THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 240, No. 5, May 1965
Printed in U.S.A.

Nucleotide Sequences in the Yeast Alanine Transfer Ribonucleic Acid*

ROBERT W. HOLLEY, GEORGE A. EVERETT, JAMES T. MADISON, AND ADA ZAMIR

From the United States Plant, Soil, and Nutrition Laboratory, Soil and Water Conservation Research Division, Agricultural Research Service, United States Department of Agriculture, and the Department of Biochemistry, Cornell University, Ithaca, New York

(Received for publication, November 27, 1964)

The purification of the yeast alanine-, tyrosine-, and valine-transfer ribonucleic acids by countercurrent distribution has been described (1), and preliminary data have been reported on the oligonucleotide compositions of these RNAs (2, 3). The present paper summarizes results of attempts to account quantitatively for all of the fragments obtained by digestion of the alanine-RNA with pancreatic RNase and with Taka-Diastase RNase T1. Results of the analyses of these two digests are consistent, one with the other, and indicate that the alanine-RNA is composed of 77 nucleotides, including nine unusual nucleotides.

EXPERIMENTAL PROCEDURE

Isolation of Alanine-RNA—The alanine-RNA was isolated by countercurrent distribution essentially as described (1, 3). In recent isolations, the first distribution has been run for 700 transfers, with the upper phase feeding in continuously at tube 1 and running out of the apparatus at tube 200. The alanine-RNA was then reisolated from tubes 150 to 190, and the peak of activity was redistributed once for 1000 transfers (1, 3). The alanine acceptor activity of the RNA was approximately that previously described, but has been somewhat variable (±30%). It is possible that there is some subtle structural change that is related to this variation in activity, but if so, it has not been detected in the analyses. The analyses described below suggest a purity of approximately 90%. The alanine-RNA, after isolation as a dry solid (1, 3), has been stored at −20°, although, as far as is known, the dry RNA is completely stable at room temperature.

Digestion with Pancreatic RNase and Chromatography on DEAE-Sephadex—The alanine-RNA (7.7 mg) was dissolved in 3.0 ml of water, and the solution was mixed with 0.37 ml of 0.10 M sodium phosphate buffer (pH 7.0), 0.37 ml of a 1 mg per ml solution of crystalline pancreatic RNase, and 0.15 ml of chloroform. The mixture was incubated for 14 hours at 37°, and then was diluted to 10 ml with water before chromatography.

DEAE-Sephadex A-25, medium (fines removed by decantation in water), was washed with 0.5 N hydrochloric acid, water, 0.5 N sodium hydroxide, water, 1 M ammonium carbonate, and finally with water. A column (0.25 × 225 cm) (two pieces of tubing 4 mm in outside diameter joined together with a small piece of Tygon tubing) was packed by gravity with a suspension of the DEAE-Sephadex in water. The RNase digest was allowed to flow onto the column by gravity and was followed by 10 ml of water. The column was then eluted with a continuous gradient produced from 240 ml of 0.04 M, 238 ml of 0.20 M, and 228 ml of 1.0 M ammonium carbonate in three chambers of a Varigrad (4). Fractions of approximately 3.4 ml were collected. A flow rate of approximately 10 ml per hour was obtained by placing the Varigrad 20 feet (6 meters) above the top of the column.

Chromatography of Pancreatic RNase Digest on DEAE-cellulose in 7 M Urea (5)—DEAE-cellulose (Carl Schleicher and Schuell, No. 70, standard) was washed thoroughly with 7 M urea and with sodium acetate in 7 M urea. A column (0.35 × 30 cm) was packed by gravity from a suspension of the DEAE-cellulose in 7 M urea. A digest of 1 mg of alanine-RNA (digested as above in a volume of 0.5 ml, but for only 4 hours at 37°) was mixed with 300 mg of urea and 2 ml of 7 M urea, and the solution was allowed to flow onto the DEAE-cellulose column. The digest was washed onto the column with 1 ml of 7 M urea, and the column was eluted with a linear gradient formed from 60 ml of 7 M urea and 60 ml of 0.6 M sodium acetate (pH 7.5) in 7 M urea in two chambers of a Varigrad. Fraction volumes were approximately 1.2 ml.

Digestion with Taka-Diastase RNase T1 and chromatography of the digest on DEAE-Sephadex and on DEAE-cellulose were performed as previously described (6). *Separation of oligonucleotides by paper electrophoresis and determination of the nucleotide compositions of oligonucleotides* were done by procedures previously described (6).

Determination of 5′-Terminal Nucleotides by Complete Degradation with Snake Venom Phosphodiesterase—The oligonucleotide, dry and free of salt (6), was dissolved in 0.01 ml of 1 M Tris-chloride buffer (pH 7.5), and the solution was mixed with 0.01 ml of 0.5 M magnesium chloride and 0.08 ml (2 mg per ml) of snake venom phosphodiesterase (7). After incubation for 4 hours at 37°, the solution was spotted directly on paper, and the products were analyzed by paper electrophoresis or two-dimensional paper chromatography (3, 6).

