RNA Codewords and Protein Synthesis

The Nucleotide Sequences of Multiple Codewords for Phenylalanine, Serine, Leucine, and Proline

Merton R. Bernfield and Marshall W. Nirenberg

Previous studies have shown that randomly ordered polyUC (*1*) templates direct phenylalanine, serine, leucine, and proline into protein in extracts of *Escherichia coli* (*2*). The base compositions of RNA codons corresponding to each amino acid have been estimated by correlating the relative proportion of each amino acid directed into protein with the relative frequency of each codon in polyUC. Triplet codons of unknown base sequence were suggested as follows: phenylalanine, UUU, (UUC); leucine (UUC), (UCC); serine (UCC), (UUC); and proline (UCC), CCC (*2*).

The sequence of bases within codons has been determined by directing the attachment of C^{14}-aminoacyl-sRNA to ribosomes with chemically defined trinucleotides (*3*). This and other studies indicated that the trinucleotides GpUpU, UpGpU, and UpUpG serve as terminal RNA codons for valine (*4*), cysteine, and leucine (*5*), respectively.

We now report the effect of eight trinucleotides, each corresponding to a polyUC triplet, upon the binding of C^{14}-aminoacyl-sRNA to ribosomes. The results indicate that the nucleotide sequences of terminal RNA codons for phenylalanine are UpUpC and UpUpU, and for serine, UpCpC and UpCpU. In addition, the data tentatively suggest that terminal codons corresponding to proline are CpCpC and CpCpU. A general pattern of degeneracy is proposed in which trinucleotide pairs containing identical bases in the first and second positions and U or C 3′-terminal nucleosides, such as XpYpU and XpYpC, serve as codons for the same amino acid.

The authors are affiliated with the Section of Biochemical Genetics of the National Heart Institute, National Institutes of Health, Bethesda, Maryland.

Methods and Materials

Chromatographic and electrophoretic methods. Descending paper chromatography was performed at room temperature with (i) solvent A, *n*-propanol: NH_4OH:H_2O; 55:10:35, by volume; or (ii) solvent B, 0.1*M* ammonium formate, *p*H 6.4. Paper electrophoretic separations were performed either with 0.05*M* ammonium formate, *p*H 2.7, at 80 v/cm for 0.5 hour, or with 0.05*M* *N*-ethyl morpholinium acetate, *p*H 7.55, at 35 v/cm for 2.5 hours. Bands were observed under ultraviolet light and were eluted from paper with H_2O.

Mono-, oligo-, and polynucleotides. Uridine-2′,3′-cyclic phosphate and cytidine-2′,3′-cyclic phosphate were prepared from 2′(3′)-UMP and 2′(3′)-CMP (*6*), or were obtained commercially (*7*), and converted to the ammonium salt by passing a solution through a column of Amberlite IR-120 (NH_4^+). The cyclic nucleotides were further purified by column chromatography on DEAE-cellulose with 0.05*M* or 0.02*M* triethyl ammonium bicarbonate, *p*H 8.0. Nucleosides (*7*) were used without further purification; polyU and polyC were obtained commercially (*8*). The polyUC (designation S-253) was prepared and its base ratio was determined (0.49 U/0.51 C) (*9*).

The dinucleoside monophosphates UpU, UpC, CpU, and CpC and the trinucleoside diphosphates UpUpU, UpUpC, UpCpU, CpUpU, CpCpU, CpUpC, UpCpC, and CpCpC were synthesized by catalyzing, with pancreatic ribonuclease A (or a derivative), the transesterification of pyrimidine-2′,3′-cyclic nucleotides to nucleosides or to dinucleoside phosphates by a modification of the procedure of Heppel, Whitfield, and Markham (*10*). The preparation and characterization of pCpCpC, pUpUpU, and pCpC has been described (*3*).

Purity of oligonucleotides. Each dinucleoside monophosphate (three O.D. units) was chromatographed on both Whatman DE-81 (DEAE) paper with solvent B and on Whatman 40 paper with solvent A. Also, each preparation was subjected to paper electrophoresis (Table 1). Since in each case only one ultraviolet-absorbing spot was found, the purity of each preparation was estimated to be greater than 98 percent.

Two O.D. units of each trinucleoside diphosphate were subjected to chromatography on DEAE paper, as described above, and two O.D. units to paper electrophoresis (Table 1). In addition, 10 to 20 O.D. units of each were separated on Whatman 3 MM paper with solvent A. Since no ultraviolet-absorbing contaminants were observed, the purity of each trinucleotide was estimated to be greater than 98 percent.

The base ratio, base sequence, chain length, and electrophoretic mobility of each oligonucleotide were established (Table 1). Each nucleoside and nucleotide product obtained after incubating an oligonucleotide with pancreatic ribonuclease A or *Crotalus adamantus* venom phosphodiesterase (*11*), free of phosphomonoesterase activity (*12*), was identified either by its chromatographic mobility in solvent A or its electrophoretic mobility on Whatman 40 paper. Each ultraviolet absorbing spot was eluted with 0.01*N* HCl and its ultraviolet spectrum was determined in a Cary recording spectrophotometer. Base ratios of oligonucleotides were calculated from the absorbancy of each separated digestion product at the appropriate maximum wavelength. The absorbancy of each eluate from paper was read against an appropriate blank eluate from paper adjacent to the ultraviolet-absorbing spot.

