4　遗传密码

（中文内容暂缺）

equipment, and to Dr. G. E. R. Deacon and the captain and officers of R.R.S. *Discovery II* for their part in making the observations.

[1] Young, F. B., Gerrard, H., and Jevons, W., *Phil. Mag.*, **40**, 149 (1920).
[2] Longuet-Higgins, M. S., *Mon. Not. Roy. Astro. Soc., Geophys. Supp.*, **5**, 285 (1949).
[3] Von Arx, W. S., Woods Hole Papers in Phys. Oceanog. Meteor., **11** (3) (1950).
[4] Ekman, V. W., *Arkiv. Mat. Astron. Fysik.* (*Stockholm*), **2** (11) (1905).

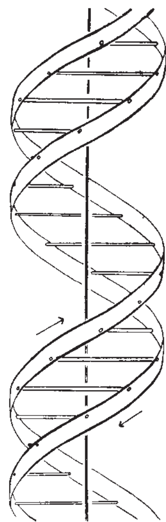
MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

WE wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.



This figure is purely diagrammatic. The two ribbons symbolize the two phosphate—sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate diester groups joining β-D-deoxyribofuranose residues with 3′,5′ linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's[2] model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There is a residue on each chain every 3·4 A. in the z-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 A. The distance of a phosphorus atom from the fibre axis is 10 A. As the phosphates are on the outside, cations have easy access to them.

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical z-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

It has been found experimentally[3,4] that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

The previously published X-ray data[5,6] on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F. Wilkins, Dr. R. E. Franklin and their co-workers at

King's College, London. One of us (J. D. W.) has been aided by a fellowship from the National Foundation for Infantile Paralysis.

J. D. WATSON
F. H. C. CRICK

Medical Research Council Unit for the
Study of the Molecular Structure of
Biological Systems,
Cavendish Laboratory, Cambridge.
April 2.

[1] Pauling, L., and Corey, R. B., *Nature*, **171**, 346 (1953); *Proc. U.S. Nat. Acad. Sci.*, **39**, 84 (1953).

[2] Furberg, S., *Acta Chem. Scand.*, **6**, 634 (1952).

[3] Chargaff, E., for references see Zamenhof, S., Brawerman, G., and Chargaff, E., *Biochim. et Biophys. Acta*, **9**, 402 (1952).

[4] Wyatt, G. R., *J. Gen. Physiol.*, **36**, 201 (1952).

[5] Astbury, W. T., Symp. Soc. Exp. Biol. 1, Nucleic Acid, 66 (Camb. Univ. Press, 1947).

[6] Wilkins, M. H. F., and Randall, J. T., *Biochim. et Biophys. Acta*, **10**, 192 (1953).

Molecular Structure of Deoxypentose Nucleic Acids

WHILE the biological properties of deoxypentose nucleic acid suggest a molecular structure containing great complexity, X-ray diffraction studies described here (cf. Astbury[1]) show the basic molecular configuration has great simplicity. The purpose of this communication is to describe, in a preliminary way, some of the experimental evidence for the polynucleotide chain configuration being helical, and existing in this form when in the natural state. A fuller account of the work will be published shortly.

The structure of deoxypentose nucleic acid is the same in all species (although the nitrogen base ratios alter considerably) in nucleoprotein, extracted or in cells, and in purified nucleate. The same linear group of polynucleotide chains may pack together parallel in different ways to give crystalline[1-3], semi-crystalline or paracrystalline material. In all cases the X-ray diffraction photograph consists of two regions, one determined largely by the regular spacing of nucleotides along the chain, and the other by the longer spacings of the chain configuration. The sequence of different nitrogen bases along the chain is not made visible.

Oriented paracrystalline deoxypentose nucleic acid ('structure *B*' in the following communication by Franklin and Gosling) gives a fibre diagram as shown in Fig. 1 (cf. ref. 4). Astbury suggested that the strong 3·4-A. reflexion corresponded to the internucleotide repeat along the fibre axis. The ~ 34 A. layer lines, however, are not due to a repeat of a polynucleotide composition, but to the chain configuration repeat, which causes strong diffraction as the nucleotide chains have higher density than the interstitial water. The absence of reflexions on or near the meridian immediately suggests a helical structure with axis parallel to fibre length.

Diffraction by Helices

It may be shown[5] (also Stokes, unpublished) that the intensity distribution in the diffraction pattern of a series of points equally spaced along a helix is given by the squares of Bessel functions. A uniform continuous helix gives a series of layer lines of spacing corresponding to the helix pitch, the intensity distribution along the nth layer line being proportional to the square of J_n, the nth order Bessel function. A straight line may be drawn approximately through

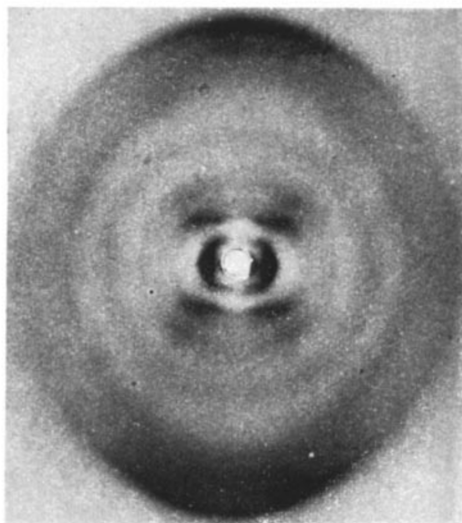


Fig. 1. Fibre diagram of deoxypentose nucleic acid from *B. coli*. Fibre axis vertical

the innermost maxima of each Bessel function and the origin. The angle this line makes with the equator is roughly equal to the angle between an element of the helix and the helix axis. If a unit repeats n times along the helix there will be a meridional reflexion (J_0^2) on the nth layer line. The helical configuration produces side-bands on this fundamental frequency, the effect[5] being to reproduce the intensity distribution about the origin around the new origin, on the nth layer line, corresponding to *C* in Fig. 2.

We will now briefly analyse in physical terms some of the effects of the shape and size of the repeat unit or nucleotide on the diffraction pattern. First, if the nucleotide consists of a unit having circular symmetry about an axis parallel to the helix axis, the whole diffraction pattern is modified by the form factor of the nucleotide. Second, if the nucleotide consists of a series of points on a radius at right-angles to the helix axis, the phases of radiation scattered by the helices of different diameter passing through each point are the same. Summation of the corresponding Bessel functions gives reinforcement for the inner-

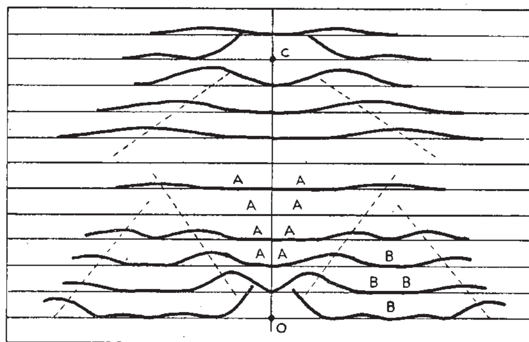


Fig. 2. Diffraction pattern of system of helices corresponding to structure of deoxypentose nucleic acid. The squares of Bessel functions are plotted about 0 on the equator and on the first, second, third and fifth layer lines for half of the nucleotide mass at 20 A. diameter and remainder distributed along a radius, the mass at a given radius being proportional to the radius. About *C* on the tenth layer line similar functions are plotted for an outer diameter of 12 A.

PNAS
Proceedings of the National Academy of Sciences of the United States of America www.pnas.org

On the Impossibility of all Overlapping Triplet Codes in Information Transfer from Nucleic Acid to Proteins

S. Brenner

PNAS 1957;43;687-694
doi:10.1073/pnas.43.8.687

This information is current as of March 2007.

