RIBONUCLEOTIDE COMPOSITION OF THE GENETIC CODE

Robert G. Martin, J. Heinrich Matthaei*, Oliver W. Jones and Marshall W. Nirenberg

Laboratory of Biochemistry and Metabolism, National Institute of Arthritis and Metabolic Diseases, National Institutes of Health, Bethesda, Maryland

Received December 4, 1961

We have developed a cell-free E. coli amino acid incorporating system dependent upon the addition of messenger (template) polyribonucleotides (Nirenberg and Matthaei, 1961, a, b). Thus this system affords a sensitive assay for messenger RNA. For example, polyuridylic acid specifically directed the synthesis of polyphenylalanine thereby demonstrating that one or more uridylic acid residues in polyuridylic acid is the coding unit for phenylalanine We have also shown that phenylalanine linked to soluble RNA is an intermediate in this process (Nirenberg, Matthaei and Jones, 1961, c). These findings established a reasonable experimental approach for deciphering the genetic code using synthetic polyribonucleotides of known constitution.

We have now studied amino acid incorporation directed by a series of randomly-mixed polynucleotides. The purpose of this communication is to report these results.

The polynucleotides were synthesized enzymatically using purified M. lysodeikticus polynucleotide phosphorylase (Singer and Guss) and 5.0μmoles of UDP/ 1.0 μmole of ADP, CDP, or GDP for poly UA**, UC and UG respectively. 5 μmoles of UDP/ 0.5 μmoles each of ADP, CDP or GDP were used to synthesize

* NATO Postdoctoral Research Fellow

** The following abbreviations are used: Poly UA, polyuridylic-adenylic acid; poly UC, polyuridylic-cytidylic acid; poly UG, polyuridylic-guanylic acid; poly UAC, polyuridylic-adenylic-cytidylic acid; poly UCG, polyuridylic-cytidylic-guanylic acid; poly UGA, polyuridylic-guanylic-adenylic acid; ADP, adenosine diphosphate; UDP, uridine diphosphate; GDP, guanosine diphosphate; CDP, cytidine diphosphate.

poly UAC, poly UCG and poly UGA. Materials and methods for measuring incorporation of C^{14}- amino acids into protein were described previously (Nirenberg and Matthaei, 1961 b). All results are based on at least two experiments each done in duplicate.

A strikingly specific stimulation of amino acid incorporation into protein due to the addition of synthetic polynucleotides was found (Table 1). The nucleotides (letters of the genetic code) corresponding to fifteen amino acids (words of the code) can be specified from these data and are summarized in Table 2. Some of these results agree with recent data cited by Dr. S. Ochoa and co-workers.+ Poly U stimulated the incorporation of phenylalanine; poly UA, tyrosine, isoleucine and lysine; poly UC, serine, proline and leucine; poly UG, leucine, valine, cysteine, tryptophan, methionine and glycine; poly UGC, arginine, serine, glutamic acid and alanine. If a polynucleotide containing two bases stimulated incorporation of an amino acid, inclusion of a third base in the polynucleotide did not prevent this stimulation.

Poly C stimulated the incorporation of small amounts of proline (Nirenberg and Matthaei, 1961 b) and similar results were obtained with poly UC (1/1 ratio, see Table 1). Lysine incorporation was stimulated by the addition of poly UA (1/4 ratio) but not by randomly-mixed poly UA containing lower proportions of A. The coding unit for lysine therefore contains U < A. The stimulation of phenylalanine incorporation by poly UA (1/4 ratio) was negligible, so these data have not been expressed as the ratio of lysine to phenylalanine incorporation.

The data of Table 1 demonstrate that a coding unit corresponding to leucine can contain either U and C or U and G. Since two "words" containing different nucleotides correspond to leucine, at least part of the code is degenerate when synthetic polynucleotides are used to direct amino acid incorporation. The data of Table 1 suggest that the coding units for other amino acids also may be degenerate. It is possible that a coding unit may contain other bases

+ We thank Dr. Ochoa for sending us manuscripts describing similar experiments (c.f. Ref.)