RESULTS

Pancreatic RNase Digest of Alanine-RNA—The results of a typical chromatogram of a pancreatic RNase digest on DEAE-Sephadex are shown in Fig. 1. The fraction of the ultraviolet-absorbing material found in each peak and the analyses of the constituents of the peaks are summarized in Table I. Recoveries of the nucleotides from paper electrophoresis and paper chromatography, as used in the analyses, were 60 to 80%. The analyses were, in general, entirely straightforward, but certain of the

* This work was supported in part by the National Science Foundation and by the National Institutes of Health.

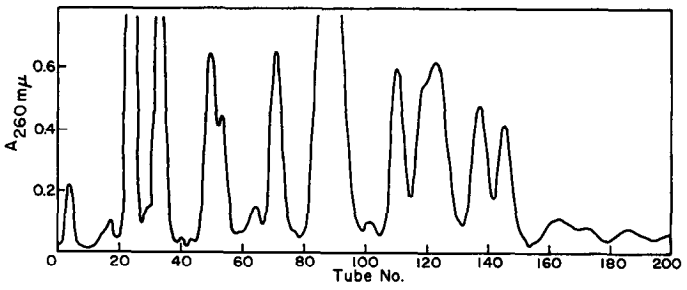

FIG. 1. Chromatography of a 14-hour pancreatic RNase digest of the alanine-RNA on a column (0.25 × 225 cm) of DEAE-Sephadex (See "Experimental Procedure").

oligonucleotides, particularly those containing unusual nucleotides, deserve additional comment.

The peak in tubes 46 to 51 of Fig. 1 contained two dinucleotides which could be separated readily by paper electrophoresis at pH 2.7. The slower moving component was identified readily as ApCp. The faster moving component frequently gave only ψp after alkaline hydrolysis with 0.5 N potassium hydroxide at 37° for 18 hours, although the spectrum of the original dinucleotide was not that of ψp. Since this dinucleotide has the highest electrophoretic mobility of any of the dinucleotides found in a pancreatic RNase digest, and a mixture of all the dinucleotides can be obtained by chromatography of an RNase digest in 7 M urea (Fig. 2, tubes 19 to 25), the unidentified dinucleotide can

TABLE I

Summary of analyses of peaks from chromatography of pancreatic RNase digest on DEAE-Sephadex (Fig. 1)

The tube numbers correspond to those in Fig. 1.

Tube No.	A_{260}[a]	Composition of peak	Yield[b]	Analyses[c]
	% total		*moles/mole*	
1–9	1.0	C	1.0	*A*
10–19	0.5	Trace A + ?		*A*
20–26	13.2	Cp	12.4	*A*
27–30	0.4	ψp!	1.0	*B*
31–36	7.7	ψp	1.0	*A*
		Up	5.0	*A*
37–45	0.5	Traces ?		*A*
46–51	4.3	ApCp	1.0[d]	*C:* Ap, 1.20; Cp, 1.00
		1-MeIpψp	0.5[d]	*C:* 1-MeIp, 1.0; ψp, 1.00 (see text)
52–58	2.5	N^2-DiMeGpCp	1.0	*B:* N^2-DiMeGp, 0.93; Cp, 1.00
59–66	1.2	ApUp	0.4	*B:* Ap, 1.10; Up, 1.00
67–78	5.0	GpCp	1.9	*B:* Gp, 0.88; Cp, 1.00; Ap, 0.15; Up, 0.06
79–98	21.5	1-MeGpGpCp	0.9[d]	*C:* 1-MeGp, 1.17; Gp, 1.30; Cp, 1.00. *D:* 5′-terminal 1-MeG
		GpUp	4.5[d]	*C:* Gp, 0.90; Up, 1.00
		ApGpCp	0.9[d]	*C:* Ap, 1.21; Gp, 1.35; Cp, 1.00. *D:* 5′-terminal A
		ApGpDiHUp	0.9[d]	*C:* Ap, 0.96; Gp, 1.2 (see text). *D:* 5′-terminal A
99–105	0.8	Traces		
106–114	5.0	GpApUp	1.1	*B:* Gp, 1.00; Ap, 0.86; Up, 1.00. *D:* 5′-terminal G
115–120	4.1	IpGpCp	1.0	*B:* Ip, 0.88; Gp, 0.90; Cp, 1.00. *D:* 5′-terminal I
121–131	7.0	GpGpTp	1.0[d]	*C:* Gp, 2.03; Tp, 1.00
		GpGpDiHUp	1.0[d]	*C:* Gp, 1.95 (see text)
132–141	4.7	GpGpApCp	0.9	*B:* Gp, 2.27; Ap, 0.96; Cp, 1.00; Up, 0.1. *E:* ApCp sequence found
142–150	3.8	pGpGpGpCp	0.7	*B:* pGp, 0.81; Gp, 2.14; Cp, 1.00; Trace Ap, Up
151–176	2.5	Mixture		
177–195	1.4	Mixture		
196–264	2.4	Mixture		
4 M Ammonium acetate	10.4	GpGpGpApGpApGpUp	1.0	*B:* Gp, 5.7; Ap, 2.0; Up, 0.47 (see text). *F:* complete sequence

[a] One absorbance unit at 260 mμ (A_{260}) is defined as the amount of material giving an absorbance reading of 1.0 at 260 mμ when dissolved in 1.0 ml and with a light path of 1.0 cm.