	This article has been cited by other articles: www.pnas.org#otherarticles
E-mail Alerts	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.
Rights & Permissions	To reproduce this article in part (figures, tables) or in entirety, see: www.pnas.org/misc/rightperm.shtml
Reprints	To order reprints, see: www.pnas.org/misc/reprints.shtml

Notes:

Proceedings of the

NATIONAL ACADEMY OF SCIENCES

Volume 43 · Number 8 · August 15, 1957

(ED

ON THE IMPOSSIBILITY OF ALL OVERLAPPING TRIPLET CODES IN INFORMATION TRANSFER FROM NUCLEIC ACID TO PROTEINS

BY S. BRENNER

MEDICAL RESEARCH COUNCIL UNIT FOR THE STUDY OF THE MOLECULAR STRUCTURE OF BIOLOGICAL SYSTEMS, CAVENDISH LABORATORY, CAMBRIDGE, ENGLAND

Communicated by G. Gamow, June 10, 1957

It is a generally accepted view that nucleic acids control the synthesis of proteins, and it has been proposed more specifically that the sequence of amino acids in a polypeptide chain is determined by the order of nucleotides in ribo- or deoxyribonucleic acid. The problem of how this determination is effected has come to be known as the "coding" problem. The formal aspects of this problem can be investigated theoretically, and most of the work done in this field has recently been reviewed by Gamow, Rich, and Yčas.[1]

Since there are only four different nucleotides in RNA or DNA to determine twenty different amino acids, it is clear that more than one nucleotide must be used to code for each amino acid. Most codes have been constructed on the basis that each amino acid is determined by a set of three nucleotides. Such triplet codes, however, have an excess of information, since there are sixty-four different triplets for the twenty amino acids. In Gamow's original diamond code,[2] several triplets, chosen in a particular way, coded for any given amino acid; the code was therefore "degenerate." This code was also of the overlapping type—that is, the number of nucleotides in the nucleic acid was equal to the number of amino acids in the polypeptide chain. Gamow's diamond code does not, in fact, code for known sequences, and the same is true for the major-minor code, another overlapping triplet code, invented by L. Orgel.[1] These are, however, only two examples of a large number of possible codes of this type which can be obtained by choosing different ways of degenerating the triplets. To test all of these systematically is clearly impossible, and hence it is necessary to have some general theorem about such codes.

The general overlapping triplet code has the following properties.

(i) The coding triplets are chosen from four nucleotides, A, B, C, and D, giving sixty-four different triplets.

(ii) Coding is overlapping, each triplet sharing two nucleotides with the succeeding triplet in a sequence. Thus the sequence $ABCDA$ codes for three amino acids: ABC for the first, BCD for the second, and CDA for the third.

(iii) An amino acid may be represented by more than one triplet; that is, the sixty-four triplets are degenerated into twenty sets.

Since any dipeptide sequence is represented by a sequence of four nucleotides, there cannot be more than 256 different dipeptides. On the other hand, if all dipeptide sequences were possible, 400 would be expected. Thus overlapping codes introduce restrictions in amino acid sequences. The number of dipeptide sequences known is less than 256, and, although statistical studies have suggested that all dipeptides are likely to be found, the significance of this result has been difficult to assess.[1] The sample of proteins studied is highly selected, a large number of sequences are fragmentary, and the methods used to study sequences further bias the data.

However, sufficient sequences are known to prove that it is impossible to code them with overlapping triplets. The proof is simple and does not depend on any special way of degenerating the triplets. It consists in the demonstration that sixty-four triplets are insufficient to code the known sequences.

Proof: Since successive triplets share two nucleotides in common, any given triplet can be preceded by only four different triplets and succeeded by only four different triplets. In an amino acid sequence *j.k.l.*, we call *j* an N-neighbor, and *l* a C-neighbor, of *k*. For every four different N-neighbors (or C-neighbors) or part thereof, *k* must have one triplet assigned to it. Thus the *minimum* number of triplet representations for each amino acid can be counted from a table of neighbors.

The available sequences are given in the Appendix. From these sequences a grid is constructed and the different neighbors counted for each amino acid. The number of triplets assigned to each amino acid is based on the larger number of its neighbors. These data are given in Table 1, from which it can be seen that seventy

TABLE 1

Amino Acid	C-Neighbors	N-Neighbors	Minimum No. of Triplets Required	Amino Acid	C-Neighbors	N-Neighbors	Minimum No. of Triplets Required
Lys	18	17	5	Pro	13	12	4
Ser	17	13	5	Tyr	12	10	3
Gly	15	15	4	Glu	11	11	3
Leu	15	15	4	Glun	12	9	3
Cys	15	14	4	Asp	10	11	3
Arg	14	16	4	Asn	9	10	3
Ala	14	15	4	Ileu	9	9	3
Val	14	12	4	His	6	9	3
Thr	13	14	4	Met	5	7	2
Phe	13	14	4	Try	3	3	1
						Total	70

triplets would be required to code the sequences. We conclude, then, that all overlapping triplet codes are impossible.

This result has one important physical implication. The original formulation of overlapping codes was based on the similarity of the internucleotide distance in DNA to the spacing between amino acid residues in an extended polypeptide chain. It was supposed that each amino acid was spatially related in a one-to-one way with each nucleotide on a nucleic acid template. The present result shows that this cannot be so and that each amino acid is stereochemically related to at least two, if not three, nucleotides, depending on whether coding is partially overlapping or nonoverlapping. The difficulties raised by this can easily be overcome by assuming that the polypeptide sequence is in contact with the nucleic acid template only at the growing point, and detailed schemes can be readily proposed.

As far as the coding problem is concerned, it now appears that all amino sequences are likely to be found and that it will not be possible to effect a "decoding" by discovering restrictions in sequences. The nonoverlapping of triplets implies that there must be some way of determining which triplets in a sequence are coding triplets and which are not, and a very interesting code has recently been proposed by Crick, Griffith, and Orgel,[3] in which this problem is dealt with in a novel manner.

APPENDIX

AMINO ACID SEQUENCES

In writing the sequences, the same conventions used by Gamow *et al.*[1] have been followed. Wherever doubt exists as to whether glutamic acid is present as such (glu) or as the amide (glun), it has been assigned as "glux," and the same rule has been followed for aspartic acid and asparagine. All the longer lysozyme sequences suggested by Thompson (*Biochem. J.*, **60,** 507; **61,** 253, 1955) have been omitted, since some of these appear to be incorrect when compared with those established by the French workers. Sequences established by carboxypeptidase digestions alone are given at the end of the list but are omitted from the grid. The same applies to the pepsin sequence of Williamson and Passmann (*J. Biol. Chem.*, **222,** 151, 1956), as there are conflicting reports about the N-terminal group (Van Vunakis and Herriott, *Biochim. et Biophys. Acta,* **23,** 60, 1957).

The grid (Table 2) shows the number of times dipeptide sequences are found. Identical sequences from the closely related proteins vasopressin and oxytocin and corticotrophin and melanophore-stimulating hormone are only recorded once. Dipeptide sequences from lysozyme are not recorded if the same sequence is found in a longer peptide. When both glu and glun are absent, glux is counted as a neighbor, and the same rule is followed for asp, asn, and ax.

