, the γ chain contains fewer of these residues—four alanyl, one cysteinyl, two histidyl, one leucyl, three prolyl, one tyrosyl, and five valyl—and more of these residues—one glutaminyl, four isoleucyl, one lysyl, one methionyl, six seryl, three threonyl, and one tryptophyl. Thus, in amino acid composition, 17 residues have been replaced by 17 others.

In the above sequence of the γ chain, the residues within parentheses are placed in their most probable sequence. There are 38 differences in the sequence of the β and γ chains on the basis of the present results and the sequences of the β chain as