[b] The number of moles of a fragment per mole of RNA is based on the observed A_{260} of the fragment divided by the A_{260} calculated for 1 mole of the fragment per mole of RNA based on the molar extinction coefficients listed below, with allowances for hypochromicity of oligonucleotides (18), and assuming a molecular weight of 26,000 for the sodium salt of the alanine-RNA, with 20 A_{260} for 1 mg of RNA (approximately 15% water). The molar extinction coefficients (pH 7) used were: Ap, 15.0 × 10^3; Cp, 7.6 × 10^3; Gp, 11.4 × 10^3; Ip, 7.1 × 10^3; Up, 10.0 × 10^3 (19); 1-MeGp, 11.4 × 10^3; N^2-DiMeGp, 15.0 × 10^3 (20); Tp, 8.9 × 10^3 (21); and ψp, 9.6 × 10^3 (22). 1-MeIp was assumed equal to Ip.

[c] Analyses: *A*, identification by paper chromatographic behavior (6) (isopropyl alcohol-water-NH_3 (23) and isopropyl alcohol-water-HCl (24)) and spectra of nucleotides; *B*, identification by alkaline hydrolysis (0.5 N potassium hydroxide at 37° for 18 hours (6)), followed by *A*; C, high voltage paper electrophoresis (pH 2.7 in 20% ammonium acetate (25)), followed by *B*; *D*, nucleotide sequence determined by complete digestion with snake venom phosphodiesterase; *E*, nucleotide sequence determined by digestion with Taka-Diastase RNase T1; *F*, nucleotide sequence determined by partial degradation with snake venom phosphodiesterase (12).

[d] The amounts of these oligonucleotides were estimated after further fractionation of the peak to determine the proportions of the different constituents.

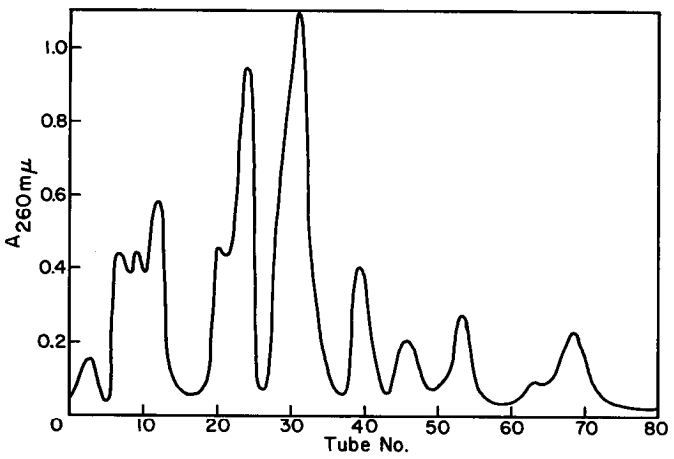

FIG. 2. Chromatography of a 4-hour pancreatic RNase digest of the alanine-RNA on a column (0.35 × 30 cm) of DEAE-cellulose in 7 M urea (see "Experimental Procedure").

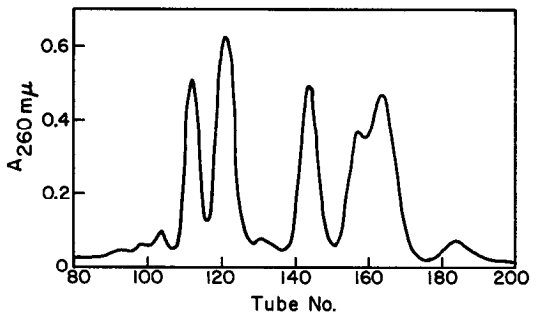

FIG. 3. Rechromatography of the pancreatic RNase trinucleotides (Fig. 2, tubes 27 to 35) on DEAE-Sephadex with a nonlinear gradient produced from 0.04 M, 0.20 M, 0.50 M, and 1.0 M ammonium carbonate in four chambers of a Varigrad (4).