SEQUENCES USED IN THE GRID*

-thr. gly. ileu[e]-
-thr. ser. ileu[c, d]-
-ala. gly. val[b]-

Insulin A: Gly. ileu. val. glu. glun. cys. cys. ala. ser. val[a]. cys. ser. leu. tyr. glun. leu. glu. asn. tyr. cys. asn.[4–6]

Insulin B: Phe. val. asn. glun. his. leu. cys. gly. ser. his. leu. val. glu. ala. leu. tyr. leu. val. cys. gly. glu. arg. gly. phe. phe. tyr. thr. pro. lys. ala.[4, 5]

Oxytocin: Cys. tyr. ileu. glun. asn. cys. pro. leu. gly. NH_2.[7]

-arg[a]-

Vasopressin: Cys. tyr. phe. glun. asn. cys. pro. lys[b]. gly. NH_2.[8]

Corticotrophin: Ser. tyr. ser. met. glu. his. phe. arg. try. gly. lys. pro. val. gly. lys. lys. arg-arg. pro. val. lys. val. tyr. pro. *asp. gly. ala. glu.*[9] asp. glun. leu. ala[b]. glu. ala. phe. pro. leu-
-glux. ala. ser[c]-
glu. phe.[10–13]

Glucagon: His. ser. glun. gly. thr. phe. thr. ser. asp. thr. ser. lys. tyr. leu. asp. ser. arg. arg. ala. glun. asp. phe. val. glun. try. leu. met. asn. thr.[14]

-ser[a]-

Melanophore-stimulating hormone: Asp. glu[b]. gly. pro. tyr. lys. met. glu. his. phe. arg. try. gly. ser. pro. pro. lys. asp.[15–17]

* [a] Cattle. [b] Pig. [c] Sheep. [d] Whale. [e] Horse. [f] Salmon. [g] Chicken. [h] Man.

TABLE 2

Grid Showing Numbers of N- and C- Neighbors

C-Neighbors	N-Neighbors: Lys	Ser	Gly	Leu	Cys	Arg	Ala	Val	Thr	Phe	Pro	Tyr	Glu	Glun	Glux	Asp	Asn	Asx	Ileu	His	Met	Try
Lys	1	1	3	1	2	1	4	2	2	2	4	2	1	1	1	1	..	1	..	1	2	..
Ser	1	..	4	1	2	..	3	1	5	..	..	1	2	1	1	3	1	1	..	1	..	..
Gly	3	4	..	2	4	4	2	2	2	1	1	1	1	1	..	1	..	1	..	..	..	1
Leu	2	3	3	1	..	2	3	2	1	..	2	2	..	3	..	..	1	1	1	3	..	1
Cys	1	1	2	2	1	2	2	5	1	1	..	2	..	2	..	..	1	2	1	..	..	..
Arg	3	5	1	1	1	2	1	..	..	2	1	2	1	1	..	3	2	..	1	..	..	1
Ala	1	2	1	2	4	1	4	3	2	..	1	..	4	1	1	1	3	2	1	..	..	..
Val	2	3	2	4	1	1	..	..	1	2	3	1	..	..	..	..	1	..	4	..	..	..
Thr	1	2	1	1	1	..	2	1	1	1	..	2	1	..	..	1	2	1	..	1	..	..
Phe	2	1	1	..	1	1	1	1	1	1	1	1	1	..	..	1	..	..	..	1	..	..
Pro	2	3	2	1	1	1	..	..	4	1	1	2	..	..	..	..	..	..	1	1	..	..
Tyr	2	2	1	2	2	..	1	2	..	1	1	..	..	..	..	..	1	1	..	..	..	..
Glu	1	1	1	2	1	2	2	3	..	1	..	..	..	..	..	2	..	..	..	..	1	..
Glun	..	2	..	..	..	..	2	3	..	2	..	2	1	..	..	1	3	..	1	..	..	..
Glux	..	1	..	..	1	..	1	2	1	1	1	..	..	..	..	1	..	1	1	..	..	..
Asp	2	1	1	1	..	..	..	1	1	2	1	..	1	1	..	3	..	..	..	..	..	..
Asn	2	..	..	1	3	1	..	1	2	..	1	..	1	1	..	..	..	..	..	..	1	..
Asx	1	2	1	..	2	1	..	3	..	..	..	..	..	..	1	..	..	1	1	..	1	..
Ileu	1	1	2	..	1	1	..	..	..	..	1	3	..	..	1	..	..	1	..	..	..	..
His	1	1	..	..	1	1	..	1	..	1	..	..	1	1	..	..	..	..	1	..	..	..
Met	1	1	1	1	..	..	1	..	..	..	..	..	..	1	..	..	..	..	..	..	1	..
Try	..	..	..	..	..	1	1	..	..	..	..	..	..	1	..	..	..	..	..	..	..	..

-ileu[e]-

Hypertensin: Asp. arg. val. tyr. val[a]. his. pro. phe. his. leu.[18, 19]

-ser[g]-

Cytochrome c: -val. glun. lys. cys. ala[a, b, e, f]. glun. cys. his. thr. val. glu. lys-[20, 21]

Trypsinogen: Val. asp. asp. asp. asp. lys. ileu. val. gly-[22, 23]

Ribonuclease: Lys. glu. thr. ala. ala. ala. lys. phe. glun. arg. glu-[24–26]

-tyr. cys. asn. glun. met. met. lys. ser. arg. asn. leu. thr. lys. asp. arg. cys.[24–27]

-lys. asn. val. ala. cys. lys. asn. thr-[26, 27]

-cys. asn. arg. glu. ser. thr. ser. gly. lys. tyr. pro. asn. ala. cys. tyr. lys. thr. thr. asn. glun. ala. lys. his-[26, 27]

-tyr. glun. ser. tyr-[24]

-phe. asp. ala. ser. val.[24, 28, 29]

Lysozyme: -lys. asx-[30]

-arg. his. lys-[31]

Lys. val. phe. gly. arg-

-ala. lys. phe. glux-

-asx. tyr. arg. gly-

-arg. gly. tyr. ileu. leu-

-asn. ala. tyr. gly. ser. leu. asn-

-thr. pro-

-leu. pro-[32]

-asn. arg-

-ileu. arg-

-thr. pro. gly. ser. arg-

-val. ala. try. arg-[33]

-gly. cys. arg. leu.

-phe. glu. ser. phe. asp. glu. ala. thr. asp. arg-

-cys. glu. ala. leu. ala. ala. met. lys. arg-[34]

-ala. ala-	-asx. ileu-	-ileu. arg-	-ser. ala-
-ala. leu-	-asx. leu-	-ileu. asx-	-ser. arg-
-ala. lys-	-cys. ala-	-ileu. val-	-ser. asx-
-ala. met-	-cys. arg-	-leu. ala-	-ser. leu-
-arg. asx-	-cys. asx-	-leu. cys-	-ser. val-
-arg. cys-	-cys. glux-	-leu. leu-	-thr. ala-
-arg. gly-	-cys. ileu-	-lys. gly-	-thr. asx-
-arg. leu-	-cys. lys-	-met. asx-	-thr. gly-
-asx. ala-	-gly. leu-	-met. lys-	-thr. glux-
-asx. arg-	-gly. met-	-phe. asx-	-thr. pro-
-asx. asx-	-glux. ala-	-phe. glux	-val. ala-
-asx. gly-	-glux. leu-		-val. asx-
			-val. cys-
			-val. glux-[35]

Ovalbumin: -ala. gly. val. asx. ala. ala-[36]

-asx. ser. glux. ileu. ala-

-glux. ser. ala-[37]

-cys. ala-	-thr. cys-	-cys. phe-
-cys. gly-	-val. cys-	-gly. cys-
-cys. val-	-cys. glux-	-asx. cys-[38]
-ser. cys-	-phe. cys-	
	-val. ser. pro.[28]	

Papain: Ileu. pro. glux-[39]
-ser. asx- -cys. asx- -val. cys-[40]
-asx. cys- -cys. gly. asx-