Components of reactions. Ribosomes and sRNA from *Escherichia coli* W 3100 were prepared by modifications of methods described (*13*). Each C^{14}-aminoacyl-sRNA was prepared in the presence of 19 C^{12}-amino acids. Yeast C^{14}-Phe-sRNA was prepared by W. Groves from baker's yeast (Fleischmann) with an aminoacyl-sRNA synthetase from *Saccharomyces cerevisiae*. C^{14}-Leu-sRNA fractions IA, IB, and II (from *E. coli* B) were separated by countercurrent distribution (*14*). The C^{14}-Leu-sRNA fractions IA and IB correspond to Leu-sRNA fraction I

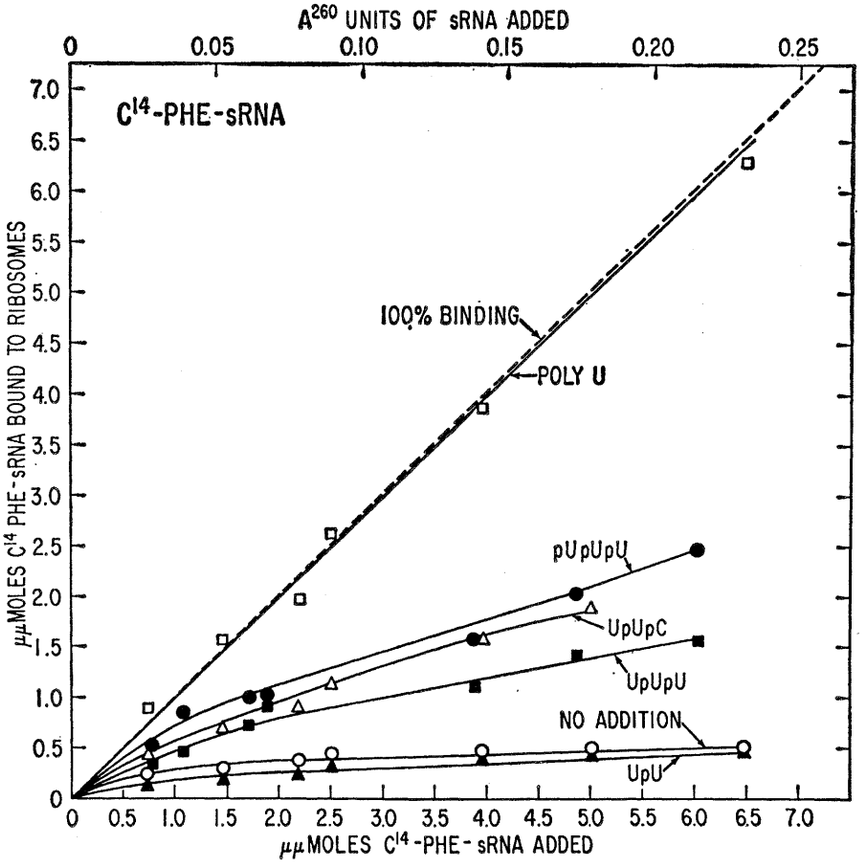

Fig. 1. Relation between C^{14}-Phe-sRNA concentration and the template activities of polyU, pUpUpU, UpUpU, UpUpC, and UpU. The broken line (----) represents the binding to ribosomes of 100 percent of the available C^{14}-Phe-sRNA. In addition to the components already described, reactions contained 0.02*M* magnesium acetate; C^{14}-Phe-sRNA, and 25 mμmole (as base residues) of polyU or oligonucleotide, as indicated. Reactions were incubated at 24°C for 15 minutes.

described by Weisblum, Benzer, and Holley (*15*). These fractions respond to polyUC in *E. coli* extracts that incorporate amino acids into protein. The C^{14}-Leu-sRNA fraction II contains two Leu-sRNA components, IIA and IIB, which recognize polyU and polyUG, respectively, in the amino acid–incorporating system (equivalent to Leu-sRNA fraction II of Weisblum, Benzer, and Holley, *15*). One component of C^{14}-Leu-sRNA fraction II binds to ribosomes in response to the trinucleotide sequence, UpUpG (*5*).

The assay for ribosomal bound C^{14}-aminoacyl-sRNA has been described (*3*). Each reaction mixture (50 μl) contained 0.10*M* tris-acetate, *p*H 7.2; 0.05*M* KCl; 0.02 or 0.03*M* magnesium acetate; and 2.21 A^{260} units of washed ribosomes from *E. coli* W 3100. Conditions of incubation, magnesium acetate concentration, and amounts of oligo- or polynucleotide and C^{14}-aminoacyl-sRNA added are indicated in the legends accompanying each table and figure.

Uniformly labeled C^{14}-L-proline (200 μc/μmole) and C^{14}-L-serine (120 μc/μmole) (*16*), and C^{14}-L-phenylalanine (282 μc/μmole) and C^{14}-L-leucine (160 μc/μmole) (*17*) were obtained commercially. Radioactivity was determined in a liquid scintillation counter with a C^{14}-counting efficiency of 55 to 65 percent (*3*).

Results

Template specificity of poly- and oligonucleotides. The effect of polyU, polyUC, and polyC upon the attachment to ribosomes of C^{14}-Phe-, C^{14}-Ser-, C^{14}-Pro-, and C^{14}-Leu-sRNA is shown in Table 2. PolyU stimulated only the binding of C^{14}-Phe-sRNA to ribosomes; polyUC stimulated C^{14}-Phe-, C^{14}-Ser-, C^{14}-Pro-, and C^{14}-Leu-sRNA binding; and polyC stimulated C^{14}-Pro-sRNA and, to a slight extent, C^{14}-Ser-sRNA binding. The fact that polyUC directs the binding to ribosomes of the four sRNA preparations tested demonstrates that there are appreciable amounts of C^{14}-aminoacyl-sRNA in each preparation capable of responding to codons in polyUC. In addition, each polynucleotide directs, with similar specificity, both the binding of sRNA to ribosomes and the incorporation of amino acid into protein (*2*). It should be noted that a response of serine to polyC has not been reported previously.

In Table 3 is shown the activity of the eight triplets expected in polyUC—UpUpU, UpUpC, UpCpU, CpUpU, UpCpC, CpUpC, CpCpU, and CpCpC—as well as trinucleotides with 5′-terminal phosphate—pUpUpU and pCpCpC—and dinucleotides, upon the binding of C^{14}-aminoacyl-sRNA to ribosomes. To our knowledge the six trinucleotides containing both U and C have not been synthesized previously.

