Complete nucleotide sequence of the 23S rRNA gene of the Cyanobacterium, Anacystis nidulans

Susan E.Douglas and W.Ford Doolittle

Department of Biochemistry, Dalhousie University, Halifax, Nova Scotia, B3H 4H7 Canada

Received 30 November 1983: Revised and Accepted 2 March 1984

ABSTRACT

ABSTRACT The nucleotide sequence of the Anacystis nidulans 235 gene, including the 5'- and 3'-Flanking regions has determined. The gene is 2876 nucleotides long and shows primary sequence homology to the 23S rRNAs of plastids than to that of E. coli (79%). The predicted rRNA tran also shares many secondary structural features with th plastids, reinforcing the endosymbiont hypothesis for the of these organelles. 23S has rRNA been ws higher ds (84.5%) transcrip those origin

INTRODUCTION

The hypothesis that chloroplasts originated from endosymbiotic bacteria is widely accepted. Overwhelming evidence supporting this idea (1) has come from nucleic acid hybridisation studies of cyanobacterial and Euglena gracilis chloroplast rRNAs from ribosomal reconstitution experiments (2);(3) from ; T_1 -oligonucleotide cataloguing of 16S rRNAs of plastids and cyanobacteria (4,5,6) and from comparative primary sequence data of the rRNAs and/or their genes. Most of the latter form of evidence has involved comparisons of chloroplast and E. coli rRNAs. However, a contemporary organism more likely to be related to the ancestor of plastids is the cyanobacterium Anacystis nidulans.

In view of this, we have studied one of the two rRNA gene clusters of A. nidulans, and have shown it to be similar to those of plastids in the complement and order of genes, except that there is no physically distinct 4.5S rRNA (7). The spacer between the 16S and 23S rRNA genes contains genes for tRNAIle and tRNA^{Ala} (8), as do those of <u>Z.</u> mays (9), <u>N.</u> tabacum (10), and chloroplasts. The A. gracilis (11)nidulans tRNAs Ε. themselves show higher primary sequence homology to those of plastids than to those of E. coli. The 5S rRNA gene of A. <u>nidulans</u> also is more homologous to plastid 5S rRNA genes than to the <u>E.</u> <u>coli</u> 5S rRNA gene (7). Sequence data on the 16S rRNA gene of <u>A.</u> <u>nidulans</u> (12) shows the same relationship.

MATERIALS AND METHODS

Materials

Restriction enzymes were purchased from Bethesda Research Laboratories and New England Biolabs and used as suggested by the manufacturer. Phage M13 cloning vectors, primer and host cells were obtained from New England Biolabs. DNA polymerase I (large fragment) was from Boehringer Mannheim. $[\alpha-3^2p]$ -dATP was synthesized according to the method of Walseth and Johnson (13). DNA Preparation and Restriction Mapping

The recombinant plasmid containing the rRNA gene cluster of <u>A. nidulans</u> was propagated on <u>E. coli</u> JF1754 and the DNA purified as described previously (8). The restriction map for enzymes <u>Eco</u> RI, <u>Bam</u> HI, <u>Sal</u> I, <u>Pst</u> I, <u>Sma</u> I, <u>Xma</u> III and <u>Hind</u> III is shown in Figure 1 (a). The 23S rRNA gene lies on <u>Sma</u> I fragments of 1.4 and 2.3 kb, <u>Hind</u> III fragments of 0.48, 0.84 and 1.23 kb, and <u>Sma</u> I-<u>Hind</u> III fragments of 0.98 and 1.2 kb. These fragments were purified from low-melting-point-agarose and



Figure 1. (a). Restriction Map of the 6.3 kb Cloned rRNA Gene Cluster of A. <u>nidulans</u>. The 16S, 23S and 5S rRNA genes are indicated by heavy bars. (b). Strategy for Determining the Sequence of the 23S rRNA Gene and Flanking Regions. Arrows indicate the origin and direction of sequencing. subcloned into phage M13 using the restriction enzymes <u>Tag</u> I, <u>Sau</u> IIIa, <u>Hpa</u> II and <u>Rsa</u> I, as shown in Figure 1 (b).

DNA Sequencing

Sequencing was done by the dideoxynucleotide chain termination method using the phage M13 system (14). DNA sequencing gels were 33 x 40 x 0.3 cm and 8% in polyacrylamide (19:1, acrylamide:bis) and contained 8 <u>M</u> urea, 50 <u>mM</u> each of tris base and boric acid, 1 <u>mM</u> EDTA (pH 8.3). The gels were run at 1500 V and several loadings were performed to obtain overlapping sequences. Most regions were sequenced several times and in both orientations.