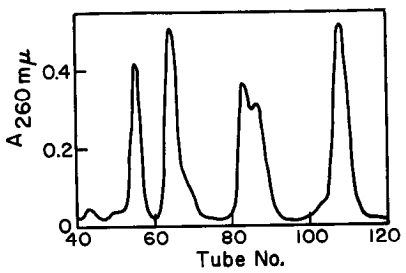

FIG. 4. Rechromatography of the pancreatic RNase trinucleotides (Fig. 2, tubes 27 to 35) on DEAE-cellulose with a linear gradient produced from 7 M urea, pH 3.8 (adjusted with formic acid), and 0.08 M sodium chloride in 7 M urea, pH 3.8.

be isolated simply by paper electrophoresis of the mixture of dinucleotides. Variations in the hydrolysis conditions of the unidentified dinucleotide indicated that a good yield of a second component could be obtained by hydrolyzing the dinucleotide in 0.3 M potassium hydroxide for 2 hours at 37°. The second component had an electrophoretic mobility equivalent to that of Ip in ammonium formate at pH 3.5. Its mobility in two-dimensional paper chromatography (3, 6) corresponded to that expected of a methylated inosinic acid. Its absorption maximum was at 250 mμ, both at pH 2 and at pH 12. Hydrolysis of the dinucleotide with N hydrochloric acid for 30 minutes at 100° gave a free base with the spectral properties reported for 1-methylhypoxanthine (8). The dinucleotide is therefore 1-MeIpψp. The isolation of 1-methylinosine from bulk yeast transfer RNA has recently been reported by Hall (9). It will be noted in Table I that the yield of 1-MeIpψp is considerably less than 1 mole per mole of RNA. However, a tetranucleotide containing the 1-MeIpψp sequence is obtained in good yield in the RNase T1 digest, indicating that this sequence undoubtedly is present in the alanine-RNA molecule.

The two trinucleotides that contain dihydrouridylic acid (DiHUp) were rather difficult problems. Initially, the only indication of something unusual was the observation that certain of the trinucleotide fractions gave "poor" analyses for nucleotide composition. Under many conditions, ApGpDiHUp moves with ApGpCp, and since DiHUp does not absorb at 260 mμ, analysis of this mixture of trinucleotides by alkaline hydrolysis and paper chromatography gave a composition approximating $(Ap)_2(Gp)_2Cp$. Although the analysis for Cp was low, the results were originally interpreted as indicating the presence of two ApGpCp sequences in the alanine-RNA (3). After extensive study of the fractionation of the trinucleotides in an RNase digest, conditions were found that gave partial resolution of each of the unusual trinucleotides (Figs. 3 and 4). Analyses of the various peaks (Tables II and III) indicated that there were two unusual trinucleotides that contained no normal, ultraviolet-absorbing pyrimidine nucleotide. Degradation of these trinucleotides with RNase T1, followed by chromatography, established the presence of a compound that contained organic phosphate, that absorbed light at 230 mμ without any characteristic spectrum, and that gave a test for a dihydropyrimidine (10). Treatment of the unknown nucleotide with alkaline phosphatase gave a compound with the chromatographic properties of authentic dihydrouridine (11). A method has been developed for the isolation of dihydrouridylic acid from bulk yeast transfer RNA (11).

TABLE II

Identities of trinucleotide peaks in Fig. 3

The tube numbers correspond to those in Fig. 3.

Tube No.	Composition of peak
109–115	1-MeGpGpCp
118–125	ApGpCp
	ApGpDiHUp
140–148	GpApUp
153–159	GpGpDiHUp
160–170	IpGpCp
	GpGpTp

TABLE III

Identities of trinucleotide peaks in Fig. 4

The tube numbers correspond to those in Fig. 4.

Tube No.	Composition of peak
53–57	ApGpCp
62–67	1-MeGpGpCp
	IpGpCp
81–84	GpApUp
86–90	ApGpDiHUp
105–111	GpGpDiHUp
	GpGpTp

The largest oligonucleotide found in a pancreatic RNase digest of the alanine-RNA has the structure GpGpGpApGpApGpUp (12). This octanucleotide is not eluted from the DEAE-Sephadex column (Fig. 1) by 1 M ammonium carbonate, but is eluted (Table I) with 4 M ammonium acetate. The yield of this octanucleotide can be calculated more accurately from the peak obtained when the digest is chromatographed on a short DEAE-cellulose column in 7 M urea (Fig. 2), and usually approximates 0.9 mole per mole of RNA. Since this octanucleotide is degraded slowly by pancreatic RNase, as might be anticipated from the reported lack of specificity of this enzyme (13), a good yield of the octanucleotide is obtained only if the time of digestion is limited. It will be noted in Table I that the analysis of this octanucleotide is low in Up. This may be due to experimental difficulties, but recent experiments suggest that part of the octanucleotide contains DiHUp in place of Up.

The yield of pGpGpGpCp (Table I and Fig. 2, tubes 52 to 55) is somewhat lower than the yields of the other fragments. This probably indicates that part of the alanine-RNA has lost the 5′-terminal phosphate, an explanation that is consistent with the finding that the tetranucleotide peak from the small urea column (Fig. 2, tubes 38 to 42) contains some GpGpGpCp in addition to the major component, GpGpApCp. The small peak of pentanucleotide obtained from the urea column (Fig. 2, tubes 44 to 48) has been shown by rechromatography on DEAE-Sephadex to be a mixture of at least six different pentanucleotides, which presumably arise from small amounts of impurities in the alanine-RNA. The small peak at tubes 61 to 64 in Fig. 2 appears to be very closely related to the octanucleotide, but because of the small amount it has not been studied extensively.