Hemoglobin: Val. glun. leu-
Val. leu-[41]
-lys. arg- -arg. leu- -val. lys-
-lys. leu- -phe. lys- -arg. lys-
-ser. arg- -arg. phe- -tyr. arg-
-ala. arg- -leu. arg- -phe. arg-[42]

Myoglobin: Gly. leu-[43, 44]

γ-Globulin: Ala. leu. val. asx. glux-[45, 46]

β-Lactoglobulin: -val. glux- -thr. lys- -ala. lys-
-val. leu- -lys. gly- -leu. lys-
-asx. lys- -pro. lys- -phe. lys-[47]
-glux. lys- -lys. pro-

Carboxypeptidase: Asn. ser-
-ser. thr-[48]
-thr[h]-

Serum albumin: Asx. ala[a]-[49]

Lactogenic hormone: Thr. pro. val. thr. pro-[50]

Tobacco mosaic virus: -thr. ser. gly. pro. ala. thr.[51]
Pro. ileu. glux-[52]

Casein: Lys. leu. val. ala. glux. asx-[53]
-leu. gly- -ser. pro- -ser. leu-[47]

Chymotrypsinogen: -gly. leu. ser. arg. ileu. val-[54, 55]
-tyr. thr. asn. ala-[56]
-gly. asp. ser. gly-[57]

SEQUENCES NOT USED IN THE GRID

Serum albumin: -gly. val. ala. leu[h].[58]
-ser. val. thr. leu. ala. ala[a].[58]

Actin: -his. ileu. phe.[59]

Tropomyosin: -ala. ileu. met. thr. ser. ileu.[59]

Pepsin: Leu. gly. asp. asp. his. glu-[60] (cf.[61])

[1] G. Gamow, A. Rich, and M. Yčas, *Advances in Biol. and Med. Physics*, Vol. **4** (New York, Academic Press, Inc., 1955).
[2] G. Gamow, *Nature*, **173**, 318, 1954; *Kgl. Danske Videnskab. Selskab Biol. Medd.*, **22**, 3, 1954.
[3] F. H. C. Crick, J. S. Griffith, and L. E. Orgel, these PROCEEDINGS, **43**, 416, 1957.
[4] A. P. Ryle, F. Sanger, L. F. Smith, and R. Kitai, *Biochem. J.*, **60**, 541, 1955.
[5] H. Brown, F. Sanger, and R. Kitai, *Biochem. J.*, **60**, 556, 1955.
[6] J. I. Harris, F. Sanger, and M. A. Naughton, *Arch. Biochem. and Biophys.*, **65**, 427, 1956.
[7] V. Du Vigneaud, C. Ressler, and S. Trippett, *J. Biol. Chem.*, **205**, 949, 1953.
[8] E. A. Popenoe and V. Du Vigneaud, *J. Biol. Chem.*, **207**, 563, 1954.
[9] Different sequences for the italicized amino acids have been reported by White and Land-

mann and Li *et al.* (see nn. 12–13 below) in pig and sheep corticotrophin, but these are doubtful (see Howard *et al.* [n. 11 below]).

[10] P. H. Bell, *J. Am. Chem. Soc.*, **76**, 5566, 1954.

[11] K. S. Howard, R. G. Shepherd, E. A. Eigner, D. S. Davies, and P. H. Bell, *J. Am. Chem. Soc.*, **77**, 3420, 1955.

[12] W. F. White and W. A. Landmann, *J. Am. Chem. Soc.*, **77**, 1712, 1955.

[13] C. H. Li, I. I. Geschwind, R. D. Cole, D. Raacke, J. I. Harris, and J. S. Dixon, *Nature*, **176**, 687, 1955.

[14] W. W. Bromer, L. G. Sinn, A. Straub, and O. K. Behreus, *J. Am. Chem. Soc.*, **78**, 3858, 1956.

[15] J. I. Harris and P. Roos, *Nature*, **177**, 527, 1956.

[16] I. I. Geschwind, C. H. Li, and L. Barnafi, *J. Am. Chem. Soc.*, **78**, 4494, 1956.

[17] I. I. Geschwind, C. H. Li, and L. Barnafi, *J. Am. Chem. Soc.*, **79**, 1003, 1957.

[18] D. F. Elliott and W. S. Peart, *Nature*, **177**, 527, 1956.

[19] L. T. Skeggs, K. E. Lentz, J. R. Kahn, N. P. Shumway, and K. Woods, *J. Exptl. Med.*, **104**, 193, 1956.

[20] H. Tuppy and G. Bodo, *Monatsh.*, **85**, 1024, 1954.

[21] H. Tuppy and S. Paleus, *Acta Chem. Scand.*, **9**, 353, 1955.

[22] E. W. Davie and H. Neurath, *J. Biol. Chem.*, **212**, 515, 1955.

[23] P. Desnuelle and C. Fabre, *Biochim. et Biophys. Acta*, **18**, 49, 1955.

[24] C. H. W. Hirs, W. H. Stein, and S. Moore, *J. Biol. Chem.*, **221**, 151, 1956.

[25] R. R. Redfield and C. B. Anfinsen, *J. Biol. Chem.*, **221**, 385, 1956.

[26] C. H. W. Hirs, *Federation Proc.*, **16**, 196, 1957.

[27] C. B. Anfinsen and R. R. Redfield, *Advances in Protein Chem.*, **11**, 1, 1956.

[28] C. B. Anfinsen, *J. Biol. Chem.*, **221**, 405, 1956.

[29] C.-I. Niu and H. Fraenkel-Conrat, *J. Am. Chem. Soc.*, **77**, 5882, 1955.

[30] R. Acher, M. Justisz, and C. Fromageot, *Biochim. et Biophys. Acta*, **8**, 442, 1952.

[31] R. Acher, J. Thaureaux, C. Crocker, M. Justisz, and C. Fromageot, *Biochem. et Biophys. Acta*, **9**, 339, 1952.

[32] R. Acher, U.-R. Laurila, and C. Fromageot, *Biochim. et Biophys. Acta*, **19**, 97, 1956.

[33] J. Thaureaux and R. Acher, *Biochim. et Biophys. Acta*, **20**, 559, 1956.

[34] J. Thaureaux and P. Jolles, *Compt. rend. Acad. Sci.*, **243**, 1926, 1956.

[35] A. R. Thompson, *Biochem. J.*, **60**, 507; **61**, 253, 1955.

[36] M. Ottessen and A. Wollenberg, *Compt. rend. trav. Lab. Carlsberg.*, *Ser. chim.*, **28**, 463, 1953.

[37] M. Flavin, *J. Biol. Chem.*, **210**, 771, 1954.

[38] M. Flavin and C. B. Anfinsen, *J. Biol. Chem.*, **211**, 375, 1954.

[39] E. O. P. Thompson, *J. Biol. Chem.*, **207**, 863, 1954.

[40] J. R. Kimmel, E. O. P. Thompson, and E. G. Smith, *J. Biol. Chem.*, **217**, 151, 1955.

[41] R. R. Porter and F. Sanger, *Haemoglobin*, ed. F. J. W. Roughton and J. C. Kendrew (London: Butterworth & Co., Ltd.), 1949.

[42] P. Mäsiar, B. Keil, and F. Sorm, *Chem. Listy*, **51**, 352, 1957.

[43] H. Fraenkel-Conrat, *J. Am. Chem. Soc.*, **76**, 3606, 1954.

[44] V. M. Ingram, *Biochim. et Biophys. Acta*, **16**, 599, 1956.

[45] R. R. Porter, *Biochem. J.*, **46**, 473, 1950.

[46] M. L. McFadden and E. L. Smith, *J. Biol. Chem.*, **214**, 185, 1955.

[47] B. A. Askonas, P. N. Campbell, C. Godin, and T. S. Work, *Biochem. J.*, **61**, 105, 1955.