RESULTS AND DISCUSSION

The nucleotide sequence of the 23S rRNA gene and the 5'- and 3'-flanking regions is presented in Figure 2. The termini of the 23S rRNA gene were determined by comparison with the published sequence of the <u>E. coli</u> 23S rRNA (15). The gene is 2876 nucleotides long, slightly shorter than that of <u>E. coli</u> (2904 nucleotides) and longer than that of <u>N. tabacum</u> chloroplast (2804 nucleotides), excluding the 4.5S rRNA region. It is 54.2% G+C.

The 23S rRNAs of <u>A</u>. <u>nidulans</u> and plastids show several additional nucleotides relative to that of <u>E</u>. <u>coli</u>, as shown in Table 1. In <u>C</u>. <u>reinhardii</u> chloroplast, there is a 7S and 3S rRNA 5' to the 23S rRNA gene (16). These small rRNAs are homologous to the 5' terminus of other chloroplast 23S rRNAs and

Table 1. Insertions in 23S rRNAs of A. <u>nidulans</u> and Plastids Relative to E. <u>coli</u>. (A) <u>A. nidulans</u>, (Z) <u>Z. Mays</u>, (N) N. <u>tabacum</u>, (C) <u>C. reinhardli</u>. The co-ordinates refer to the location of the additional nucleotides in the <u>A. nidulans</u> sequence relative to the <u>E. coli</u> sequence.

Location	<u>A</u>	<u>Z</u>	N	<u>c</u>
270-278	9	-	-	9
292-313	22	25	26	25
1245-1249	5	5	5	?
1496-1501	6	7	7	?
1601-1605	5	6	6	?
1834-1838	5	5	5	?

