The nucleotide sequence of phenylalanine tRNA from Bacillus subtilis

H.-H. Arnold* and G. Keith

Physiologisches-chemisches Institut der Universität Erlangen-Nürnberg, 852 Erlangen, Wasserturmstr. 5, GFR and, Institut de Biologie Moléculaire et Cellulaire du C.N.R.S., 15, rue Descartes, Université Louis Pasteur, 6700 Strasbourg France.

Received 23 May 1977

ABSTRACT

The nucleotide sequence of $tRNA^{Phe}$ from Bacillus subtilis W 23 has been determined using ^{32}P labeled tRNA. This is the second B. subtilis tRNA so far reported. The nucleotide sequence was found to be pG-G-C-U-C-G-G-U-A-G-C-U-C-A-G-U-D-G-G-D-A-G-A-G-C-A-A-C-G-G-A-C-U-Gm-A-A-ms^2i^6A-A-ψ-C-C-G-U-G-U-m^7G-U-C-G-G-C-G-G-T-ψ-C-G-A-U-U-C-C-G-U-C-C-C-G-A-G-C-C-A-C-C-A_{OH}.

INTRODUCTION

Most of the primary structures of phenylalanine specific tRNAs determined so far have been those of eucaryotic organisms (1). The nucleotide sequences of $tRNA^{Phe}$ from E. coli (2) and from B. stearothermophilus (3) are the only bacterial phenylalanine tRNA structures known. Since $tRNA^{Phe}$ from B. subtilis can be charged by the phenylalanine tRNA ligases of E. coli, B. stearothermophilus and yeast to the same extent as by the homologous ligase (unpublished results), we wanted to check whether the nucleotide sequences believed to be involved in the aminoacylation process are similar to those proposed for yeast $tRNA^{Phe}$ (4) and yeast $tRNA^{Val}$ (5), which is also charged by the phenylalanine tRNA ligase.

In previous papers the existence of additional isoacceptors of $tRNA^{Phe}$ in B. subtilis as a result of certain growth conditions was described (6-8). The analysis of both isoacceptors showed differences solely in the content of two modified nucleosides. Nevertheless, the knowledge of the total primary structure is necessary for deriving conclusions on the origin of the newly arising species. In addition to the reasons above, the unavailability of known tRNA structures from non-thermophilic organisms, but closely related to the thermophilic B. stearothermophilus for comparison of their tRNA structures, prompted the determination of the tRNA structure of the grampositive microorganism B. subtilis.

MATERIALS AND METHODS

Pancreatic RNAse and snake venom phosphodiesterase were obtained from

Worthington Biochemicals Corp. RNAse T_1, T_2 and U_2 (Sankyo) were purchased from Calbiochem. Bacterial alkaline phosphatase (E. coli) was from Boehringer. DEAE-cellulose paper was Whatman DE 81. Cellulose acetate strips and cellulose thin layer plates were purchased from Schleicher and Schuell. All other reagents and solvents were from Merck, Germany. [^{32}P] orthophosphate was obtained from Radiochemical Centre, Amersham.

Nucleic acids labeled with [^{32}P] were extracted from B. subtilis W 23 grown in [^{32}P] orthophosphate containing minimal medium (9) under aerobic conditions. The harvested cells were suspended in buffer containing sodium dodecylsulfate and further treated with phenol (10). Crude tRNA was prepared as described (11). The phenylalanine specific tRNA was purified by two successive column chromatographies: (1) BD-cellulose and (2) RPC-5 material as described previously (11). Phenylalanine tRNA together with $tRNA^{Tyr}$ and one species each of $tRNA^{Leu}$ and $tRNA^{Ser}$ were separated from the bulk tRNAs on BD-cellulose. These four tRNAs are eluted together in 1,5 M sodium chloride buffer containing 0-30% gradient of ethanol, due to their hydrophobicity. In the second step, the tRNAs were further separated on a RPC-5 column at pH 7.5. Final purification of $tRNA^{Phe}$ was achieved by disc electrophoresis on 10% acrylamide slab gels 40 cm in length (12). In some preparations fractionation by RPC-5 chromatography was replaced by two dimensional gel electrophoresis as described earlier (12).