The trinucleotide with 5′-terminal phosphate, pUpUpU, is a more active template for Phe-sRNA than UpUpU (*3*). The binding of C^{14}-Phe-sRNA to ribosomes (Table 3) was also stimulated by UpUpC, whereas other trinucleotides containing U and C, or dinucleotides, were without effect. The template activity of UpUpC for C^{14}-Phe-sRNA was slightly higher than that of UpUpU.

Of the eight triplets, only UpCpU and UpCpC stimulated the binding of C^{14}-Ser-sRNA to ribosomes. The degree of stimulation was dependent upon the percentage of sRNA that was acylated in different C^{14}-Ser-sRNA preparations. In experiments not reported here, C^{14}-Ser-sRNA which had become attached to ribosomes in the presence of UpCpU or UpCpC was not released rapidly from ribosomes upon further incubation with a fivefold excess of deacylated sRNA (compare with *5*).

The binding of C^{14}-Pro-sRNA was stimulated by pCpCpC (*3*) and, slightly, by CpCpC and CpCpU. The higher template activity of pCpCpC as compared to that of the trinucleotide without terminal phosphate, CpCpC, resembles previous results obtained with oligoU and oligoA templates (*3*).

None of the trinucleotides markedly stimulated the binding to ribosomes of unfractionated C^{14}-Leu-sRNA (Table 4). There was slightly more binding in the presence of the remaining two trinucleotides, CpUpU and CpUpC, than in the presence of the other oligonucleotides. As already reported, the binding to ribosomes of the separated C^{14}-Leu-sRNA fractions IA and IB was stimulated by polyUC, but not by polyU or polyUG; whereas the binding of C^{14}-Leu-sRNA fraction II was stimulated by polyUG, UpUpG, and polyU,

Table 1. Oligonucleotide preparation and characterization.

Compound	Synthesized from	Relative electrophoretic mobility*			Digestion		
		pH 2.7 †		pH 7.55 ‡			
		To anode	To cathode	To anode	Enzyme	Products	Base ratio
UpU	U>p + U	0.82			RNase §	Up/U	1.0/1.0
CpC	C>p + C		0.64		RNase	Cp/C	1.0/1.0
UpC	U>p + C	0.05			RNase	Up/C	0.9/1.0
CpU	C>p + U	0.05		0.34	RNase	Cp/U	1.2/1.0
UpUpC	U>p + C	0.60			RNase	Up/C	1.9/1.0
CpCpU	C>p + U	0.04		0.56	RNase	Cp/U	2.2/1.0
UpCpC	U>p + CpC	0.05			SVD ||	U/pC	1.0/2.1
CpUpU	C>p + UpU	0.61			SVD	C/pU	1.0/2.2
UpCpU	U>p + CpU	0.61			RNase	Up/Cp/U	1.0/0.9/1.0
UpCpU	U>p + CpU	0.61			SVD	U/pC/pU	1.1/0.8/1.0
CpUpC	C>p + UpC	0.04		0.57	RNase	Cp/Up/C	1.2/1.0/0.9
CpUpC	C>p + UpC	0.04		0.57	SVD	C/pU/pC	1.0/0.9/1.1
UpUpU	U>p + U	1.16			RNase	Up/U	2.0/1.0
CpCpC	C>p + C		0.48		RNase	Cp/C	1.9/1.0
2′(3′)-UMP*		1.00					
Cytidine*			1.00				
2′(3′)-CMP*				1.00			

* Mobility of reference markers defined as 1.00. † 0.05*M* ammonium formate, *p*H 2.7; 80 v/cm for 0.5 hr. ‡ 0.05*M* *N*-ethyl morpholinium acetate, *p*H 7.55; 35 v/cm for 2.5 hr. Oligonucleotides purified with this buffer were converted to the ammonium salt prior to assaying their template activity with ribosomes. § Incubated with 5 μg of pancreatic ribonuclease A (RNase) in 0.05*M* NH_4HCO_3, *p*H 7.5 for 4 hr at 37°C. || Incubated with 10 μg of *Crotalus adamantus* venom phosphodiesterase (SVD) in 0.05*M* NH_4CO_3, *p*H 9.0 and 0.01*M* magnesium acetate for 4 hr at 37°C.

but not by polyUC (*5*). None of the oligonucleotides had a discernible effect (Table 4) upon the binding to ribosomes of C^{14}-Leu-sRNA fractions IA, IB, or II.

In other experiments, each trinucleotide was without marked effect upon the binding to ribosomes of 14 other C^{14}-aminoacyl-sRNA preparations, each acylated with a different C^{14}-amino acid; (the trinucleotides were not tested with labeled Try- and Cys-sRNA).

Percentage of Phe-, Ser-, and Pro-sRNA binding to ribosomes in response to RNA codons. PolyUC directs the stable binding to ribosomes of approximately 6 percent of the available unfractionated C^{14}-Leu-sRNA (*5*). In order to estimate the minimum proportion of C^{14}-Phe-sRNA which can respond either to UpUpU, UpUpC, pUpUpU, or polyU, the effect of each template upon the binding of C^{14}-Phe-sRNA to ribosomes was determined in reactions containing different concentrations of C^{14}-Phe-sRNA (Fig. 1). PolyU directed the binding to ribosomes of essentially 100 percent of the available C^{14}-Phe-sRNA within the range 0 to 6.7 μμmole of C^{14}-Phe-sRNA; whereas, UpUpU, UpUpC, and pUpUpU stimulated the binding of approximately 50, 60, and 70 percent, respectively, of the available C^{14}-Phe-sRNA. At all C^{14}-Phe-sRNA concentrations studied, UpU was inactive. These data demonstrate the template activity of UpUpC for C^{14}-Phe-sRNA and indicate that at least 60 percent of the available C^{14}-Phe-sRNA can respond to UpUpC.