			1	
	CTTTGCAAGCAGGATGTCAGCGGTTCGAGTCCGCTAACCTCCACCAAAAA	tRNAAla	A	COTAGGETGAAGCGOTAGCGGAAGCAGCCGTGGACGAAACCGA AAGCCTCCGAAGGTGTCTGTGAGGCATGCTGGAGGTATCAGAA
	GACTGCTTAAAAATTCAAAAAATCAGTTCAGCATCTTAAGTTTTCGATTTT			
	CGTATCGGAAGGCTTAGGAATGCCTGAGTTGAAAGACTCAGCAAGAAČČŤ		Ê	GTĢČGAATGĊTGĄĊATAAGTAAÇGATAAAGĊGĢGTGAAAAGĊÇĊGĊŢĊĠĊ 1297
	TGAAAACTGCATAGAGATGAGAGTGTAGGTATCACAGACACCATCTCATC		A	CGAAATCCCAAGGGTTCCTCCGGAAGGCTCGTCCGCGGGGGGGTTAGTCGG CGGAAGACCCAAGGGTTCCTGTCCAACGTTAATCGGGGCAGGGTGAGTCG
A	GOTTANGCANCENAGEGETTANCGTGGTGGTGCTACCTAGGCACACAGAGGCGAT	235	A	TCCTAAGECGAGCAGAAGTGCGTAGTCGACGACAACAGGTTAATATT
À	50 99 GAAGGACGTCGCTACCGA-CGATACGCCTCGGGAGCTCGAAGCAAGCAA		Ă	1397 COTOTACCOATTITICGATTOTGCAGA-OGGACOGGAGAAOGCTAGOCIAGO COTOTACCOATTITICACTACTICACAGACOGGAGAAOGCTAGOCIAGOCIAGOCIAGOCIAGOCIAGOCIAGOCIAGOCI
5			•	AGGATGTTGGTT-ACCTGTCCAAGTGTCCGAGGCGTT
Ë	ŢĂŦĂĂĊĊĞĠĊĢĂŦŦŦĊĊĠĂĂŢĠĠĠĠĂĂĂĊĊĊĂĠŦĠŦĠŦŦŦĊĠĂĊĂĊĊŖ 149 185		E	CGGGCGACGGTTGTCCCGGTTTAAGCGTGTAGGCTGGTTTTCCAGGCAAA 1496 1541
A E	ACCTGAATCCATAGGGTGGCCGCGACGAACCCGGCGAATTGAAAC TCATTAACTGAATCCATAGGGTTAATGAGG-CGAACCGGGGGAACTGAAAAC 198		B	GAGGAGCGGCGAAAACGCTCCCGAGCTGAGGCGTGATGCGACCCGCTACG TCCGTCAAAGCTGAGGCGTGATGACGAGGCGTTGATGACGAGGCGTTGA 1537
Å	АТСТТАБТАСССССАССАВСААСАСААААСААААСТСАТТСССТСАСТАССС АТСТААСТАСССССАССАССАССАССАСТАССССАСТАСССС 246		Å	GCGGGGAAAGTGGTTGATGTCAAGCTTCCAAGAAAAGC-TCTAAACA-1588 GTGCTGAAGCAACAAATGCCCTGCTTCCAGGAAAAGCCTCTAAGCATCAG 1587
A E	285 CGAGCGAACGGGGACCAGCCTAAACCAAACTCCACGGAGTTTGGGGTCGT CGAGCGAACGGGGAGCAGCCCCAGAGCCTGAATCAGTGT-GT 288		Å	TTAA-TCCAAAATTGCCTGTACCCTAAACCGACACAGGTGGGACGGTAGG TAACATCAAATCGTACCCCAAACCGACACAGGTGGGACGGTAGG 1632
A E	GGGACAGCAATGTGGACTGTGAATGTTAGACGAAGCAGCTGAAAA-CTGC GTGTTAGTGGAAGCGTCTGGAAAGCGCG 316		A E	GTATACCAAGGGGCGCGAGGTAACTCTCTCTAAGGAACTCGGCAAAATGA GAATACCAAGGGGCGCTGAGAGAACTCGGGTGAAGGAACTAGGCAAAATGG 1682
A E	ACCAGAGAAGGTGAAAGTCCTGTAGTCGAAAATTGAAACA-GCC-T-AGC GCGATACAGGGTGACAGCCCCCGTACACAAAAAT-G-CACATGCTGTGAGA 364		A B	TCCCTAACTTCCCCACAAAGCCACGCTGATATGTAACTGCCCCCCCC
A	TGAATCCCCAGTAGCACGGAGCACGTGAAATTCCGTGTGAATCCGCGAGG TCGATGAGTAGGGCGGGACACGTGGTATCCTGTCTGAATATGGGGG 411		Å	ACOTOGAÇCTGAAATCAÇTCGAACATGAAGAGGCCCAGGCCACTOTTAC
AB	АССА-ССТСЕТЛАЛССТАЛАТАСТССТЕТСТЕСССАТАСТСАЛССАСТА АССАТССТС-САЛЕВСТАЛАТАСТССТЕЛСТЕЛСССАТАСТСАЛССАСТА 460		Å	AAAAACACAGGTCTCCCCCTAAGTCGTAAGACGATGTATGGGGGCCTGACGC AAAAACACAGGACTGTCGCAAACTCGTAAGTCGGACGTATACGGGGCTGACGC 1832
Å	CCGCCAGGGAAAGGTGAAAAGAACCCCCGGAA-GGGGAAGTGAAATAGAACA CCGTCAGGGAAAGGCGAAAAGAACCCCCGGCAAGGGGAAGTGAAAAAAGAACA 510		A B	CTGCCCAGTGCCGGAAGGTTAAGGAAGCTGGTCAGCGCAAGTGAA CTGCCCGGTGCCGGAAGGTTAATTGATGGGGTTAGCGCAAGCGAA 1877
A E	TGAAACCGTGAGCTTACAAGCAGTCGGAGCCCCGATTCAACGGGTGACGGG TGAAACCGTGTACGTACAAGCAGTGGGAGCACGCTTAGGCGTGTGACTGG 560		Å	GCTGGCGACCGAAGCCCCGGTGAACGGCGGCCGTAACTATAACGGTCCTA GCTCTTGATCGAAGCCCCCGGTAAACGGCGGCCGTAACTATAACGGTCCTA 1927