Enzymatic degradation of the [^{32}P] tRNA and also identification of the resulting oligonucleotides were done by standard procedures (13,14). Larger oligonucleotides derived by limited T_1 treatment of the tRNA were separated by polyacrylamide gel electrophoresis (15% or 20%) containing 7 M urea (15).

For identification of modified nucleotides, the complete tRNA and the oligonucleotides were digested extensively with RNAse T_2. The hydrolysates were subjected to two dimensional thin layer chromatography in solvent systems described (16). The identity of 2-thiomethyl,6-isopentenyl adenosine was determined by the UV absorption spectrum of the nucleoside isolated from non-radioactive $tRNA^{Phe}$ as described (8).

RESULTS

The separation of the mononucleotides obtained by complete digestion of B. subtilis $tRNA^{Phe}$ with RNAse T_2 is shown in figure 1. All modified nucleotides occurring in $tRNA^{Phe}$ are well separated. The dinucleotide GM-Ap was cochromatographed with the same fragment obtained from unlabeled $tRNA^{Phe}$. The identity has been proven after digestion with snake venom phosphodiesterase and bacterial phosphates on two dimensional TLC and by the position on the original chromatogram (16).

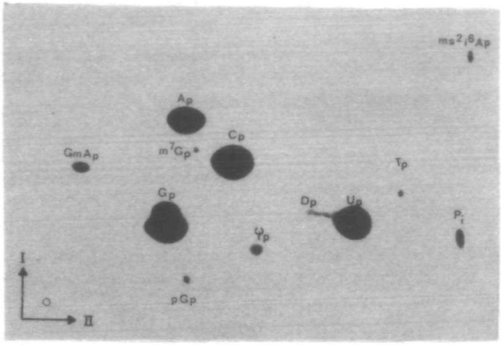

Figure 1.
Autoradiograph of a two dimensional separation of a complete RNAse T_2 digestion of B. subtilis $tRNA^{Phe}$. Cellulose thin layer chromatography was carried out using isobutyric acid - 0.5 M ammonia (5:3, v/v) as the solvent in the 1. dimension and isopropanol - conc. HCl - water (70:15:15, v/v) in the 2. dimension (16). All nucleotides found are indicated in the figure.

The [^{32}P] $tRNA^{Phe}$ was digested with RNAse T_1 and the oligonucleotides were separated by bidimensional high voltage electrophoresis. Figure 2 shows the resulting fingerprint. Seventeen fragments were found.

The analysis of T_1 RNAse digestion products are described in Table I. Those oligonucleotides which could not be sequenced by the methods described in Table I, were investigated further as indicated in Table III. The deduced sequence of each spot is also given in Fig. 2.

Four of the T_1 RNAse digest endproducts had to be analyzed by partial digestion with venom phosphodiesterase in the presence of alkaline phosphatase in order to establish the questionable sequences of pyrimidines. Either by determination of the M-values (13) or by compositions analysis of each intermediate found, we could determine unambiguously the fragments C-U-C-Gp, C-U-C-A-Gp, U-m^{7}G-U-C-Gp, and A-U-U-C-C-Gp.

The sequence A-C-U-Gm-A-A-ms^{2}i^{6}A-A-ψ-C-C-Gp could be deduced from the combined results of Table I and Table III. Digestion with RNAse U_2, micrococcal nuclease and pancreatic RNAse provided sufficient overlapping fragments to establish the correct sequence. After U_2 RNAse digestion the 3' end fragment C-C-A-C-C-A_{OH} gave C-C-Ap and C-C-A separated by paper electrophoresis.

The digestion of $tRNA^{Phe}$ with pancreatic RNAse and subsequent electrophoresis resulted in the fingerprint shown in Figure 2. Fourteen spots were obtained corresponding to eighteen fragments since some of them are not separated in the system used. The sequences indicated in Figure 3 were derived from the results shown in Tables II and III. The isomeric structures of tetranucleotides G-G-A-Cp and G-A-G-Cp, not separated in the fingerprint, could be established easily according to the T_1 RNAse digestion products A-Cp, A-Gp, 3Gp and 1Cp.