Taka-Diastase RNase T1 Digest of Alanine-RNA—The results of chromatography of a Taka-Diastase RNase T1 digest of the alanine-RNA on DEAE-Sephadex are shown in Fig. 5. The identities of the peaks are summarized in Table IV.

The results of chromatography of a similar digest on a long DEAE-cellulose column in 7 M urea are shown in Fig. 6. The recovery of large fragments is better on this column and the resolution of the oligonucleotides is somewhat different. A detailed summary of the analyses of the oligonucleotide fractions from the DEAE-cellulose column (Fig. 6) is given in Table V. Certain of the oligonucleotides require special comment.

As is shown by the analyses of the first two peaks of Fig. 6 (tubes 37 to 41 and 42 to 48; see Table V), the methylated guanylic acids are obtained as dinucleotides terminating in 2′,3′-cyclic phosphates. These cyclic phosphates are opened very slowly, if at all, under the T1 digestion conditions used. The particularly slow rate of attack on the linkage next to 1-MeGp is shown also by the isolation of a small amount of Up1-MeGpGp among the trinucleotides. This behavior of the methylated guanylic acids is similar to the behavior described by others (14, 15).

The dihydrouridylic acid (DiHUp) residues described above are also found in trinucleotides in the T1 digest. These trinucleotides chromatograph as separate peaks on the DEAE-cellulose column in 7 M urea (Fig. 6, tubes 78 to 82 and 83 to 87). Dihydrouridylic acid can be isolated readily after enzymatic digestion of these peaks.

The amino acid acceptor end of the alanine-RNA is the fragment with the structure UpCpCpApCpC, since this is the only T1 digest fragment that does not have a terminal Gp residue. The 3′-terminal cytidine is consistent with the finding of cytidine in the RNase digest (Table I) and indicates that most of the alanine-RNA molecules isolated from commercial bakers' yeast have lost the terminal adenylic acid residue. (The terminal adenylic acid residue is replaced under the usual assay conditions before alanine is attached.)

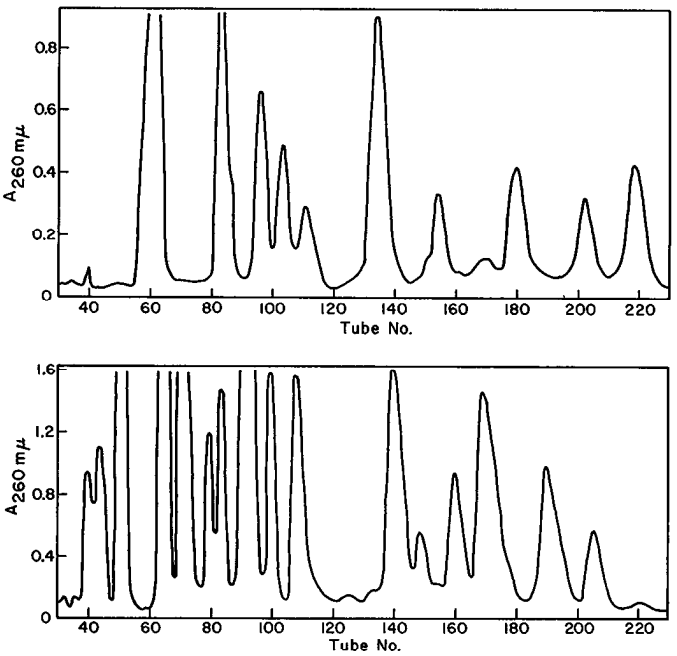

FIG. 5 (*upper*). Chromatography of a Taka-Diastase RNase T1 digest of the alanine-RNA on DEAE-Sephadex with a nonlinear gradient produced from 0.04 M, 0.2 M, 0.4 M, and 1.0 M ammonium carbonate in four chambers of a Varigrad.

FIG. 6 (*lower*). Chromatography of a Taka-Diastase RNase T1 digest of the alanine-RNA on DEAE-cellulose with a nonlinear gradient produced from 7 M urea, 0.4 M sodium acetate (pH 7.5) in 7 M urea, 7 M urea, and 1.0 M sodium acetate (pH 7.5) in 7 M urea in four chambers of a Varigrad.

TABLE IV

Identities of peaks from chromatography of RNase T1 digest on DEAE-Sephadex (Fig. 5)

The tube numbers correspond to those in Fig. 5.