A E	GTGCCTGTTGAAGAATGAGCCGGCGACTTATAGGCACTGGCA-GGTTAAG GTACCTTTTGTATAATGGGTCAGCGACTTATATTCTGTAGCAAGGTTAA- 609		A B	AGGTAGCGAAATTCCTTGTCGGGTAAGTTCCGACCCGCACGAAAGGCGTA AGGTAGCGAAATTCCTTGTCGGGTAAGTTCCGACCTGCACGAATGGCGTA 1977
A E	GCGGAAATGCCGAAGCCAAAGCGAAAGCGAGTCTGAATAGGGCGATAGT -CCGAATAGG-GGAGCCGAAGCGAAACCGAGTCTTAACTGGGCGTTAAGT 657		Ă	ACGATCTGGGCGCTGTCTCCAGAGAGAGGCTCGGCGAAATAGGAGTGTC ATGATGGCCAGGCTGTCTCCACCCGAGACTCAGTGAAATTGAACTGGC 2025
A E	CAGTGTTTATAGACCCGAACCCGGGTGATCTAACCATGGCCAGGATGAAG TGCAGGGTATAGACCCGAAACCCGGGTGATCTAACCATGGGCAGGTGAAG 707		Å	TOTGAAGATACCGGACTACCTGCACCCGGACAGAAAGACCCCTATGAACCTT TOTGAAGATGCAOTGTACCCCGCGGCAAGACCGCAAAGACCCCCGTGAACCT 2075
AE	CTTGGGTAACACCAAGTGGAGGTCCGAACCGACCGACGATGTTGAAAAATTAG GTTGGGTAACACTAACTGGAGGACCCGAACCGACCGACTAATGTTGAAAAATTAG 757	- - -	Ă	TACTOTAGCTTGGCTATTGG-CTTCGGCCTTTGACTGCGCAGGATAGCTGG TACTATAGCTTGACACTGAACACTTGAGCCTTGA-TGTGTAGGATAGCTGG 2124
A E	CCGATGACTGTGGGTTAGGGGTGAAATGCCAATCGAACCCGGAGCTAGCT		Å	CAGGCTATGAGACTTTCCTTGTGGGGGGAAGTGGAGCCAACGGTGAGTAG CAGGCTTTGAAGTGTGGACGCCAGTCTGCATGGAGCCGACCTTGAAGTAG 2174
Ă	GGTTCTCCCCCGAAATACGTTGAGGCGTAGCGGGTATG-G-ATT-ATAGCGC GGTTCTCCCCCGAAAGCTATTTAGGTAGCGCCTCGTGAATTCATCCATCA 855		Å	CACTCTTTCAAAGCTAGAAGT-CTAACTTTGAGCCCGTAATCCGGCCCAAG CACCCTTT-AAAGCTAGATGT-CTAACCTTGACCCGTAATCCGGCCCAAG
AE	TGGGGTAGAGCACTGATTCGGTGCGGGCTGCGA-GAGCGGTACCAACCG -GGGGTAGAGCACTGTTCGGCAAGGGGGTCATCCCGACTTACCAACCG 904		Ă	CAGTATCAGGTGGGCAGTTTGACTGGGGCGGTCGCCTCCTAAAAGGTA GACAGTGTCTGGTGGGTAGTTTGACTGGGGCGGTCTCCTCCTAAAAGGTA 2273
AE	AGTCAAACTCCGAATACGCCGTGTACACCATGCCAGTCAGACTGG ATGCAAACTGCGAATAC-CGGAGAATGTTATCACGGGAGACACACGGGGG 953		Å	ACGGAGGCGCGCAAAGGTTCCCTCAGGCTGGTCGGAAATCAGCCGACGAG ACGGAGGCGCGCCAAAGGTTGGCTAATCCTGGTCGGACATCAGGCGCGAG 2323
AE	GGGATAAGCTCCATGGTCAAGAGGGAAACAGCCCAGACCACCAGCTAAGG GTGCTAACGTCCGTCGTGAAGAGGGGAAACAACCCAGACCGCCAGCTAAGG 1003		Å	TGCAAAGGCATAAGGGAGCTTGACTGCAAGACCTACAAGTCGAGCAGGG TGCAATGGCATAAGCCAGCTTGACTGCAGGCGTGACGGCGCGGCAGCAGGTG 2373
AE	TCCTCANATCAGAACTTAGTGATAAAGGAGGTGGGAGTGCAATAGACAAC TCCCAAAGTCATGGTTAATGGGAAAGGATGTGGGAAGGCCCAGACGACGACGACGACGACGACGACGACGA		ě	CGAAAGTCGGCCTTAGTGATCCGACGGTTCTGAGTGGAAGGGCCCGCCC
A	AGGAGGTTTGCCTAGAAGCAGCCATCCTTAAAAGAAGTGCGTAATAGCTCA AGGATGTTGCCTTAGAAGCAGCCATCCTTTAAAAGAAAGCGTAATAGCTCA 1103		Ă	CAACGGATAAAAGTTACTCTAGGGATAACAGGCTGATCTCCTCCAAC450 CAACGGATAAAAGGTACTCCGGGGGATAACAGGCTGATACCGGCCCAAC407 2473
A	CTGGTCAAGCGCTCCTGCGCCGAAAATG-AACGGGGCTAAGTTCTGACC CTGGTCGAGTCGGCCTGCGCCGAAGATGTAACGGGGCTAAACCATGGACG 1153		Å	TCACATCGACGAGGAGGTTTGGCACCTCGATGTCGGCTCATCGCAACCTG TCATATCGACGGCGGGGGTGTTTGGCACCTCGATGTCGGCTCATCACAACCTG 2523
A	GAAGCTGTGGAATTGCTGTGCAATT-GGTAGGGGAGCGTTCGT GAAGCTGCGGCAGCGACCGTTATGCGTAGGGGAGCGTTCGTG 1203		A E	GCCCTCAACTCCCCAACCCTTTCCCCCATTAAACCCCTTCCCCCATTAAACCCCTCCCCACCCCCC