The spot containing G-G-Up, G-G-Dp, G-G-Tp and pG-G-Cp provided, after

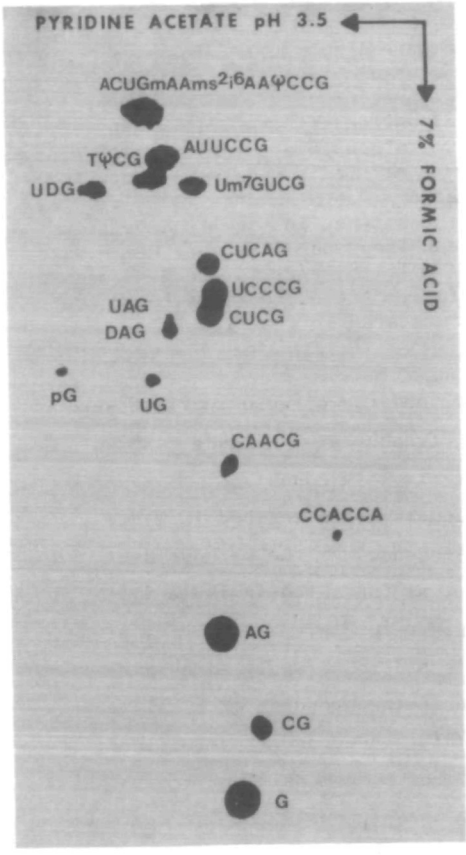

Figure 2
Fingerprint of complete RNAse T_1 digest of phenylalanine tRNA.

treatment with alkaline phosphatase, two separate spots on the electrophoresis.

These spots after digestion with venom phosphodiesterase gave 3pG, pU, pD, pT, for one of the spots and pG and pC for the other. This analysis together with the respective position of the oligonucleotides on the fingerprint proved the sequence of the individual fragments.

The sequence containing the anticodon was determined as described for the analogous structure of the T_1 RNAse fingerprint. All other oligonucleotides derived by complete pancreatic RNAse digestion were unambiguously determined by the methods given in Table II. The exact molar amounts of the mono- and dinucleotides were calculated for the T_1 products from the pancreatic oligonucleotides and for the pancreatic products from the T_1 oligonucleotides

Overlapping fragments of the $tRNA^{Phe}$ (see Figure 4) were obtained by partial T_1 RNAse digestion. The larger oligonucleotides were separated on polyacrylamide gels. One half of the eluted material from each band was digested

Table I

Analysis of endproducts of complete T_1-RNase digestion.

Fragments	T_2-RNase products					panc. RNase products	5' end	molar yield	
	A	C	G	U	others			found	calc.
Gp			X			-	-	8	6
AGp	1		1			A-G-	-	3	2
CGp		1	1			C-(1), G-(1)	-	1	1
UGp			1	1		U-(1), G-(1)	-	1,5	1
pGp					pG-(1)	-	-	0,8	1
U-D-Gp			1	1	D-(1)	U-(1), D-(1), G-(1)	Up	1	1
D-A-Gp	1		1		D-(1)	D-(1), A-G-(1)	-	1	1
U-A-Gp	1		1	1		U-(1), A-G-(1)	-	1	1
C-U-C-Gp*		2	1	1		C-(2), U-(1), G-(1)	Cp	1	1
T-ψ-C-Gp		1	1		ψ-(1), T-(1)	T-(1), ψ-(1), C-(1), G-(1)	Tp	1	1
C-A-A-C-Gp	2	2	1			C-(1),A-A-C-(1), G-(1)	Cp	1	1
U-C-C-C-Gp		3	1	1		U-(1), C-(3), G-(1)	Up	1	1
C-U-C-A-Gp*	1	2	1	1		C-(2), U-(1), A-G-(1)	Cp	1	1
U-m^7G-U-C-Gp*		1	1	2	m^7C-(1)	U-(1), C-(1), G-(1), m^7G-U-(1)	Up	0,7	1
A-U-U-C-C-Gp*	1	2	1	2		A-U-(1), U-(1), C-(2), G-(1)	Ap	1	1
A-C-U-Gm-A-A-ms^2i^6A-A-ψ-C-C-Gp*	3	3	1	1	ms^2i^6A-(1) GmA-(1), ψ-(1)	A-C-(1), U-(1), C-(2), $_2C_6$(1), Gm-A-A-ms^2i^6A-A-ψ-(1)	Ap	1	1
C-C-A-C-C-A_{OH}	1	4				A-C-(1), C-(3)	Cp	0,7	1