Tube No.	Composition of peak
55–66	Gp
	CpN^2-DiMeGp!
	Up1-MeGp!
80–89	CpGp
	pGp
93–100	ApGp
101–107	UpGp
	DiHUpCpGp
108–117	DiHUpApGp
130–140	Up1-MeGpGp
	UpCpCpApCpC
	UpApGp
	Cp1-MeIpψpGp
152–160	Tpψ pCpGp
175–187	ApCpUpCpGp
198–208	Up(Cp,Up)CpCpGp
214–225	ApUpUpCpCpGp

TABLE V

Summary of analyses of peaks from chromatography of RNase T1 digest on DEAE-cellulose (Fig. 6)

Tube No.	A_{260}[a]	Composition of peak	Yield[b]	Analyses[c]
	% total		*moles/mole*	
37–41	2.1	CpN^2-DiMeGp!	0.9	*B:* Cp, 0.92; N^2-DiMeGp, 1.00
42–48	3.7	Up1-MeGp!	1.0	*B:* Up, 1.3; 1-MeGp, 1.00; trace CpMeIp!
49–58	15.9	Gp	10.5	*A*
59–68	8.3	CpGp	3.5	*B:* Cp, 1.1; Gp, 1.00
69–77	10.3	ApGp	2.3	*B:* Ap, 1.0; Gp, 1.00
		UpGp	0.9	*B:* Up, 0.9; Gp, 1.00
78–82	2.8	DiHUpCpGp	1.2	*B:* Cp, 1.4; Gp, 1.00 (see text). *D:* 5′-terminal DiHU
83–87	3.2	DiHUpApGp	0.9	*B:* Ap, 0.9; Gp, 1.00 (see text). *G:* ApGp sequence found
88–97	12.4	UpApGp	1.0	*C:* Up, 0.9; Ap, 0.8; Gp, 1.00. *D:* 5′-terminal U
		pGp	1.0	*A, C*
		UpCpCpApCpC	1.0	*C:* Up, 0.8; Cp, 2.9; Ap, 1.0; C, 1.00. *D:* 5′-terminal U. *F:* complete sequence
		Up1-MeGpGp		*C:* trace only; mostly as dinucleotide above
98–104	4.1	Cp1-MeIpψpGp	0.9	*B:* Cp, 0.7; 1-MeIp, 0.8; ψp, 0.9; Gp, 1.00. *G:* MeIpψp sequence found
105–120	5.9	TpψpCpGp	0.8	*B:* Tp, 1.2; ψp, 0.9; Cp, 1.2; Gp, 1.00. *H:* complete sequence
121–135	1.5	Traces ?		
136–145	6.7	ApCpUpCpGp	1.0	*B:* Ap, 0.8; Cp, 1.8; Up, 1.0; Gp, 1.00. *D:* 5′-terminal A. *G:* ApCp sequence found. *F:* complete sequence
146–152[d]	2.0			
153–156	0.6	Traces ?		
157–165	3.7	Up(Cp,Up)CpCpGp	0.7	*B:* Up, 2.0; Cp, 2.9; Gp, 1.00; Ap, 0.3. *D:* 5′-terminal U. *F:* partial sequence
166–183	8.2	ApUpUpCpCpGp	1.1	*B:* Ap, 0.9; Up, 1.9; Cp, 1.9; Gp, 1.00. *D:* 5′-terminal A. *G:* ApUp sequence found. *F:* complete sequence
184–200	5.7	Cp(Cp,Cp,Up)CpUpUpIp!	0.9	*B:* Cp, 3.3; Up, 3.0; Ip, 1.00. *D:* 5′-terminal C. *F:* Partial sequence
201–214	2.7	Cp(Cp,Cp,Up)CpUpUpIp		
215–280	0.6	Traces ?		

[a] See Table I, footnote *a*.

[b] See Table I, footnote *b*.

[c] Analyses *A* through *F* are defined in Table I. Analysis *G*, nucleotide sequence determined by digestion with pancreatic RNase; *H*, nucleotide sequence determined by partial digestion with micrococcal nuclease (6).

[d] Probably related to hexanucleotide in tubes 157–165 (see text).

The identification of the tetranucleotide TpψpCpGp and evidence that this is a common sequence in all yeast transfer RNAs have been described (6).

The hexanucleotide Up(Cp,Up)CpCpGp is obtained in somewhat low yield (Fig. 6 and Table V, tubes 157 to 164). This hexanucleotide, however, is very closely related to the oligonucleotide in the preceding peak at tubes 146 to 152. There is some indication that the two peaks differ only in the 5′-terminal nucleotide, the first peak containing DiHUp, and the second, Up. The sum of the two peaks approximates 1 mole per mole of RNA.

The largest fragment obtained in the T1 digest, the octanucleotide Cp(Cp,Cp,Up)CpUpUpIp, is not eluted from the DEAE-Sephadex column by the gradient of ammonium carbonate, but is obtained readily from the DEAE-cellulose column in 7 M urea. From the latter column, the octanucleotide is obtained as a double peak (Fig. 6, tubes 184 to 200 and 201 to 214) in which the first peak has a terminal 2′,3′-cyclic phosphate. Continued digestion with RNase T1 results in a decrease in the size of the first peak and an increase in the second.

Two short sequences of pyrimidine nucleotides, enclosed within parentheses in the above formulas, remain to be established in these last two large oligonucleotides.