 a)
 STERNETTERSONTERRANCETCOTERRANCESTERSONTERRANCESTERSO



Figure 2. Nucleotide Sequence of the 23S rRNA Gene and Flanking Regions. The 3'-terminus of the tRNAAla, the 23S rRNA gene and the 5'-terminus of the 5S rRNA gene are boxed. The sequence of the E. coli 23S rRNA gene (E) is given below that of <u>A.</u> <u>nidulans (A)</u>.

<u>A. nidulans</u>, as shown in Figure 3. In fact, the insertion of 22 nucleotides in <u>A. nidulans</u> and plastids relative to <u>E. coli</u> (position 292-313) corresponds to the spacer between the 7S and 3S rRNA genes of <u>C. reinhardii</u>. On the other hand, the second spacer, that between the 3S and 23S rRNAs, is present in the <u>E. coli</u> 23S rRNA gene. The deletion of the 7S-3S rRNA gene spacer in <u>E. coli</u> 23S rRNA is therefore puzzling. The region cannot be of any functional importance since it is deleted in <u>E. coli</u>, post-transcriptionally removed in <u>C. reinhardii</u> chloroplast and forms variable secondary structures in <u>A. nidulans</u> and plastids (see Figure 4).

The 3S equivalent is very highly conserved in secondary structure, although primary sequence homology is quite variable (see Table 2). A. nidulans shows by far the highest homology to region of C. reinhardii chloroplast, and also shares a the 3S 9-nucleotide insertion in the 7S region relative to the other organisms (see Figure 3). This may reflect a closer evolutionary nidulans and the plastids of the relationship between A. reinhardii and Euglena as C. such unicellular green algae, nidulans and E. gracilis plastids share the gracilis. Both A. feature of uninterrupted spacer tRNAs (8,17). However, the 23S rRNA gene of C. reinhardii chloroplast is interrupted by a 940 bp intervening sequence near the 3' end (18). A. nidulans 23S rRNA contains no intervening sequence here, but neither do the Z. mays or N. tabacum plastid 23S rRNAs (19,20).

The 3' terminus of the 23S rRNA gene of A. nidulans does

T.A.C.GAAC	.G.TG.A	G.GCAATACAAGCGTCGTGCTGCTA	.GCA
	.u.u.icu.A		
AT.AATGGCAAGATGT/		CCACAACATAAAAATCCCTATTTTA	TAC
CT.A.C.AAACTCCGGAGTTT	G.G.TCG.A	C.GCAATGT GGACTGTGAATGTTA	.AC
T.A.C.GAAC	G.G.TG	G.GCAATACAAGCGTTGTGCTGCTA	.GC
GGGAGCAGCCCAGAGCCTGAATCA	GTG TGTGTGT	ТА	GTGGAAGC
	GGGAGCAGCCCAGAGCCTGAATCA T.A.C.GAAC CT.A.C.AAACTCCGGAGTTT	GGGAGCAGCCCAGAGCCTGAATCA GTG TGTGTGT T.A.C.GAAC G.G.TG. T.A.C.AAACTCCGGAGTTT G.G.TCG.A	GGGAGCAGCCCAGAGCCTGAATCA GTG TGTGTGTTA

	35	spacer
N.	CG.CCT.CTA.CCG.TCATAGC	GGCATTATATTCCC
C.	AGA. TCCTA.CG.TATAGT	GAAAATTTA.A.T.TT.T AATCCC
Α.	AGA CTA.C.G.G.AATTGT	.GTAACGAATCCC
z.	.GTGT.CA.CCG.TCT.TTAGC	.G.T.GCATTACTCCC
E.	GTCTGGAAAGGCGCGCGATACAGGGTGACAGCCCCGTACA	CAAAAA TGC ACATGCTGTGAG CTC GA TGAGTAG

Figure 3. Alignment of 75-3S rRNA Region of <u>Chlamydomonas</u> reinhardii Chloroplast With Homologous Region of 23S rRNAS, (E) E. <u>coli</u>, (Z) Z. <u>mays</u>, (A) <u>A. nidulans</u>, (C) <u>C. reinhardii</u>, (N) N. tabacum. This region corresponds to nucleotides 246-395 of the <u>A. nidulans</u> sequence.

not contain a sequence homologous to the 23S-4.5S spacer of \underline{Z} . mays and N. tabacum chloroplasts (7). The region where the 23S-4.5S spacer occurs in plastids is found near a helix which is altered in both plastids and <u>A.</u> <u>nidulans</u> relative to <u>E.</u> <u>coli</u>.



Figure 4. Potential Secondary Structures of the Regions Homologous to the 7S-3S rRNA Spacer of C. reinhardii. (A) Z. mays, (B) C. reinhardii, (C) A. nidulans, (D) N. tabacum.