*Analysis for final sequence determination is described in Table III.

Table II

Analysis of endproducts of complete panc. RNAse digestion.

Fragments	T_2-RNAse products					T_1-RNAse products	5' end	molar yield	
	A	C	G	U	others			found	calc.
Cp		X				-	-	13	14
Up, ψp, Dp,				X	ψ-(1),D-(1)	-	-	4	6
m^7G-Up				1	m^7G-(1)	-	m^7Gp	0,5	1
A-Cp	1	1				-	-	0,6	1
G-Up			1	1		-	-	3	3
A-A-Cp	2	1				-	-	1,5	1
A-G-Cp	1	1	1			A-G-(1),C-(1)	Ap	1,4	1
G-G-Cp		1	2			-	-	1,3	1
G-A-Up	1		1	1		C-(1),A-U-(1)	Gp	1	1
A-G-Up	1		1	1		A-G-(1),U-(1)	Ap	1	1
G-G-Up*			2	1		-	-	1	1
C-G-Dp*			2		D-(1)	-	-	1	1
G-G-Tp*			2		T-(1)	-	-	1	1
pG-G-Cp*		1	1		pGp(1)	-	pGp	0,7	1
G-A-G-Cp+	1	1	2			G-(1),A-G-(1) C-(1)	Gp	1	1
G-G-A-Cp+	1	1	2			G-(2),A-C-(1)	Gp	1	1
A-G-A-G-Cp	2	1	2			A-G-(2),C-(1)	Ap	1	1
Gm-A-A-ms^2i^6A-A-ψp*	2				ms^2i^6A-(1) GmA-(1)	-	-	1	1

+Since not separated both fragments were analyzed together.

*Analysis for final sequence determination is given in table III.

completely with T_1 RNAse, the other half with pancreatic RNAse.

The determination of the sequences of the oligonucleotides taken together with information from partially digested fragments enabled us to derive the primary structure of B. subtilis $tRNA^{Phe}$ (Figures 4 and 5).

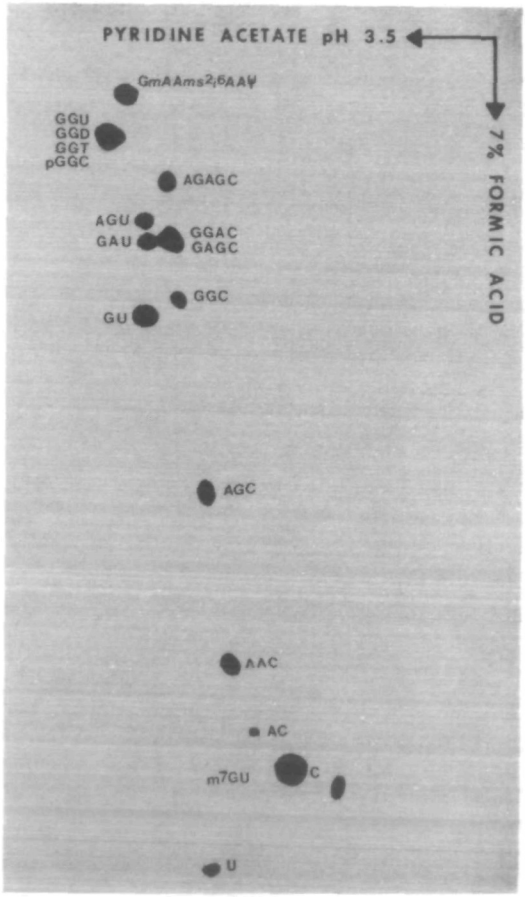

Figure 3
Fingerprint of complete pancreatic RNAse digest of phenylalanine tRNA. Spot m^7GUp is lost by the photographic reproduction.