[48] E. O. P. Thompson, *Biochim. et Biophys. Acta*, **10**, 633, 1953.

[49] E. O. P. Thompson, *J. Biol. Chem.*, **208**, 565, 1954.

[50] R. D. Cole, I. I. Geschwind, and C. H. Li, *J. Biol. Chem.*, **224**, 399, 1957.

[51] C.-I. Niu and H. Fraenkel-Conrat, *Arch. Biochem. and Biophys.*, **59**, 538, 1955.

[52] G. Braunitzer, *Naturwissenschaften*, **42**, 371, 1955.

[53] N. Seno, K. Murai, and K. Shimura, *J. Biochem. (Japan)*, **42**, 699, 1955.

[54] W. J. Dreyer and H. Neurath, *J. Biol. Chem.*, **217**, 527, 1955.

[55] M. Rovery, M. Poilroux, A. Curnier, and P. Desnuelle, *Biochim. et Biophys. Acta*, **17**, 565, 1955.

[56] M. Rovery, M. Poilroux, A. Yoshida, and P. Desnuelle, *Biochim. et Biophys. Acta*, **23**, 608, 1957.

[57] N. K. Schaffer, L. Simet, S. Harshman, R. R. Engle, and R. W. Driske, *J. Biol. Chem.*, **225**, 197, 1957.

[58] W. F. White, J. Shields, and K. S. Robbins, *J. Am. Chem. Soc.*, **77**, 1267, 1955.

[59] R. H. Locker, *Biochim. et Biophys. Acta*, **14**, 533, 1954.

[60] M. B. Williamson and J. M. Passmann, *J. Biol. Chem.*, **222**, 151, 1956.

[61] H. Van Vunakis and R. M. Herriott, *Biochim. et Biophys. Acta*, **23**, 60, 1957.

INFECTION OF PROTOPLASTS BY DISRUPTED T2 VIRUS

BY JOHN SPIZIZEN*

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

Communicated by H. G. Wood, May 29, 1957

INTRODUCTION

The functions of the components of virus particles have been the subject of much investigation in two widely different virus systems. In the case of tobacco mosaic virus (TMV) it has recently been demonstrated that the ribonucleic acid isolated from the virus retains infectivity[1, 2] and possesses the property of genetic determination.[3] In contrast with the low infectivity of TMV, bacteriophage T2 can be studied with very high precision. The protein coat of T2 isolated after osmotic rupture has been found to adsorb to the receptor site on the cell wall of the host and to stop multiplication of the host cell.[4] After injection through the tail of T2 virus into the host cell, the deoxyribonucleic acid (DNA) component is found to be associated with the replicating activity of the virus.[5] However, no direct biological activity of the isolated DNA has been demonstrated. It would be of much interest to determine whether the biological activity of the protein moiety subsequent to attachment[4, 6] is a necessary prerequisite for the function of the DNA portion or, indeed, whether DNA can be active in the complete absence of protein.

Since cell wall components and the protein tip of T2 virus interact prior to injection of DNA, it was considered possible that the removal of the cell wall might permit the replicating moiety of the virus to enter the cytoplasm directly. Removal of cell walls can be achieved by lysozyme treatment of *Escherichia coli* B in hypertonic sucrose to produce forms known as "protoplasts."[7, 8] Intact T2 virus will not infect these protoplasts. We have attempted to infect protoplasts directly with DNA isolated from T2, without success. However, when osmotically disrupted preparations of T2 virus[9] were mixed with protoplasts of *E. coli* B, significant increases in the numbers of infective particles were observed. In fact, similar increases in infective units were obtained with protoplasts of a number of bacteria whose cells are normally resistant to infection. Although the osmotically shocked preparations employed in these studies contained all the components of the original virus, we have obtained evidence to indicate that the activity observed may be due to DNA units protected by a protein shell. This communication describes some of the data obtained on the interaction of protoplasts and disrupted T2 virus, as they relate to the functional requirements of virus-cell interaction.

[6] Rose, T. H., D. Neubert, and A. L. Lehninger, *J. Biol. Chem.* (in press).
[7] Lehninger, A. L., *Ann. N. Y. Acad. Sci.*, **86**, 484 (1960).
[8] Lehninger, A. L., *Physiol. Rev.* (in press).
[9] Neubert, D., *Fed. Proc.*, **20**, 146 (1961).
[10] Neubert, D., and A. L. Lehninger, *Biochem. Pharmacol.*, **8**, 126 (1961).
[11] Lehninger, A. L., *J. Biol. Chem.*, **234**, 2187, 2465 (1959).
[12] Ward, D. N., and R. Guillemin, *Proc. Soc. Exp. Biol. Med.*, **96**, 568 (1957).
[13] Oncley, J. L., E. Ellenbogen, D. Gitlin, and F. R. N. Gurd, *J. Phys. Chem.*, **52**, 88 (1952); Creeth, J. M., *Nature*, **170**, 210 (1952).
[14] Melhuish, A. H., and A. L. Greenbaum, *Biochem. J.*, **78**, 392 (1961).
[15] Tapley, D. F., *J. Biol. Chem.*, **222**, 325 (1956).
[16] Wojtczak, L., and A. L. Lehninger, *Biochim. et Biophys. Acta*, **51**, 442 (1961).
[17] Acher, R., *Ann. Rev. Biochem.*, **29**, 550 (1960).
[18] Li, C. H., *Adv. in Protein Chem.*, **11**, 101 (1956).
[19] Levine, R., and M. S. Goldstein, *Recent Progress Hormone Res.*, **11**, 343 (1955); Wick, A. N., and D. R. Drury, *Am. J. Physiol.*, **173**, 229 (1953); Klein, S. P., *Am. J. Physiol.*, **163**, 70 (1950); Kipnis, D. M., and C. F. Cori, *J. Biol. Chem.*, **224**, 681 (1957); Park, C. R., H. Kaji, M. Smith, D. Orth, and H. E. Morgan, *Abstr. 5th Internat. Congr. Biochem., Moscow* (1961), p. 263.
[20] Barrnett, R. J., and E. G. Ball, *Science*, **129**, 1282 (1959); Barrnett, R. J., and E. G. Ball, *J. Biophys. Biochem. Cytol.*, **8**, 83 (1960).
[21] Fuhrman, F. A., and H. H. Ussing, *J. Cell. Comp. Physiol.* , **38**, 109 (1951); Sawyer, W. H., *Am. J. Physiol.*, **164**, 44 (1951); Andersen, B., and H. H. Ussing, *Acta Physiol. Scand.*, **39**, 228 (1957); Morel, F., J. Maetz, and C. Lucarain, *Biochim. et Biophys. Acta*, **28**, 619 (1958).
[22] Sawyer, W. H., and R. M. Schisgall, *Am. J. Physiol.*, **187**, 312 (1956).
[23] Smith, H. W., *Fed. Proc.*, **11**, 701 (1952); Wirz, H., *Helv. Physiol. Acta*, **14**, 353 (1956); Gottschalk, C. W., *Circulation*, **21**, 861 (1960); Lassiter, W. E., C. W. Gottschalk and M. Mylle, *J. Clin. Invest.*, **39**, 1004 (1960).
[24] Fong, C. T. O., L. Silver, D. R. Christman, and I. L. Schwartz, these PROCEEDINGS, **46**, 1273 (1960); Rasmussen, H., I. L. Schwartz, M. A. Schoessler, and G. Hochster, *ibid.*, 1278 (1960); Rasmussen, H., M. A. Schoessler, L. Silver, and C. T. O. Fong, *ibid.*, 1288 (1960).
[25] Huggins, C. B., D. F. Tapley, and E. V. Jensen, *Nature*, **167**, 592 (1951); Jensen, E. V., *Science*, **130**, 1319 (1959).
[26] Cadenas, E., H. Kaji, C. R. Park, and H. Rasmussen, *J. Biol. Chem.*, **236**, PC63 (1961).
[27] Mirsky, I. A., and G. Perisutti, *Biochim. et Biophys. Acta*, **50**, 603 (1961).
[28] Peters, R. A., *Nature*, **177**, 426 (1956).
[29] Hechter, O., and G. Lester, *Recent Progress in Hormone Res.*, **16**, (1960).
[30] Lazarow, A., *Glutathione, A Symposium* (New York: Academic Press, 1954), p. 231.
[31] Mize, C. E., and R. G. Langdon, *Fed. Proc.*, **20**, 231 (1961); Mize, C. E., and R. G. Langdon, *J. Biol. Chem.* (in press).
[32] Langdon, R. G., *J. Biol. Chem.*, **235**, PC15 (1960).