Tab.	le	2.	Pei	cen	t :	Sequence	Homo	logy	of	the 3S	rRNA	Gene	of C.
rei	nhar	dii	and	its	Ho	moļogs.	(C) C	. îre	einł	nardii,	(E)	Ε.	colt,
(2)	Ζ.	may	<u>'s</u> ,	(A) [Α.	nidulans	5, (N)	<u>-N.</u> -	tal	bacum.	• •		

		E	A	<u>Z</u>	<u>c</u>	N
<u>E.</u>	<u>coli</u>		63	58	58	56
<u>A.</u>	nidulans		,	67	85	71
Z. Cħ1	mays oroplast				71	79
C. ch1	<u>reinhardii</u> oroplast					75

Other helices which are altered (or even missing) in A. nidulans and Z. mays chloroplast relative to E. coli occur at positions 131 (where the helix is absent), 257-312/365-389 (where the equivalent of the 7S-3S and 3S-23S spacers of C. reinhardii chloroplast are found) and 1721 (where the helix is shortened) of the A. nidulans sequence (see Figure 5). For the most part, the secondary structures of the rest of the 23S rRNA molecules of E. coli, A. nidulans and chloroplasts are superimposable. The potential helices are of three types: highly conserved (involving only Watson-Crick base pairs), irregular (containing G:U pairs) and highly irregular (containing non-Watson-Crick base pairs, especially G:A pairs and bulged residues). Some helices are highly conserved phylogenetically, while others are extremely variable but maintain secondary structural features.

the following calculations, each potential helix in the In nidulans 23S rRNA molecule was tabulated according to whether Α. it was highly conserved, variable or highly irregular with respect to the 23S rRNAs of E. coli and Z. mays chloroplast (data not shown). Of the highly conserved helices (greater than 85% homology among the three organisms), A. nidulans 23S rRNA shares approximately equal homology with those of E. coli and Z. mays chloroplast. Of the variable helices, 80% show more similarity to those of the Z. mays chloroplast 23S rRNA than to to those of the E. coli 23S rRNA. Of the highly irregular helices, 67% show more similarity to those of the Z. mays chloroplast 23S rRNA than to those of the E. coli 23S rRNA. Of: 13 positions where bulged residues occur in A. nidulans, the nine are common to all three organisms. Of the remaining four,





Figure 5. Potential Secondary Structure of the 23S rRNA of A. nidulans (adapted from the model of Noller et al [21]). The region between nucleotides 258 and 389 is highly variable in different organisms and does not conform to Noller's model. An alternative secondary structure is shown in brackets. one is shared with \underline{Z} . mays chloroplast, one with \underline{E} . <u>coli</u> and two are different in all 3 organisms. Those nucleotides found in single-stranded regions of the molecule tend to be highly conserved, although this is not as pronounced as in the case of the 16S rRNA (21).

Certain regions of the 23S rRNA molecule are postulated to interact with proteins, tRNAs or other rRNAs. Nucleotides 72-83 of the E. coli 5S rRNA are complementary to nucleotides 143-154 of the E. coli 23S rRNA (15) and an interaction between these two regions has been proposed. There is no complementarity between maize 23S and 5S rDNA in this region. In fact, part of this region has been deleted in the 23S rRNA genes of both A. nidulans and Z. mays chloroplast. A region complementary to nucleotides 68-74 of maize 5S rRNA has been found between nucleotides 1846 and 1855 of maize 23S rRNA (19). Comparable sequence homology between this region of A. nidulans 5S rRNA and nucleotides 1933-1938 is also present (see Figure 6). The region of the 23S rRNA molecule of E. coli corresponding to that delimited by nucleotides 2220-2347 in the A. nidulans 23S rRNA molecule, is also believed to interact with 5S rRNA and proteins L5, L18 and L25 (22). This region shows very similar secondary structure in A. nidulans and plastids and is also the region which is deleted in mammalian mitochondrial large subunit rRNA. Mammalian mitochondrial ribosomes lack a 5S rRNA.

Several interactions have also been postulated to occur coli and Z. between 16S and 23S rRNAs in both E. mays chloroplast (19). These interactions are supported by chemical modification studies and similar interactions could occur in A. Figure 6). Compensatory base changes have nidulans (see sometimes occurred in the A. nidulans sequence which maintain In addition, some nucleotide substitutions interactions. the strengthen these interactions. For example, an A to C change at position 2438 of the A. nidulans 23S rRNA molecule relative to that of E. coli allows additional base-pairing to occur with the 16S rRNA molecule.