Comparison of Analyses of Two Digests—Examination of the results of the analyses of the two digests summarized in Tables I and V indicates that there is excellent agreement between the two analyses, both in nucleotide composition and in nucleotide sequences. For example, the pancreatic RNase digest fragments indicate that the RNA contains two sequences of ApGp preceded by a pyrimidine, and these are found in the RNase T1 digest as UpApGp and DiHUpApGp. The structure of the octanucleotide obtained in the pancreatic RNase digest indicates that 2 moles of the dinucleotide ApGp should be present in the T1 digest, and they are. The GpApUp in the pancreatic RNase digest indicates that a fragment in the T1 digest should begin with ApUp–, and such a sequence, ApUpUpCpCpGp, appears. The GpGpApCp in the pancreatic digest indicates that a T1 fragment should begin with ApCp–, and ApCpUpCpGp is found. The only inconsistency between the two digests is the small amount of ApUp in the pancreatic RNase digest which is not accounted for in the T1 RNase digest. Presumably, the ApUp comes from impurities in the RNA.

The excellent agreement between the two digests is summarized in simplified form in Fig. 7, in which the fragments from the pancreatic RNase digest are arranged above the corresponding RNase T1 fragments. The minor nucleotides and certain unique sequences give a number of overlaps between the two digests. Present data give the structure of the RNA in terms

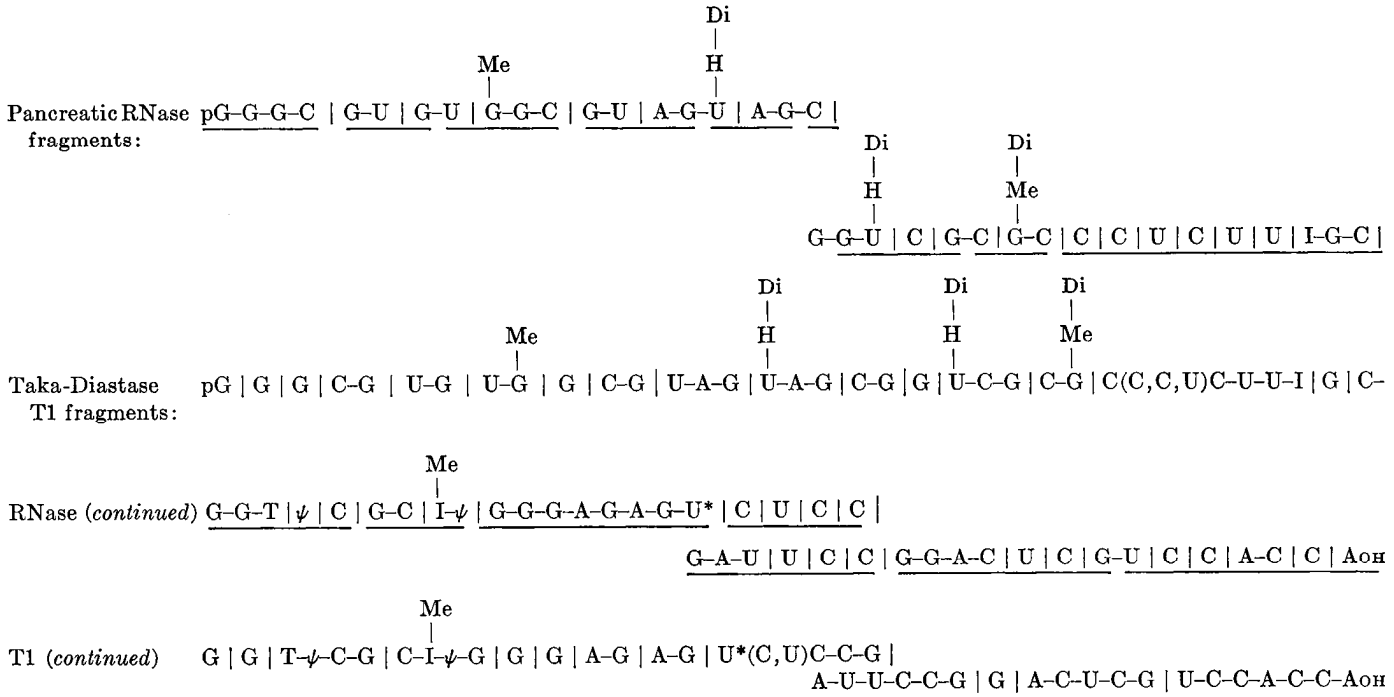

FIG. 7. One of many possible arrangements of the pancreatic RNase and RNase T1 digest fragments that shows the overlaps between the two digests. The RNA molecule is accounted for by the 16 oligonucleotide sequences indicated by the *solid lines*. Only the positions of the two terminal sequences are known. *Vertical lines* indicate the position of enzymatic attack. The *asterisk* indicates that the uridine may be partially substituted by DiHU.

of 16 fragments, which are shown by the *solid line segments* drawn between the *lines of fragments*. The positions of the 5′- and 3′-terminal fragments are known (because of the 5′-phosphate in the first and the absence of a 3′-phosphate in the latter), but otherwise the over-all arrangement of the fragments in Fig. 7 is only one of many possible arrangements.