Similarly with C^{14}-Ser- and C^{14}-Pro-sRNA and their appropriate polynucleotide templates (Fig. 2, *A* and *B*) polyUC directed the stable binding to ribosomes of approximately 10 percent of the available C^{14}-Ser-sRNA. It is not known whether C^{14}-Ser-sRNA binds to ribosomes with less stability than Phe-sRNA or whether most of the available C^{14}-Ser-sRNA responds to codons which are not present in polyUC.

Both polyUC and polyC directed the binding of C^{14}-Pro-sRNA to ribosomes. The percentage of binding of C^{14}-Pro-sRNA is difficult to estimate because, in contrast to the stability of other aminoacyl-sRNA preparations in this system, there was a pronounced deacylation of C^{14}-Pro-sRNA during the incubation. For this reason the amount of C^{14}-Pro-sRNA (count/min in the fraction precipitated by TCA at 0°C) before incubation, as well as the amount found after incubation, are shown.

Template activity of UpUpU and UpUpC. The relation between the concentration of UpUpU or UpUpC in the reaction mixture and the amount of C^{14}-Phe-sRNA bound to ribosomes is shown in Fig. 3. The reactions were incubated for 15 minutes at 24°C in order to determine the maximum amount of C^{14}-Phe-sRNA bound to ribosomes rather than the rate of binding. Under these conditions the C^{14}-Phe-sRNA binding was maximum when the concentration of UpUpC was $10^{-4}M$. Even at limiting concentrations, the template activity of UpUpC for C^{14}-Phe-sRNA was slightly higher than that of UpUpU.

Both UpUpC and UpUpU stimulated the binding of yeast C^{14}-Phe-sRNA to *E. coli* ribosomes (Table 5). At equimolar concentrations, UpUpC was a more active template for yeast Phe-sRNA than pUpUpU or UpUpU. Most of the available yeast C^{14}-Phe-sRNA formed complexes with ribosomes in response to polyU.

Table 2. Template specificity of polynucleotides in stimulating C^{14}-aminoacyl-sRNA binding to ribosomes. Where indicated, 20 mμmole (as base residues) of the appropriate polynucleotide, 13.7 μμmole C^{14}-Phe-sRNA (0.54 A^{260} unit), 12.7 μμmole C^{14}-Ser-sRNA (1.23 A^{260} unit), 9.38 μμmole C^{14}-Pro-sRNA (0.47 A^{260} unit) (2.77 μμmole after incubation), or 31.5 μμmole C^{14}-Leu-sRNA (0.26 A^{260} unit) was added to reaction mixtures containing 0.03*M* magnesium acetate and other components already described. The reaction mixtures were incubated for 15 minutes at 24°C.

Polynucleotide	C^{14}-Aminoacyl-sRNA bound to ribosomes (μμmole)			
	C^{14}-Phe-sRNA	C^{14}-Ser-sRNA	C^{14}-Pro-sRNA	C^{14}-Leu-sRNA
None	1.11	0.40	0.50	1.24
PolyU	12.30	0.40	0.30	1.22
PolyUC	4.68	3.00	1.87	3.64
PolyC	0.81	0.87	1.37	1.18

Table 3. Template specificity of oligonucleotides in stimulating C^{14}-aminoacyl-sRNA binding to ribosomes. The conditions of incubation and the amounts of C^{14}-Phe-sRNA, C^{14}-Ser-sRNA, and C^{14}-Pro-sRNA added to reaction mixtures are described in the legend to Table 2. In experiment 2, reactions contained 6.7 μμmole of C^{14}-Pro-sRNA (0.72 A^{260} unit), or 19.1 μμmole of C^{14}-Ser-sRNA (0.39 A^{260} unit). Each reaction mixture contained 6.7 mμmole of the oligonucleotide specified (as molecules, not as base residues) and the components as already described (Table 2 and Methods).

Oligonucleotide	C^{14}-Aminoacyl-sRNA bound to ribosomes (μμmole)		
	C^{14}-Phe-sRNA	C^{14}-Ser-sRNA	C^{14}-Pro-sRNA
	Experiment 1		
None	1.11	0.40	0.50
pUpUpU	2.86	0.42	0.37
UpUpU	2.16	0.39	0.41
UpUpC	2.82	0.38	0.42
UpCpU	1.09	0.76	0.43
CpUpU	1.11	0.29	0.42
UpCpC	0.92	0.72	0.41
CpUpC	1.03	0.42	0.34
CpCpU	1.08	0.46	
CpCpC	0.78	0.40	
pCpCpC	0.91	0.32	
UpC	0.81	0.39	0.44
CpU	0.97	0.40	0.41
CpC	0.96	0.45	0.39
pCpC	0.99	0.39	0.36
	Experiment 2		
None			0.14
CpCpU			0.29
CpCpC			0.22
pCpCpC			0.42
UpCpU		2.49	
UpCpC		1.60	

Discussion

The use of poly- or trinucleotides to direct the binding of C^{14}-aminoacyl-sRNA to ribosomes provides a technique for estimating the minimum proportion of an aminoacyl-sRNA species which can respond to an RNA codon, whereas, the use of polynucleotides to direct cell-free protein synthesis may reflect the catalytic recognition of codons by sRNA (*3, 5*). In accord with the template activity of polyUC during protein synthesis in cell-free *E. coli* extracts (*2*), polyUC directed the binding to ribosomes of appreciable amounts of C^{14}-Pro-sRNA, of at least 10 percent of the available C^{14}-Ser-sRNA and, as reported (*5*), 6 percent of the available C^{14}-Leu-sRNA. Essentially all of the available C^{14}-Phe-sRNA binds to ribosomes in response to polyU (*5*).