---

*SYNTHETIC POLYNUCLEOTIDES AND THE AMINO ACID CODE**

BY PETER LENGYEL, JOSEPH F. SPEYER, AND SEVERO OCHOA

DEPARTMENT OF BIOCHEMISTRY, NEW YORK UNIVERSITY SCHOOL OF MEDICINE

Communicated October 25, 1961

The problem of coding in protein biosynthesis, i.e., of how a certain sequence of four different nucleotides in an RNA[1] chain can specify a given sequence of 20 different amino acids in a polypeptide chain, has been considered by several investigators in the last decade. Until now these studies have been either theoretical[2, 3] or statistical[4–6] in nature.

Increased knowledge of the mechanism of protein biosynthesis, the concept of messenger RNA as the actual template,[7] and especially the observation of Brenner *et al.*[8] (cf. also Gros *et al.*[9]) that after infection of *Escherichia coli* with T_2 phage the newly formed messenger RNA is added to preexisting ribosomes (on which most of the protein synthesis of the infected cell occurs) suggested the use of synthetic polyribonucleotides, with known nucleotide sequences, as messengers and a possible method of experimental approach to the nucleotide code. The simplest possible sequence is that in homopolynucleotides which, if active, should prescribe the formation of homopolypeptide chains. These homopolynucleotides as well as copolymers containing two or more different nucleotide species in predetermined ratios can be synthesized by polynucleotide phosphorylase.[10]

We wish to report that various kinds of synthetic polynucleotides are active as messengers in an *E. coli* system and indeed determine the incorporation of different amino acids into an acid-insoluble product. That is the case for phenylalanine, serine, and tyrosine with polymers containing uridylic acid, uridylic acid and cytidylic acid, and uridylic acid and adenylic acid respectively.

Recently Nirenberg and Matthaei[11] succeeded in obtaining a poly U-dependent synthesis of polyphenylalanine in a system of *E. coli* supernatant and ribosomes.

Preparations.—*E. coli* ribosomes were prepared from log phase cells grown on enriched media (0.5 per cent glucose, 0.5 per cent trypticase, 0.5 per cent yeast extract). The cells were washed twice with ice-cold 0.2 *M* KCl and once with 0.02 *M* Tris-HCl buffer, pH 7.9, containing 0.01 *M* $MgCl_2$ and 0.005 *M* mercaptoethylamine. They were then resuspended in 2 volumes of the same medium and disrupted for 3 min at maximum power in a 10 kc Raytheon oscillator at 5°. The suspension was centrifuged twice for 30 min at 3–4° and 30,000 *g* to remove debris and intact cells. The extract was centrifuged at low temperature for 2 hr in the preparative Spinco centrifuge at 105,000 *g*. The clear supernatant fluid was collected, the sediment containing the ribosomes was rinsed with Tris-HCl buffer, pH 7.9, containing 0.01 *M* $MgCl_2$, 0.02 *M* KCl, 0.2 *M* sucrose, and 0.5 per cent Lubrol (Imperial Chemical Industries) and suspended in this medium, with use of a glass homogenizer, to give a concentration of 20 mg of protein/ml. This suspension was layered over a solution containing 0.034 *M* Tris-HCl buffer, pH 7.9, 0.01 *M* $MgCl_2$, 0.6 *M* KCl, and 0.3 *M* sucrose, and again centrifuged for 3 hr at 105,000 *g*. The particles were then suspended in 0.1 *M* Tris-HCl buffer, pH 7.9, containing 0.01 *M* $MgCl_2$ and 0.02 *M* KCl, to give a concentration of 10 mg of protein/ml, frozen in dry ice, and kept in a freezer at −18°. The ribosomes are stable under these conditions. They can also be kept after lyophilization. Rat liver ribosomes were prepared by the method of Rendi and Hultin.[12] Rat liver supernatant was prepared by the method of Zamecnik and Keller.[13] For some experiments the precipitate obtained by bringing the *E. coli* supernatant to pH 5.0 was dissolved in 0.05 *M* Tris-HCl buffer, pH 7.9, containing 0.005 *M* mercaptoethylamine, and used in place of the supernatant. This treatment aimed at reducing the level of endogenous amino acids in the system. Transfer RNA (S-RNA) was prepared from *E. coli* supernatant as described by Hoagland *et al.*[14]

Polyribonucleotides were prepared with *Azotobacter vinelandii* polynucleotide phosphorylase (specific activity, 62) as previously described.[10] After precipitation with ethanol, the polymers were dissolved in 0.04 *M* NaCl–0.005 *M* sodium citrate and the solution was extracted with phenol according to Gierer and Schramm.[15] The aqueous solution was dialyzed against distilled water for 5 hr at 3° and the polymer recovered by lyophilization. Poly UC was prepared from a mixture of UDP and CDP in molar ratio 5:1, poly UA from a mixture of UDP and ADP in a molar ratio 5:1, and poly CU from UDP and CDP in molar ratio 1:5. The preparation of polythio U has been previously described.[16] 5-Fluorouridine 5′-diphosphate, used in the preparation of polyfluoro U,[7] was a synthetic product kindly supplied by Dr. Charles Heidelberger, University of Wisconsin, Madison, Wisconsin. The sedimentation coefficients of the polynucleotides used in this work are listed in Table 1. Nucleoside diphosphates were obtained from the Schwarz Laboratories, Mount Vernon, N. Y., and the Pabst Laboratories, Milwaukee, Wisconsin. C^{14}-

labeled amino acids were obtained from the California Biochemical Corporation under allocation from the Atomic Energy Commission.

Methods.—In order to decrease the "blank" incorporation of labeled amino acids in the absence of added polynucleotides, *E. coli* ribosomes and supernatant were preincubated for 15 min at 37° in a mixture containing in 1.0 ml the following components (in μmoles unless otherwise specified): Tris-HCl buffer, pH 7.9, 37; Lubrol, 1.8 mg; KCl, 55; $MgCl_2$, 13; mercaptoethylamine, 11; ATP, 0.9; GTP, 0.22; creatine phosphate, 12; creatine kinase, 44 μg; each of 20 nonlabeled amino acids, 0.09; ribosomes with 4 mg of protein; and supernatant with 4 mg of protein. For amino acid incorporation each sample contained 0.1 ml of the above preincubated mixture and 0.15 ml of a basic reaction mixture of the following composition (in μmoles unless otherwise specified): Tris-HCl buffer, pH 7.9, 10.5; Lubrol, 0.5 mg; KCl, 15.3; $MgCl_2$, 2.5; mercaptoethylamine, 3.1; ATP, 0.25; GTP, 0.06; creatine phosphate, 3.3; and creatine kinase, 12.5 μg. Other additions, as noted in the tables, included C^{14}-labeled amino acid (specific radioactivity 1 to 10 μc/μmole), 0.025 μc; transfer RNA, 0.5–1.0 mg; and polynucleotide, 0.04 μmole as mononucleotide. The latter amount was found to be optimal for poly U; higher concentrations were inhibitory. After incubation for 1 hr at 37° the reaction was stopped by the addition of 5 ml of 10 per cent trichloroacetic acid and the samples were worked up according to the procedure of Zamecnik *et al.*[18] The radioactivity of the acid-insoluble residue was measured with a windowless gas-flow counter. Amino acid incorporation is expressed throughout in mμmoles/mg ribosomal protein.