The region corresponding to nucleotides 1962-1978 of <u>A.</u> <u>nidulans</u> 23S rRNA in <u>E. coli</u> has been postulated to interact with tRNAf^{Met}. Very strong base-pairing could occur in <u>E. coli</u>

55	68 CGGCAACG
235	GCC UUG 1938 1933
16S 23S	726 AGGGAUUAGAU ********* CCUCUAGUCGG 2441 2431
16S 23S	743 GUAGUCCUAG *** GACUAAAGGU 2292 2283
16S 23S	1344 GCCCGUCACACCAUG CUAGUCGG 2439 2432
tRNA 23S	⁵ GGGUAGAG CAG CCUGGUA AGAGACUCUGUCGCGGGGUC 1980 1962

Figure 6. Potential Interactions Between 23S rRNA and 16S rRNA, 5S rRNA and tRNAfMet in A. nidulans. Asterisks indicate nucleotides which could partake in base pairing.

(17 bp with a single-base bulge), but this is much weaker in both nidulans and Z. mays chloroplast (12 bp with 3 insertions Α. 1 mismatch, and 11 bp with 3 insertions and 2 mismatches, and respectively). Much stronger base-pairing could occur in A. nidulans between tRNA^{Met} and 23S rRNA between nucleotides 2426-2442. Strong base-pairing could also occur here between the tRNAf^{Met}s of both E. 23S rRNAs and coli and z. mays chloroplast. However, this is not substantiated by chemical modification cross-linking studies is therefore or and speculative.

<u>A.</u> <u>nidulans</u> and chloroplast 23S rRNAs can form very similar secondary structures in the regions believed to interact with proteins L24 and L1 in <u>E.</u> <u>coli</u> (22) as shown in Figure 7.

The region between nucleotides 2339 and 2562 is believed to be responsible for peptidyl transferase function in <u>E. coli</u> 23S rRNA. Very similar secondary structures can be formed by the large subunit rRNAs of <u>A. nidulans</u>, chloroplasts and yeast mitochondria. Chloramphenicol mutants map in this region in



Figure 7. Potential Secondary Structure of the Regions of the 23S rRNA of A. nidulans Proposed to Interact with Proteins Ll and L24 (22). Dots represent identical nucleotides to those found in <u>E. coli</u> 23S rRNA. Letters indicate nucleotide differences.

yeast mitochondrial 23S rRNA (23) and have been found to contain single-base substitutions at positions corresponding to nucleotides 2424 and 2479 of the <u>A. nidulans</u> sequence. These positions are conserved in <u>A. nidulans</u>, <u>E. coli</u> and <u>Z. mays</u> chloroplast 23S rRNAs and lie in long stretches of identical nucleotides.

The 23S rRNAs of some plastids contain "hidden breaks" which may result in the formation of the 4.5S rRNA of higher plants and 3S the and 7S rRNAs of C. reinhardii. Broad bean chloroplast 23S rRNA has been shown to be unstable at high temperatures and dissociate into three fragments (24). This may be a common to feature of chloroplast 23S rRNAs, and the sequences in these organelles and in A. nidulans corresponding to the 7S-3S and 3S-23S rRNA spacers of C. reinhardii may also be removed. It is interest that the 23S rRNA of A. nidulans also contains a of hidden break which results in its dissociation into two fragments Since the fragments formed are 0.17 and 0.88 daltons, the (25). break is likely to occur after the first domain or within domain V. The exact location is under investigation. Cyanobacteria are Table 3. Percent Sequence Homology of 23S rRNA Genes. The 23S rRNA genes from E. coli (E), A. nidulans (A), N. tabacum chloroplast (N) and Z. mays chloroplast (Z) were compared. The spacer between the 23S and 4.5S rRNAs of plastids was not included in the comparison.

		E	<u>A</u>	<u>N</u>	<u>_</u>
<u>E.</u>	<u>coli</u>		79	67	71
<u>A.</u>	nidulans			85	84
<u>N.</u>	tabacum				92

one of the few groups of prokaryotes known to contain hidden breaks in their 23S rRNA, another feature shared by chloroplast 23S rRNAs.

The 3'-flanking region of the 23S rRNA gene does not contain any typical eubacterial transcription stop or start signals (26) and transcription probably proceeds through the spacer and 5S rRNA gene, as is usual in eubacteria and some, if not all, plastids (27). Potential transcription termination signals have been located downstream from the 5S rRNA gene (7). The 5'- and regions of the 23S rRNA gene can partake in base 3'-flanking This occurs in E. pairing. coli and probably serves as a recognition signal for processing by RNase III (28). Base-pairing also occurs between the regions flanking the 23S rRNA (and 4.5S rRNA in those organisms which contain it) in chloroplasts (20).

In conclusion, from primary sequence studies (see Table 3), secondary structural features and physical characteristics of the large subunit rRNA of <u>A. nidulans</u>, it is clear that it shows more homology to chloroplast 23S rRNAs than to <u>E. coli</u> 23S rRNA, thus giving further support to the claim that plastids arose from cyanobacterium-like ancestors.