DISCUSSION

The pehnylalanine specific tRNA from B. subtilis has a chain length of 76 nucleotides including eight modified ones. The extent of modification thus appears to be similar as in $tRNA^{Phe}$ of E. coli and B. stearothermophilus (2,3) but different from the initiator tRNA of B. subtilis which contains only three modified bases (17). The hydrophobic nucleotide in the position adjacent to the 3' end of the anticodon in B. subtilis $tRNA^{Phe}$, like that of E. coli, is occupied by the hypermodified adenosine ms^2i^6A. It should be noted that also an isoaccepting $tRNA^{Phe}$ can be found in B. subtilis which is missing both modifications in the anticodon loop (8). Whether the lack of these modifications is the only difference between the isoacceptors of $tRNA^{Phe}$ will be investigated by comparison of fingerprints of these species.

The primary structures of B. subtilis $tRNA^{Phe}$ and E. coli $tRNA^{Phe}$ differ in 24 residues. Much less variation can be observed between $tRNA^{Phe}$ of B. subtilis and B. stearothermophilus. The well known microbiological

pGGCUCGGUAGCUCAGUDGGDAGAGCAACGGACUGmAAms$^{2}i^{6}$AAUCCGUGUm^{7}GUCGGCGGTψCGAUUCCGUCCCGAGCCACCA

Figure 4
Summary of overlapping fragments from complete and partial RNAse digestions. Overlaps yielded from the complete digestion products are drawn above the sequence.

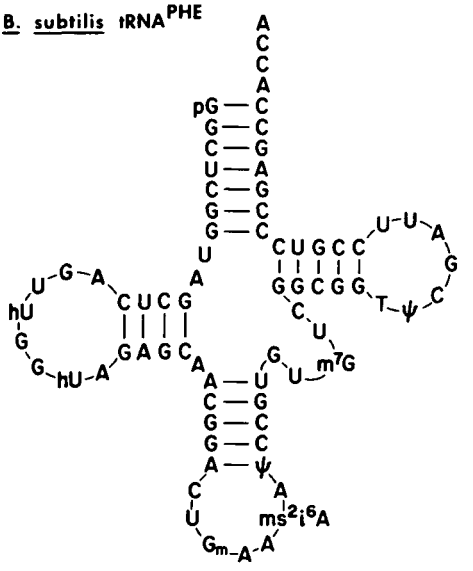

Figure 5
The nucleotide sequence of tRNA^{Phe} of B. subtilis arranged in the cloverleaf model.

relationship of the two grampositive organisms is thus confirmed by the structural similarities of at least one of their tRNAs. The seventeen G-C base pairs found in tRNA^{Phe} of B. subtilis, arranged in the cloverleaf, amounts to 80% of the entire base pairs and is thus similar in all three bacterial tRNA^{Phe} known so far. Phenylalanine tRNA of B. subtilis contains, in the fourth position, counted from the 3' end, an adenosine residue. The D-stem sequence is identical with all known tRNAs^{Phe}. The possibility to charge this tRNA with heterologous enzymes is in agreement with the hypothesis on the involvement of both these parts of the tRNA in the recognition by the

Table III

Further analysis of some T_1 and pancreatic RNase products.