DISCUSSION

The good yields of the various oligonucleotide fragments obtained in the pancreatic RNase and RNase T1 digests, as well as the excellent agreement found between the compositions of the two digests, indicate that the alanine-RNA structure is accounted for by the 16 oligonucleotide sequences summarized in Fig. 7. The alanine-RNA is therefore composed of approximately 77 nucleotides.

The major problem that remains is to arrange the 16 oligonucleotide sequences in longer sequences until the complete nucleotide sequence is established.

The results reported here are in general agreement with the partial analyses reported by Ingram and Sjöquist (16) and Armstrong *et al.* (17).

SUMMARY

The fragments obtained by pancreatic ribonuclease and Taka-Diastase ribonuclease T1 digestion of the yeast alanine ribonucleic acid have been separated and identified. The analyses indicate that the alanine-RNA is composed of a polynucleotide chain containing 77 nucleotide residues. Nine unusual nucleotide residues are present, including inosinic acid, 1-methylinosinic acid, and 5,6-dihydrouridylic acid. Overlaps of sequences found in fragments obtained from the two digests are sufficient to permit description of the structure in terms of 16 oligonucleotide sequences. The identities of the two end fragments are known, but the arrangement of the other 14 fragments in the alanine-RNA remains to be established.

REFERENCES

1. APGAR, J., HOLLEY, R. W., AND MERRILL, S. H., *J. Biol. Chem.*, **237**, 796 (1962).
2. HOLLEY, R. W., APGAR, J., MERRILL, S. H., AND ZUBKOFF, P. L., *J. Am. Chem. Soc.*, **83**, 4861 (1961).
3. HOLLEY, R. W., APGAR, J., EVERETT, G. A., MADISON, J. T., MERRILL, S. H., AND ZAMIR, A., *Cold Spring Harbor Symposia Quant. Biol.*, **28**, 117 (1963).
4. PETERSON, E. A., AND SOBER, H. A., *Anal. Chem.*, **31**, 857 (1959).
5. TOMLINSON, R. V., AND TENER, G. M., *J. Am. Chem. Soc.*, **84**, 2644 (1962); *Biochemistry*, **2**, 697 (1963).
6. ZAMIR, A., HOLLEY, R. W., AND MARQUISEE, M., *J. Biol. Chem.*, **240**, 1267 (1965).
7. KELLER, E. B., *Biochem. and Biophys. Research Communs.*, **17**, 412 (1964).
8. ELION, G., *J. Org. Chem.*, **27**, 2478 (1962).
9. HALL, R. H., *Biochem. and Biophys. Research Communs.*, **13**, 394 (1963).
10. FINK, R. M., CLINE, R. E., MCGAUGHEY, C., AND FINK, K., *Anal. Chem.*, **28**, 4 (1956).
11. MADISON, J. T., AND HOLLEY, R. W., *Biochem. and Biophys. Research Communs.*, **18**, 153 (1965).
12. HOLLEY, R. W., MADISON, J. T., AND ZAMIR, A., *Biochem. and Biophys. Research Communs.*, **17**, 389 (1964).
13. BEERS, R. F., JR., *J. Biol. Chem.*, **235**, 2393 (1960).
14. DÜTTING, D., AND ZACHAU, H. G., *Z. physiol. Chem.*, **336**, 132 (1964).
15. STAEHELIN, M., *Biochim. et Biophys. Acta*, **87**, 493 (1964).
16. INGRAM, V. M., AND SJÖQUIST, J. A., *Cold Spring Harbor Symposia Quant. Biol.*, **28**, 133 (1963).

17. Armstrong, A., Hagopian, H., Ingram, V. M., Sjöquist, I., and Sjöquist, J., *Biochemistry*, **3**, 1194 (1964).
18. Staehelin, M., *Biochim. et Biophys. Acta*, **49**, 20 (1961).
19. Beaven, G. H., Holiday, E. R., and Johnson, E. A., in E. Chargaff and J. N. Davidson (Editors), *The nucleic acids*, *Vol. I*, Academic Press, Inc., New York, 1955, p. 493.
20. Smith, J. D., and Dunn, D. B., *Biochem. J.*, **72**, 294 (1959).
21. Littlefield, J. W., and Dunn, D. B., *Biochem. J.*, **70**, 642 (1958).
22. Cohn, W. E., *J. Biol. Chem.*, **235**, 1488 (1960).
23. Markham, R., and Smith, J. D., *Biochem. J.*, **52**, 552 (1952).
24. Wyatt, G. R., *Biochem. J.*, **48**, 584 (1951).
25. Litt, M., and Ingram, V. M., *Biochemistry*, **3**, 560 (1964).

Nucleotide Sequences in the Yeast Alanine Transfer Ribonucleic Acid

Robert W. Holley, George A. Everett, James T. Madison and Ada Zamir

J. Biol. Chem. 1965, 240:2122-2128.