Since an appreciable amount of C^{14}-Phe-, C^{14}-Ser-, C^{14}-Pro-, and C^{14}-Leu-sRNA was bound to ribosomes in response to codons present in polyUC, the effect of each trinucleotide containing U and C upon the formation of sRNA-ribosome complexes was determined. Each of 18 sRNA preparations was acylated with a different C^{14}-amino acid. UpUpU and pCpCpC specifically direct C^{14}-Phe- and C^{14}-Pro-sRNA binding to ribosomes, respectively (*3*). In this study UpUpC, in addition to UpUpU, stimulated the binding of C^{14}-Phe-sRNA; UpCpU and UpCpC stimulated the binding of C^{14}-Ser-sRNA; whereas pCpCpC and, to a much lesser extent, CpCpC and CpCpU stimulated the binding of C^{14}-Pro-sRNA. The dinucleotides, pUpU, UpU, UpC, CpU, and pCpC were inactive. If one assumes that trinucleotides are recognized in phase, these data suggest that the nucleotide sequences of 3′-terminal RNA codons for phenylalanine are UpUpC and UpUpU, and for serine, UpCpU and UpCpC. Also, the data tentatively suggest that the sequences CpCpU and CpCpC correspond to proline.

The possibility must be considered that correct recognition of two out of three bases in a trinucleotide, in or out of phase, or two bases in one trinucleotide and one base in an adjacent trinucleotide occasionally may permit the attachment of C^{14}-aminoacyl-sRNA to ribosomes. This may account for relatively small stimulations of aminoacyl-sRNA binding by some trinucleotides. However, since UpUpC is a very active template for Phe-sRNA, and UpCpU, CpUpU, GpUpU, UpUpG and the dinucleotides, UpU and UpC are not, each base in UpUpC is probably recognized in correct phase (*4, 5*). The rate of binding, and also the affinity of an sRNA for a ribosome, may be greater when a codon is recognized correctly and in proper phase than when codon recognition is only partially correct. Ultimately it will be necessary to compare the relative template activities of the 64 trinucleotides with each C^{14}-aminoacyl-sRNA.

Although polyUC was an active template for C^{14}-Leu-sRNA binding, trinucleotides containing U and C were almost inactive. Binding of C^{14}-Leu-sRNA to ribosomes was slightly higher in the presence of the two unassigned