TABLE 1

SEDIMENTATION COEFFICIENTS OF SYNTHETIC POLYNUCLEOTIDES

Polynucleotide	$S_{20,w}$
Poly A	17.7
Poly U (sample 1)	10.3
Poly U (sample 2)	4.5
Poly C	8.6
Poly UC	9.5
Poly UA	9.1
Poly thio U	19.0
Poly fluoro U	9.4

In the experiment in which rat liver and *E. coli* ribosomes were compared, the preincubation described above was omitted and no transfer RNA was added. Protein was determined by the method of Lowry *et al.*[19]

Results.—*Poly U-dependent incorporation of phenylalanine in E. coli system:* Investigation of the incorporation of different C^{14}-labeled amino acids into an acid-insoluble product by an *E. coli* supernatant plus ribosomes system, with and without the addition of poly U, showed that out of 19 amino acids (C^{14}-labeled asparagine was not available to us) tested individually only the incorporation of phenylalanine was markedly stimulated by poly U. A small stimulation of the incorporation of leucine and isoleucine was also observed. The effect of poly U on the incorporation of phenylalanine is shown in Table 2. It may also be seen that addition of *E. coli* transfer RNA brought about a further pronounced increase of phenylalanine incorporation. Since poly U did not affect the loading of transfer RNA with phenylalanine, catalyzed by the amino acid activating enzymes in the supernatant, it appears that poly U affects the transfer of activated phenylalanine residues from the specific phenylalanine-transfer RNA to the ribosomes; this indicates that poly U acts as messenger RNA in this system. As shown by Nirenberg and Matthaei,[11] polyphenylalanine, which is exceedingly insoluble, is formed under these conditions. With the highest concentration of transfer RNA used (Table 2) one mole of phenylalanine was incorporated for every 3.25 uridylic acid

TABLE 2

EFFECT OF *E. coli* TRANSFER RNA ON THE POLY U-DEPENDENT INCORPORATION OF PHENYLALANINE*

Additions to basal system		
Poly U	Transfer RNA (mg/ml)	Phenylalanine incorporation†
−	0	0.1
−	9	0.3
+	0	3.6
+	2	15.2
+	6	23.5
+	9	24.6

* Incubation 30 min at 37°.
† mμmoles/mg of ribosomal protein.

residues in the poly U added. The incorporation of phenylalanine was inhibited by puromycin and, to a lesser extent, by chloramphenicol.

Experiments with other homopolymers: The effect of poly A, poly thio U, and poly fluoro U was investigated in experiments similar to those of Table 2 but *without the addition of transfer RNA*. Poly A did not stimulate the incorporation of any of 19 amino acids tested. Addition of poly A to a system containing poly U completely inhibited the effect of the latter. This is undoubtedly due to formation of the double-stranded, helical poly A + U complex. Poly thio U (which appears to be multi-stranded[20, 21]) had no effect,[22] and poly fluoro U (which like poly U is single-stranded[20]) had but a small effect on phenylalanine incorporation. Poly C had a small but consistent effect on the incorporation of proline (cf. footnote 11) but had no influence on that of any other amino acid. It may be of interest that a short-chain poly U (poly U sample 2, Table 1) was about 40 per cent as effective as an equimolar amount of the longer-chain poly U used throughout this work (poly U sample 1, Table 1) in promoting phenylalanine incorporation.

Experiments with rat liver ribosomes: Poly U had no effect on the incorporation of phenylalanine by a system of rat liver supernatant plus ribosomes (Table 3).

TABLE 3

POLY U-DEPENDENT INCORPORATION OF PHENYLALANINE BY *E. coli* AND RAT LIVER RIBOSOMES

		Phenylalanine incorporation*		
Supernatant	Ribosomes	(a) No poly U	(b) With poly U	Ratio b/a
Liver	Liver	0.40	0.46	1.2
E. coli	*E. coli*	0.07	3.29	47.0
Liver	*E. coli*	0.06	0.17	2.8
E. coli	Liver	0.16	1.81	11.3

* mμmoles/mg of ribosomal protein.

However, no transfer RNA was added in these experiments and, considering the pronounced effect of transfer RNA on the *E. coli* system, this negative result cannot be considered conclusive. With the unsupplemented rat liver system there was no effect on either poly A, poly U, or poly C on the incorporation of any of 19 amino acids tried singly. However, as seen in Table 3, poly U had a fairly marked effect on the incorporation of phenylalanine by a system of *E. coli* supernatant plus rat liver ribosomes. This was not true for the opposite combination, i.e., liver supernatant plus *E. coli* ribosomes. Thus, the ineffectiveness of poly U with the rat liver system is related to the soluble components of this system. It is too early to speculate on the significance of this observation, but the finding that

rat liver ribosomes can be substituted for *E. coli* ribosomes for poly U-dependent incorporation of phenylalanine is not without interest.

Effect of synthetic homo- and copolymers on amino acid incorporation in E. coli system: The results of these experiments are summarized in Table 4. The in-

TABLE 4

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

	Polynucleotide					
Amino acid	None	Poly U	Poly C	Poly UC	Poly UA	Poly CU
Phenylalanine	0.03	13		7	3	0.02
Serine	0.02	0.02	0.01	1.6	0.01	
Tyrosine	0.02			0.02	0.75	
Leucine	0.02	0.3		1.5	0.46	0.03
Isoleucine	0.01	0.09		0.32	0.62	0.007
Proline	0.02	0.02	0.06	0.6	0.03	0.14

* mμmoles/mg of ribosomal protein. 19 amino acids were tested individually in all cases, but the ones giving negative results have been omitted from the table. All values (except those for poly CU) are averages of at least two separate experiments.

corporation of phenylalanine was stimulated to a decreasing extent by poly U, poly UC, and poly UA. Most significant is the fact that some of the copolymers stimulated the incorporation of amino acids other than phenylalanine. Thus, whereas poly UC, and only this polymer, stimulated the incorporation of serine, only poly UA stimulated the incorporation of tyrosine. Less clearcut results were obtained with leucine and isoleucine, the incorporation of both of which was significantly stimulated by poly U, poly UC, and poly UA. Highest incorporation of leucine was promoted by poly UC whereas poly UA brought about the highest incorporation of isoleucine. Ambiguous results were also obtained with proline, the incorporation of which was stimulated by poly C, poly CU, and poly UC in order of increasing effectiveness.

Discussion.—In contrast to poly UC (U:C = 5:1) poly CU (U:C = 1:5) was almost ineffective in promoting amino acid incorporation (Table 4). The reason for this pronounced discrepancy may be sought in the fact that, contrary to natural messenger RNA (or to poly U), the synthetic copolymers do not provide an unbroken sequence of code units. Hence, as far as amino acid incorporation is concerned, there must be many gaps along their chains. This situation might result in the formation of short-chain, acid-soluble polypeptides which would escape detection by the procedure used here. Copolymers such as poly UC and poly UA containing relatively long unbroken U sequences together with some other code units might give rise to short, but acid-insoluble, polyphenylalanine chains with occasional serine, leucine, tyrosine, or isoleucine residues. Viewed in this way, a high proportion of U in a synthetic copolymer would provide a convenient "handle" for incorporation of amino acids other than phenylalanine into short, acid-insoluble polyphenylalanine chains.