Acknowledgements

This work was supported by a grant from the Medical Research Council to W.F.D. and a post-graduate scholarship from the Natural Science and Engineering Research Council to S.E.D.

REFER	ENCES									
1.	Gray, M. Revs. 4	6. ^W i-4	and	Dooli	ittle	e, W	. F.	(198	2) Mic	robiol.
2.	Piggott,	Ğ. F	i. and	l Car	:r,	N.	G.	(1972)	Scienc	e 177,
3.	Gray, G.	99. ^E .	and He	erson,	, D.	Ε.	(1976) Arch	. Mic	robiol.
4.	Bonen, L	. jan	d Dool	ittle	د. ^w .	F.	(1975) Proc.	Natl.	Acad.
5.	Bonen, 669-673.	L.S.	and D	0011t	tle,	W.	F.	(1976)	Natur	e 261,
6.	Zablen,	L. B (19	75) ^{Kis}	sil, P roc.	M. § Nat	5., W	oese, cad.	C. R. Sci. U	and J. S.	Buetow, A. 72;
7.	Douglas,	s.	Ε.	and Do	ooli	ttle,	W. B	. (198	3) FEBS	Lett.,
8.	Williams	on,	ș. ₂ e	•a	nd Do	polit	tle, W	I. F.	(1983)	Nucleic
9.	Koch, W	i., E	dwards	-235. , K.	aı	nd K	ossel,	н.	(1981) (Cell 25,
10.	Takaiwa,	F	and Su	giura	, М.	(19	82) Nu	cleic A	Acids Re	es. 10,
11.	Graf, L	., к	ossel,	н.	and	d St	utz, E	. (198	30) Natu	re 186,
12.	Tomioka,	N.	and	Sugi	ura,	м.	(198	3) Mol.	Gen.	Genet.
13.	Walseth,	-363. T.	F	and	Johr	nson,	R.	A. ((1979) E	siochim.
14.	Messing,	Act J.	a. 56 Crea	$\frac{2}{R}$	-31. and	Seeb	urg, E	. н.	(1981)	Nucleic
15.	Acids Re Brosius,	s. 9 j.,	, 309- Dull,	321. T. (J. a	and N	oller.	н. г.	(1980) Proc.
16.	Natl. A Rochaix,	cad.	Sci. D. a	U. g nd Dai	S. A rlix,	А. 7 , J. ⁷	7, 201	-204 1982) J	. Mol.	Biol.
17.	Orozco,	 	м.	Jr.,	Rus	ņow,	K.	E., Dod	ld, J.	R. and
18.	Rochaix,	J.	$\begin{array}{c} B. & (1) \\ D. & an \end{array}$	980) . d Malı	J. E noe.	B101. P.	(1978)	Cell 1	10997-	11003.
19.	Edwards	, К.	and Ro	ssel,	н.	(198	1) Nuć	leic Ac	ids Re	es. 9,
20.	Takaiwa,	F .	and Su	giura	, М.	. (1982)	Eur.	J. B	iochem.
21.	Noller,	H. F	Kop	, J.,	Whea	aton,	V., E	rosius,	J.,	Gutell,
	Nucleic'	Acids	Res.	9, 6	167-6	5189.	е, г.	and ne	err, w.	(1981)
22.	Branlant (1976) E	, C.	J. Bi	l, A Ochem	. 7	Srwid	ada, J	. and	Brimaco	ombe, R.
23.	Dujon, B	, (<u>1</u>	980) ŢĈ	ell 20	ļ, 18	<u>35-19</u>	7.			
25.	Doolitf1	C. J	• ^E (13	/3) B: /1073	loçne	em. Bac	J. 13	5, 237-	·240	263
26:	Gilbert, Chamberi	W. in, I	(1976) M. E	in ds., 1	RNA 0193	Pol , Col	ymeras d Spri	ie, Los ng Hart	sick, Foor Labo	and oratory,
27.	Surzycki	, s.	J.	and	Roc	chaix	, J.	D. ((1971) J	. Mol.
28.	BIOL. 6 Bram, R.	2, 89 J.,	-109. Young	, R.	Α.	and	Steitz	, J. A	. (198	0) Cell
29.	19, 393- Kumano, 219-225.	401. M.,	Tomio	ka, 1	N. a	and S	ugiura	, M. ((1983) G	ene 24,

Note added in proof

While this manuscript was being reviewed, the nucleotide sequence of the 23S rRNA gene from the other rRNA gene cluster of <u>A. nidulans</u> was published (29). That sequence differs from that presented here in eight positions, mostlyin the 3'-terminal 600 nucleotides. These positions have been determined absolutely unambiguously by us by sequencing in both directions. This may reflect true differences between the two 23S rRNA gene copies.