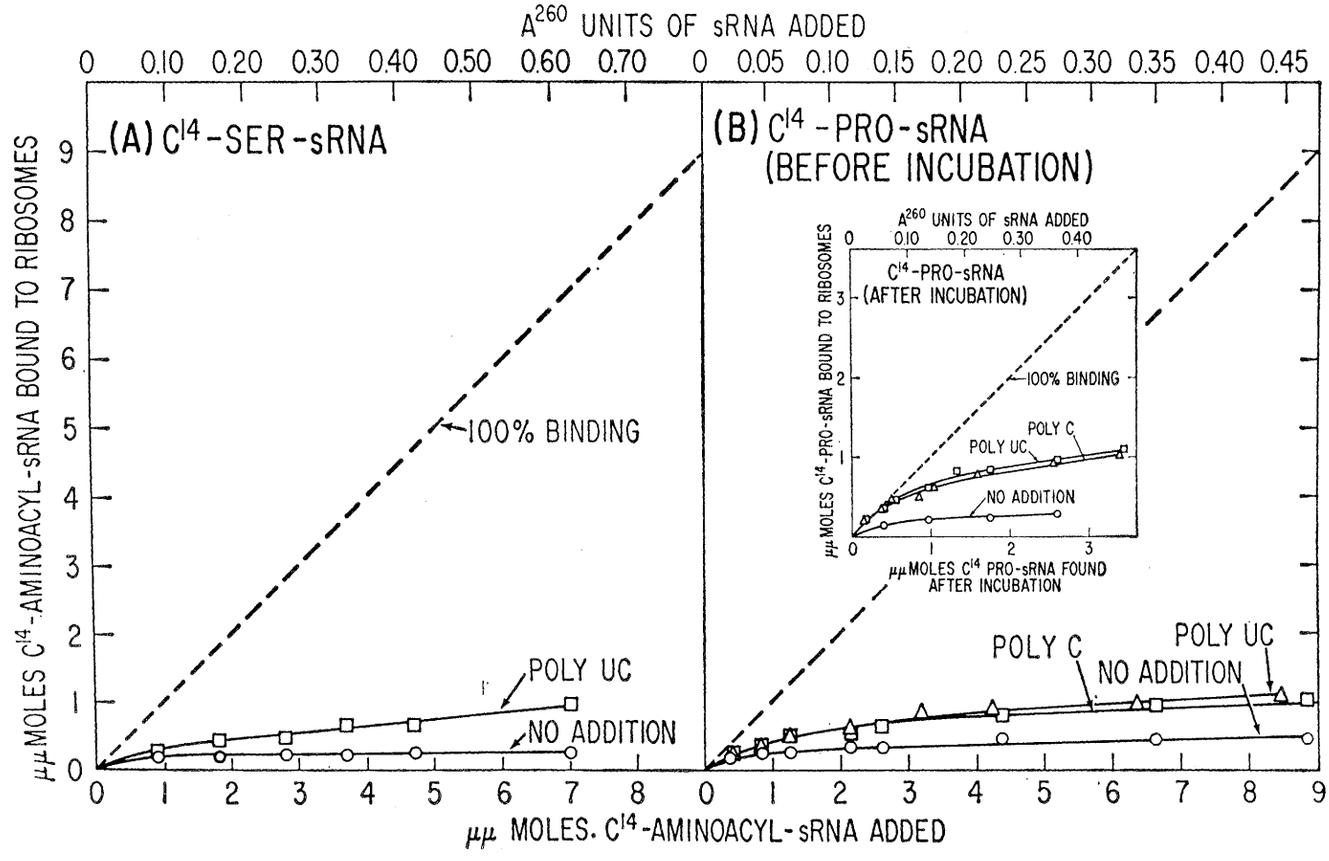

Fig. 2. Relation between the template activities of polynucleotides and C^{14}-aminoacyl-sRNA concentration. In each figure, the broken line (— — — —) represents the binding of 100 percent of the available C^{14}-aminoacyl-sRNA. In Fig. 2*A*, reactions contained, in addition to the components described in the text, C^{14}-Ser-sRNA and 20 mμmole (as base residues) of polyUC as indicated, and 0.02*M* magnesium acetate. Reactions were incubated for 20 minutes at 24°C. In Fig. 2*B*, reactions contained C^{14}-Pro-sRNA and 20 mμmole (as base residues) of polyUC or polyC, as indicated; 0.03*M* magnesium acetate, and other components described in the text. Reactions were incubated for 15 minutes at 24°C. The amount of C^{14}-Pro-sRNA added to each reaction is shown in the large figure; the amount of C^{14}-Pro-sRNA found in each reaction after incubation for 15 minutes at 24°C is shown in the insert.

trinucleotides, CpUpU and CpUpC, than in the presence of the other trinucleotides. This suggests that a base sequence in a polynucleotide may be readable, whereas the expected sequence in a trinucleotide may be read poorly or not at all. However, other possibilities also deserve consideration. It should be noted that CpUpU and CpUpC are the first trinucleotides without terminal phosphate to be found which resemble nonreadable (that is, nonsense) terminal codons.

The template efficiency of oligoU and oligoA is influenced greatly by the chainlength of the oligonucleotide and by terminal phosphate (*3*). Trinucleotides with 5′-terminal phosphate are more active templates than trinucleotides without terminal phosphate. The addition of phosphate to a 3′(2′)-terminal hydroxyl group of a trinucleotide greatly reduces its template activity and resembles a sense-nonsense interchange (*3*). Since a triplet may occur in three different chemical forms—as a 3′-terminal, 5′-terminal, or internal codons (*3*) —its template efficiency, and perhaps also its specificity for aminoacyl-sRNA, may be influenced by other codons in the same chain. It should be emphasized that the trinucleotides used in this and previous studies (*4–6*) contain both 5′- and 3′-terminal nucleosides, and their template activities in this system may reflect recognitions of both terminal nucleosides. Such considerations suggest that the sequences CpUpU and CpUpC may be readable internal codons for Leu-sRNA, but may be less efficient or nonreadable at a terminal position.

Inasmuch as polyUC directs the binding to ribosomes of only a small proportion of the available C^{14}-Ser-sRNA, additional serine codons, possibly composed of three different bases, are suggested. It should be noted that two other codons of unspecified sequence, (UCG) and (ACG) have been proposed for serine (*2*).

Two dimensional oligonucleotide patterns of ribonuclease digests of yeast Ser-sRNA countercurrent fractions II and III have been shown by Rushizky *et al.* to be almost identical (*18*). Ser-sRNA II contained 1.06 moles of (ApGp)GpUp per mole of sRNA and 1.27 moles of ApUp per mole of sRNA; whereas Ser-sRNA III contained 2.32 moles and 0.52 mole per mole of sRNA, respectively (*18*). Cantoni *et al.* reported that purified, unfractionated yeast Ser-sRNA contains the sequence ApGpApΨp (*19*). Although these data may provide insight into the mechanism of codon recognition, further work is necessary to determine whether these sequences are related to the codon specificity of each sRNA.

A general pattern of degeneracy is suggested by the observations that multiple codons for *E. coli* Phe-, Ser-, and perhaps Pro-sRNA, as well as yeast Phe-sRNA, may contain either a U or a C at the 3′-terminus. Certainly, terminal nucleosides of 3′- or 5′-terminal codons may be recognized with less fidelity than internal nucleotides. A terminal nucleoside may have greater freedom of movement and perhaps its orientation on a ribosome may be different from that of an internal nucleotide. Similar considerations also may apply to internal codons because the interaction between an internal codon and a ribosome may be influenced by neighboring nucleotides which bind to ribosomes at other sites. We suggest that part of the degeneracy pattern observed for each amino acid may be due to the effect of neighboring nucleotides upon the binding of codons to ribosomes.

Studies with randomly-ordered polynucleotides and cell-free protein synthesizing systems have shown that multiple codons corresponding to one amino acid often differ in base composition by only one base (*2, 9, 20*). These data suggested that (i) nucleotides common to multiple words may occupy identical positions within each triplet, and (ii) either recognition of two out of three nucleotides in a triplet may in some cases suffice during protein synthesis, or (iii) a nucleotide at one position in a triplet may be recognized correctly in two or more ways. Eck has extended these concepts by suggesting that U and C are equivalent and A and G are equivalent at one unspecified position in a triplet (*21*). Our results show that U and C are equivalent at the 3′-terminal nucleoside position.

The 32 3′-terminal RNA codons with either U or C 3′-terminal nucleosides may be arranged in 16 pairs of triplets, each pair containing the same nucleotide sequence in the first and second positions and either U or C as the 3′-terminal nucleoside (for example, XpYpU and XpYpC). As Eck suggested (*21*), if one trinucleotide is a template for an amino acid, the remaining trinucleotide in that pair also will serve as a template for the same amino acid. Such predictions are for 3′-terminal codons only. Further work is necessary to determine whether similar predictions can be made for internal and 5′-terminal codons. However, trinucleotides with U or C as the 5′-terminal nucleoside and with identical bases in the second and third positions may specify different amino acids because CpUpU and CpUpC do not serve as templates for C^{14}-Phe-sRNA and are not equivalent to UpUpU or UpUpC.