If poly U codes for polyphenylalanine, a short sequence of three or more U residues would be the code letter for phenylalanine. If for the sake of simplicity we assume a triplet code, the ratio of UUU to UUC (or UCU, or CUU) triplets in a random copolymer of the composition of poly UC would be the same as the U:C ratio of the polymer, i.e., 5:1. The ratio of UUU to UCC (or CUC, or CCU) triplets would be 25:1. From Table 4 the ratio of phenylalanine to serine incorporation with poly UC was 4.4:1. On the basis of this result the triplet code letter

for serine would be either UUC, UCU, or CUU. The ratio of phenylalanine to tyrosine incorporated with poly UA (U:A = 5:1) was 4.0. This would make either UUA, UAU, or AUU as the likely triplet code letter for tyrosine.

The overlapping results obtained with leucine and isoleucine are difficult to explain, but, if maximal stimulation is taken as the meaningful result, the triplet code letters for these two amino acids would contain 2U and 1C and 2U and 1A, respectively. Difficulties are also encountered with proline (Table 4). Using the same criterion as for leucine and isoleucine, one would tentatively assign a triplet code letter containing 1U and 2C to this amino acid. On the other hand, the greater effectiveness of poly UC as compared to poly C might be due to the provision of an insoluble polyphenylalanine "handle" for the proline residues. Thus it is possible that CCC is really the triplet code letter for proline. It is of interest in this connection that, in a nitrous acid mutant of tobacco mosaic virus described by Tsugita and Fraenkel-Conrat,[23] a proline residue was replaced by leucine. As the nitrous acid effect is due to deamination of C to U, replacement of C by U would be in line with the code letters suggested for these amino acids. Our results would also be compatible with the fact that proline:phenylalanine ratio of wild cucumber and tobacco mosaic virus protein varies in the same direction as the C:U ratio of their respective nucleic acids.[24]

Experiments now in progress with other homo- and copolymers prepared with polynucleotide phosphorylase and deductions from amino acid replacement data as used by Woese[5] may be expected to be of further help in deciphering the nucleotide code.

Summary. Polyribonucleotides synthesized with polynucleotide phosphorylase stimulated the incorporation of certain amino acids into an acid-insoluble product by a system of consisting of *E. coli* high-speed supernatant and ribosomes. Addition of *E. coli* transfer RNA brought about a further pronounced increase of the incorporation of phenylalanine in the presence of poly U. This indicates that the polymers affect the transfer of activated amino acid residues from transfer RNA to ribosomes and act as messenger or template RNA in this system. In one experiment with poly U and phenylalanine, without addition of transfer RNA, rat liver ribosomes could be substituted for their *E. coli* counterparts.

While poly U promoted phenylalanine incorporation, poly UC promoted the incorporation of phenylalanine and serine, and poly UA stimulated the incorporation of phenylalanine and tyrosine. These and other results reported in this paper would appear to open up an experimental approach to the study of the coding problem in protein biosynthesis.

We are indebted to Carlos Basilio for the preparation of poly CU, to R. C. Warner and Yo-Yun Chen for the sedimentation coefficients of the synthetic polynucleotides, and Albert Lenny and Horace Lozina for skillful technical assistance.

Note added in proof.—The results of further experiments to be reported in these PROCEEDINGS, with use of other copolymers (UG, UAC, where G stands for guanylic acid residues), extend to eleven the list of amino acids incorporated into an acid-insoluble product by the *E. coli* system with different polymers. It now includes cysteine, histidine, isoleucine, leucine, lysine, phenylalanine, proline, serine, threonine, tyrosine, and valine.

* 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.

[1] Abbreviations: RNA, ribonucleic acid; the capital letters A, U, and C are used for the nucleotides adenylic, uridylic, and cytidylic acid, respectively, or their corresponding residues in polynucleotide chains; ADP, UDP, and CDP, the 5′-diphosphates of adenosine, uridine, and cytidine; ATP and GTP, the 5′-triphosphates of adenosine and guanosine; Tris, tris(hydroxymethyl)aminomethane.

[2] Gamow, G., *Nature*, **173**, 318 (1954).

[3] Crick, F. H. C., J. S. Griffith, and L. E. Orgel, these PROCEEDINGS, **43**, 416 (1957).

[4] Yčas, M., *Nature*, **188**, 209 (1960).

[5] Woese, C. R., *Biochem. Biophys. Res. Comm.*, **5**, 88 (1961)

[6] Sueoka, N., these PROCEEDINGS, **47**, 1141 (1961).

[7] Jacob, F., and J. Monod, *J. Mol. Biol.*, **3**, 318 (1961).

[8] Brenner, S., F. Jacob, and M. Meselson, *Nature*, **190**, 576 (1961).

[9] Gros, F., H. Hiatt, W. Gilbert, C. G. Kurland, R. W. Risebrough, and J. D. Watson, *Nature*, **190**, 581 (1961).

[10] Grunberg-Manago, M., P. J. Ortiz, and S. Ochoa, *Biochim. Biophys. Acta*, **20**, 269 (1956).

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

[12] Rendi, R., and T. Hultin, *Expl. Cell Research*, **19**, 253 (1960).

[13] Zamecnik, P. C., and E. B. Keller, *J. Biol. Chem.*, **209**, 337 (1954).

[14] Hoagland, M. B., M. L. Stephenson, J. F. Scott, L. I. Hecht, and P. C. Zamecnik, *J. Biol. Chem.*, **231**, 241 (1958).

[15] Gierer, A., and G. Schramm, *Nature*, **177**, 702 (1956).

[16] Lengyel, P., and R. W. Chambers, *J. Am. Chem. Soc.*, **82**, 752 (1960).

[17] Lengyel, P., and C. Basilio, to be published.

[18] Zamecnik, P. C., R. B. Lotfield, M. L. Stephenson, and J. M. Steele, *Cancer Research*, **11**, 592 (1951).

[19] Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. S. Randall, *J. Biol. Chem.*, **193**, 265 (1951).

[20] Lengyel, P., and R. C. Warner, unpublished results.

[21] A. Rich, personal communication.

[22] A. Rich (personal communication) also observed no effect of this polymer.

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

[24] Yamazaki, H., and P. Kaesberg, *Nature*, **191**, 97 (1961).

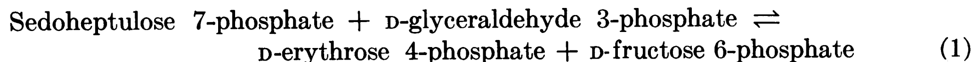
THE PREPARATION OF CRYSTALLINE TRANSALDOLASE FROM CANDIDA UTILIS

BY S. PONTREMOLI, B. D. PRANDINI,* A. BONSIGNORE, AND B. L. HORECKER

THE INSTITUTE OF BIOLOGICAL CHEMISTRY, UNIVERSITY OF GENOA, GENOA, ITALY AND THE DEPARTMENT OF MICROBIOLOGY, NEW YORK UNIVERSITY SCHOOL OF MEDICINE

Communicated October 23, 1961

Transaldolase was first isolated from brewer's yeast[1] and shown to catalyze the reversible reaction:



$$\text{Sedoheptulose 7-phosphate} + \text{D-glyceraldehyde 3-phosphate} \rightleftharpoons \text{D-erythrose 4-phosphate} + \text{D-fructose 6-phosphate} \quad (1)$$

In the course of the reaction a three-carbon unit equivalant to dihydroxyacetone is transferred from the donor, which may be either sedoheptulose 7-phosphate or D-fructose 6-phosphate, to the acceptor, either D-glyceraldehyde 3-phosphate or