Treatment	Fragments	Products
(1) alkaline phosphatase and partial venom phosphodiesterase	C-U-C-Gp C-U-C-A-Cp U-m^7G-U-C-C A-U-U-C-C-Gp	CUCG, CUC, CU, CUCAG, CUCA, CUC, CU, Um^7GUCG, Um^7GUC, Um^7GU, AUUCCG, AUUCC, AUUC,
(2) U_2RNase	A-C-U-Gm-A-A- ms^2i^6A-A-ψ-C-C-Gp C-C-A-C-C-A	CUGmAp(1), UCCGp(1), Ap(3) ms^2i^6A(1), CCAp(1), CCA_{OH}(1),
(3) micrococcal nuclease	Gm-A-A-ms^2i^6A-A- ψp	GmAp, Ams^2i^6Ap, Aψp,
(4) alkaline phosphatase followed by electrophoresis or DEAE-paper in 7% formic acid.	G-G-Up G-G-Tp G-G-Dp pG-G-Cp (one spot)	GGU GGD (one spot) GGT GGC

The sequences of products found after treatment (1) were determined by the M - values (13) and by T_2 RNase digestion of the products. Both spots found after treatment (4) were analyzed after venom phosphodiesterase digestion on 2-dimensional tlc.

tRNA ligase, as proposed for the yeast enzyme by Dudock et. al. (4) and Kern et. al. (5).

ACKNOWLEDCEMENTS

We are gratefully endebted to Dr. G. Dirheimer in whose laboratory most of this work was performed and Dr. H. Kersten for generous support and helpful discussion. We thank the EMBO for the award of a short term fellowship to H.H.A. We are also grateful to Mrs. C. Fix for skillful technical assistance. This work was supported by grants of the Deutsche Forschungsgemeinschaft (Ke 98/11), INSERM France (contract Nr. 7610613) and from the Fondation pour la Recherche medicale Francaise.

*To whom all correspondence should be addressed

Present address: Roche Institute of Molecular Biology, Nutley, NJ 07110, USA.

REFERENCES

1. Barrell, B.G. and Clark, B.F.C. (1974) in: Handbook of Nucleic Acid Sequences, (Johnsons-Bruvvers Ltd., Oxford), 1. ed.
2. Barrell, B.G. and Sanger, F. (1969) FEBS Letts. 3, 275-278.
3. Geurrier-Takada, C., Dirheimer, G., Grosjean, H. and Keith, G. (1975) FEBS Letts. 60, 286-289.
4. Dudock, B.S., di Peri, C., Scileppi, K. and Reszelbach, R. (1971) Proc. Natl. Acad. Sci. USA 68, 681-684.
5. Kern, D., Giege, R. and Ebel, J.P. (1972) Eur. J. Biochem. 31, 148-155.
6. Kersten, H. and Kersten, W. (1975) in: FEBS Proc. 9th Meeting, Budapest 1974 Vol. 34, Postsynthetic Modification of Macromolecules pp. 99-llo, North-Holland, Amsterdam.
7. Arnold, H.H., Schmidt, W., Raettig, R., Sandig, H., Domdey, H. and Kersten, H. (1976) Arch. Biochem. Biophys. 176, 12-20.

8. Arnold, H.H., Raettig, R. and Keith, G. (1977) FEBS Letts. 73, 210-214.
9. Arnold, H.H. and Ogilvie, A. (1977) Biochem. Biophys. Res. Commun. 74, 343-349.
10. Doi, R.H. (1971) in: Methods in Molecular Biology (Laskia, A.J. and Last, J.A., eds.(Vol. 1, pp. 67-88, Marcel Dekker, New York.
11. Raettig, R., Schmidt, W., Mahal, G., Kersten, H. and Arnold, H.H. (1976) Biochim. Biophys. Acta 435, 109-118.
12. Fradin, A., Gruhl, H. and Feldmann, H. (1975) FEBS Letts. 50, 185-189.
13. Barrell, B.G. (1971) in: Procedures in Nucleic Acid Research (Cantoni, G.L. and Davis, D.R., eds.) Harper and Row, New York Vol. 2, pp. 751-779.
14. Keith, G., Roy, A., Ebel, J.P. and Dirheimer, G. (1972) Biochimie 1405-1415.
15. Meza, L., Araya, A., Leon, G., Krauskopf, M., Siddiqui, M.A.Q. and Garel, J.P. (1977) FEBS Letts. in press.
16. Nishimura, S., Harada, F., Narushima, U. and Seno, T. (1967) Biochim. Biophys. Acta 142, 133-148.
17. Yamada, Y. and Ishikura, H. (1975) FEBS Letts. 54, 155-158.