Table 1

Amino Acid Incorporation into Protein
Stimulated by Synthetic Polynucleotides*

C^{14}- Amino Acid	Polynucleotides					
	UA	UC	UG	UAC	UGC	UGA
Phenylalanine	100	100	100	100	100	100
Valine	0.6	0.7	69	0.4	50	60
Leucine	4.9	14	67	5.1	64	44
Cysteine	4.9	4.5	53	0	36	33
Tryptophan	7.1	7.5	94	8.6	42	36
Glutamic Acid	1.5	1.2	0	1.2	4.9	-
Methionine	0.6	0	13	0.6	5.5	11
Glycine	4.7	0	21	0.5	9	1.4
Arginine	0	0	5.5	0	13	2.9
Alanine	1.9	0.2	1.6	1.0	15	0.9
Serine	0.4	5.0 14**	8.4	3.6	28	6.2
Proline	0	1.4 101‡	0	0	3.7	0
Tyrosine	13	0	0	1.4	0	8.6
Isoleucine	12	1.0	1.0	4.8	4.4	8.4

* The figures represent the percent of any amino acid incorporated compared to phenylalanine incorporation $\left(\frac{\text{m}\mu\text{moles amino acid incorporated}}{\text{m}\mu\text{moles phenylalanine incorporated}} \times 100\right)$
The components of the reaction mixtures were described previously (Nirenberg and Matthaei, 1961 b). 0.2 μmole each of 19 L-amino acids minus the appropriate C^{14}-amino acid, 0.150 μmole of the C^{14}-amino acid and approximately 25 μg of each polynucleotide were added/ml of reaction mixture. Samples were incubated at 37° for 15 minutes. Incorporation of C^{14}-phenylalanine in counts per minute due to the addition of polynucleotides UA, UC, UG, UAC, UGC, and UGA were 731, 745, 714, 804, 2144 and 2744 respectively.
The reproducibility of the above percentage figures was ± 3.

** 10 μmoles UDP/ 1.0 μmole CDP were used to synthesize this polynucleotide. 670 counts per minute of C^{14}-phenylalanine were incorporated due to the addition of this polymer.

‡ 1.0 μmole of UDP/ 1.0 μmole CDP were used to synthesize this polynucleotide. 343 counts per minute of C^{14}-phenylalanine were incorporated due to the addition of this polymer.

in addition to or replacing the ones indicated in Table 2. The proposed correlations between nucleotides and amino acids are in excellent agreement with mutant amino acid replacement data (Yčas, 1958).

Table 2

Genetic Code for Fifteen Amino Acids

Amino Acid	Nucleotide Composition of Coding Unit *
Phenylalanine	UUU...
Valine	UG (U > G)
Leucine	UG, UC, (U > G) (U > C)
Cysteine	UG
Tryptophan	UG (U ≲ G)
Glutamic Acid	UGC
Methionine	UG
Glycine	UG
Arginine	UGC
Alanine	UGC
Serine	UC, UGC, (U > C) (U > G or C)
Proline	UC (U < C)
Tyrosine	UA (U > A)
Isoleucine	UA
Lysine	UA (U < A)

* The order of the nucleotides in a coding unit is not specified.

The stimulation of valine and leucine incorporation compared to phenylalanine by UG polymers of different base-ratios was directly proportional to the increase in G content of the polymer, whereas, incorporation of tryptophan was an exponential function of G content.* These data strongly suggest that the coding unit for tryptophan is G≫U.

The coding units for arginine and alanine each contain three different nucleotides. Assuming that all amino acids have the same coding ratio, these data rule out the possibility of singlet and doublet codes. The minimum

* Unpublished data, Nirenberg, Martin, Matthaei and Jones.

coding ratio must be three; very possibly it is larger. It should be noted that the sequential order of the nucleotides in any coding unit, with the exception of UUU... cannot be specified at this time. Further conclusions and coding constraints derived from these and additional data will be discussed in a subsequent communication.

Summary. Randomly-mixed polynucleotides were found to direct the incorporation of different amino acids into protein. Using this technique the ribonucleotide composition of the RNA code corresponding to fifteen amino acids was determined. A minimum coding ratio of three nucleotides per amino acid was demonstrated. Two coding units corresponding to leucine were found; thus a part of the code was shown to be degenerate.

Acknowledgement

The authors wish to express their appreciation for the valuable technical assistance of Mrs. Linda Greenhouse. They also thank Drs. Leon Heppel and Maxine Singer for invaluable advice.

References

Lengyel, P., Speyer, J. F. and Ochoa, S. Proc. Nat. Acad. Sci., In Press

(a) Nirenberg, M. W. and Matthaei, J. H., V Internat. Cong. of Biochem., Moscow, Aug. 9-16 (1961)

(b) Nirenberg, M. W. and Matthaei, J. H., Proc. Nat. Acad. Sci. 47, 1588 (1961)

(c) Nirenberg, M. W., Matthaei, J. H. and Jones, O. W., Proc. Nat. Acad. Sci., In Press

Singer, M. F. and Guss, J. K., J. Biol. Chem., In Press.

Speyer, J. F., Lengyel, P., Basilio, C. and Ochoa, S., Proc. Nat. Acad. Sci., In Press

Yčas, M. in "Symposium on Information Theory in Biology." Pergamon Press, New York (1958)