From the foregoing predictions, the nucleotide sequences found for Phe-, Ser-, Pro-, Val-, Cys-, and Leu- codons (*4, 5*), as well as from amino acid

Table 4. The effect of oligonucleotides upon the binding of C^{14}-Leu-sRNA to ribosomes. In addition to the components described in the text, reaction mixtures contained 0.03*M* magnesium acetate and 6.7 mμmole (molecules, not base residues) of oligonucleotide. Reactions with C^{14}-Leu-sRNA fraction II were incubated for 10 minutes and others for 15 minutes at 24°C. In experiment 1, reaction mixture contained, where indicated, 31.5 μμmole of unfractionated C^{14}-Leu-sRNA (0.26 A^{260} unit); 3.92 μμmole of C^{14}-Leu-sRNA fraction IA (0.035 A^{260} unit); 4.72 μμmole of C^{14}-Leu-sRNA fraction IB (0.036 A^{260} unit); and 3.56 μμmole of C^{14}-Leu-sRNA fraction II (0.017 A^{260} unit). In experiment 2, reaction mixtures, where specified, contained 1.78 μμmole of C^{14}-Leu-sRNA fraction IA (0.020 A^{260} unit); 1.41 μμmole of C^{14}-Leu-sRNA fraction IB (0.011 A^{260} unit); and 3.56 μμmole of C^{14}-Leu-sRNA fraction II (0.017 A^{260} unit). Leu-sRNA was purified by countercurrent fractionation.

Oligonucleotide	C^{14}-Leu-sRNA bound to ribosomes (μμmole)			
	Unfractionated	Fraction IA	Fraction IB	Fraction II
	Experiment 1			
None	1.24	0.31	0.38	0.68
pUpUpU	0.83	0.27	0.38	0.68
UpUpU	1.00			0.56
UpUpC	0.92	0.27	0.40	0.68
UpCpU	1.13	0.30	0.35	0.64
CpUpU	1.28	0.30	0.35	0.62
UpCpC	1.18	0.29	0.34	0.50
CpUpC	1.27			
CpCpU	1.13			
CpCpC	1.12			
	Experiment 2			
None		0.21	0.27	0.40
CpUpC		0.21	0.23	0.40
CpCpU		0.24	0.21	0.43

Table 5. Trinucleotide specificity in directing the binding of yeast C^{14}-Phe-sRNA to *E. coli* ribosomes. Reactions contained, in addition to the components already described, 0.03*M* magnesium acetate; 11.5 μμmole of yeast C^{14}-Phe-sRNA (0.306 A^{260} unit); 15 mμmole (as base residues) of oligonucleotide or 20 mμmole (as base residues) of polyU where specified. Reactions were incubated for 15 minutes at 24°C.

Addition	Yeast C^{14}-Phe-sRNA bound to ribosomes (μμmole)
None	0.28
UpUpU	.84
UpUpC	1.37
pUpUpU	1.07
PolyU	8.32

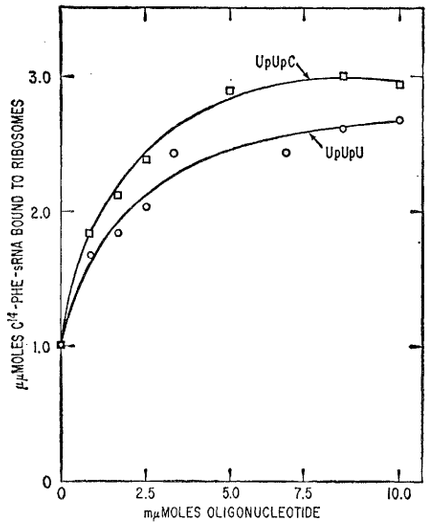

Fig. 3. Relation between UpUpU and UpUpC concentration and C^{14}-Phe-sRNA binding to ribosomes. Each reaction mixture contained 7.68 $\mu\mu$mole of C^{14}-Phe-sRNA (0.43 A^{260} unit); 0.03M magnesium acetate; mμmole (as molecules) of UpUpU or UpUpC as specified; in addition to the components described in the text. Reactions were incubated for 15 minutes at 24°C.

replacement data reported for *E. coli* (*22*) and tobacco mosaic virus mutants (*23, 24*), nucleotide sequences of other RNA codons can be derived. Although many additional sequences can be predicted, only the most probable are presented (Table 6). Trinucleotides with terminal bases were used in these studies, and it is not known whether 5′-terminal, 3′-terminal, and internal codons are recognized in identical ways.

Possibly the trinucleotides, UpUpU and UpUpC, may be recognized by different Phe-sRNA molecules, each with specificity for one trinucleotide sequence. Alternatively, one molecule of Phe-sRNA may respond to both trinucleotides. The data of Fig. 2 show that polyU can direct the binding to ribosomes of 99 to 100 percent of the available C^{14}-Phe-sRNA. Since at least 60 percent of the C^{14}-Phe-sRNA also can be directed to bind to ribosomes by UpUpC, most C^{14}-Phe-sRNA molecules can probably respond both to UpUpU and to UpUpC. The template activity of UpUpC was slightly higher than that of UpUpU with both *E. coli* and yeast C^{14}-Phe-sRNA, and this suggests that C may be recognized with equal or greater affinity than U. It is not possible at this time to distinguish between the many mechanisms which might conceivably permit a dual recognition of U and C.

To investigate nucleotide sequences of RNA codons, the trinucleotides UpUpU, UpUpC, UpCpU, CpUpU, CpCpU, CpUpC, UpCpC, and CpCpC, and the dinucleotides UpU, UpC, CpU, and CpC, were used as templates to direct the binding of C^{14}-aminoacyl-sRNA to *E. coli* ribosomes. The results suggest that both UpUpU and UpUpC serve in *E. coli* as terminal RNA codons for phenylalanine. Yeast C^{14}-phenylalanine sRNA also became attached to *E. coli* ribosomes in response to UpUpU and UpUpC. One molecule of Phe-sRNA may respond to both UpUpU and UpUpC. In addition, UpCpU and UpCpC serve in *E. coli* as terminal RNA codons for serine.

The template efficiencies of the remaining trinucleotides were low, consequently only tentative estimates were made concerning their specificity. The data indicate that CpCpU and CpCpC correspond to proline; and CpUpU and CpUpC to leucine. PolyUC was a more efficient template for serine-, proline-, and leucine-sRNA than any of the aforementioned trinucleotides. Thus trinucleotides vary widely in template efficiency and the properties of a terminal codon may be different from that of an internal codon.

A general pattern of degeneracy was suggested for 3′-terminal codons in which codon pairs, containing identical bases in the first and second positions and U or C 3′-terminal nucleosides (such as XpYpU and XpYpC), correspond to the same amino acid.

Table 6. Predicted nucleotide sequences of terminal RNA codons. The template activities of the trinucleotides written in italics have been studied experimentally in extracts of *E. coli* W 3100. Codon assignments marked with an asterisk (*) are tentative. Nucleotides within parentheses have not been arranged in sequence. Amino acid replacements used for these predictions were found in *E. coli* by Yanofsky and co-workers (*22*) or were induced by HNO_2 in tobacco mosaic virus by Wittman *et al.* (*23*) and by Tsugita (*24*). No attempt has been made to distinguish between codons for aspartic and glutamic acids.

Amino acid	Predicted sequences			
Phenylalanine	*UpUpU*	*UpUpC*		
Serine	*UpCpU*	*UpCpC*	UpCpG	
Leucine	*CpUpU**	*CpUpC**	*UpUpG*	CpUpG
Proline	*CpCpU**	*CpCpC*		
Valine	*GpUpU*	GpUpC		
Cysteine	*UpGpU*	UpGpC		
Isoleucine	ApUpU	ApUpC	ApUpA	
Tyrosine	UpApU	UpApC		
Threonine	ApCpU	ApCpC	ApCpA	
Histidine	CpApU	CpApC		
Arginine	CpGpU	CpGpC	Cp(Gp)A	ApGpA
Alanine	GpCpU	GpCpC		
Glycine	GpGpU	GpGpC	GpGpA	
Aspartic, glutamic acids	ApGpU	ApGpC		
	GpApU	GpApC		
Asparagine	ApApU	ApApC		
Lysine	*ApApA*			

References and Notes

1. The following abbreviations are used: polyU, polyuridylic acid; polyC, polycytidylic acid; polyUC, copolymer of uridylic (UMP) and cytidylic (CMP) acids; polyUG, copolymer of uridylic and guanylic acids; U, uracil; A, adenine; C, cytidine; G, guanine; TCA, trichloroacetic acid; DEAE, diethylaminoethyl; O.D., optical density; A^{260}, absorbancy at 260 mμ; sRNA, transfer RNA; and Leu-, leucine-, Phe-, phenylalanine-, Ser-, serine-, Try-, tryptophane-, Cys-, cystine-, Pro-, proline-sRNA. For mono- and oligonucleotides of specific structure, the p to the left of a terminal nucleoside initial indicates a 5′-terminal phosphate; the p to the right, a 3′-(2′)-terminal phosphate. Internal phosphates of oligonucleotides are (3′, 5′)-phosphodiester linkages. Oligonucleotides whose nucleoside initials are enclosed within parenthesis are of unspecified sequence. The symbol > indicates the cyclic form of the nucleotide.
2. M. W. Nirenberg, O. W. Jones, P. Leder, B. F. C. Clark, W. S. Sly, S. Pestka, *Cold Spring Harbor Symp. Quant. Biol.* **28**, 549 (1963); J. F. Speyer, P. Lengyel, C. Basilio, A. J. Wahba, R. S. Gardner, S. Ochoa, *ibid.*, p. 559.
3. M. W. Nirenberg and P. Leder, *Science* **145**, 1399, (1964).
4. P. Leder and M. W. Nirenberg, *Proc. Natl. Acad. Sci. U.S.* **52**, 420 (1964).
5. ———, *ibid.*, p. 1521.
6. W. Szer and D. Shugar, *Biochemical Preparations* **10**, 139 (1963).
7. Schwarz BioResearch, Inc.
8. Miles Chemical Co.
9. O. W. Jones and M. W. Nirenberg, *Proc. Natl. Acad. Sci. U.S.* **48**, 2115 (1962).
10. L. A. Heppel, P. R. Whitfield, R. Markham, *Biochem. J.* **60**, 8 (1955).
11. Worthington Biochemical Corporation.
12. R. Hilmoe and M. F. Singer, personal communication.
13. M. W. Nirenberg, J. H. Matthaei, O. W. Jones, *Proc. Natl. Acad. Sci. U.S.* **48**, 104 (1962).
14. Gift of Drs. B. Weisblum and G. von Ehrenstein.
15. B. Weisblum, S. Benzer, R. W. Holley, *Proc. Natl. Acad. Sci. U.S.* **48**, 1449 (1962).
16. New England Nuclear Corp.
17. Nuclear-Chicago Corp.
18. G. W. Rushizky, H. A. Sober, C. M. Connelly, B. P. Doctor, *Biochem. Biophys. Res. Commun.*, in press.
19. G. L. Cantoni, H. Ishikura, H. H. Richards, K. Tanaka, *Cold Spring Harbor Symp. Quant. Biol.* **28**, 123 (1963).
20. M. W. Nirenberg and O. W. Jones, in *Symposium on Informational Macromolecules*, H. Vogel, V. Bryson, J. Lampen, Eds. (Academic Press, New York, 1963), p. 451.
21. R. V. Eck, *Science* **140**, 477 (1963).
22. C. Yanofsky, *Cold Spring Harbor Symp. Quant. Biol.* **28**, 581 (1963), and personal communication.
23. H. G. Wittman and B. Wittman-Liebold, *ibid.*, p. 589.
24. A. Tsugita, personal communication.

25. We thank Theresa Caryk and Norma Zabriskie for assistance.