JOURNAL OF BACTERIOLOGY, Aug. 1988, p. 3448–3458
0021-9193/88/083448-11$02.00/0

Vol. 170, No. 8

Genes Encoding the Alpha, Gamma, Delta, and Four F_0 Subunits of ATP Synthase Constitute an Operon in the Cyanobacterium *Anabaena* sp. Strain PCC 7120†

DAVID F. McCARN, RICHARD A. WHITAKER, JAWED ALAM,‡ JACQUELINE M. VRBA, AND STEPHANIE E. CURTIS*

Department of Genetics, Box 7614, North Carolina State University, Raleigh, North Carolina 27695-7614

Received 26 January 1988/Accepted 21 April 1988

A cluster of genes encoding subunits of ATP synthase of *Anabaena* sp. strain PCC 7120 was cloned, and the nucleotide sequences of the genes were determined. This cluster, denoted *atpI*, consists of four F_0 genes and three F_1 genes encoding the subunits *a* (*atpI*), *c* (*atpH*), *b'* (*atpG*), *b* (*atpF*), δ (*atpD*), α (*aptA*), and γ (*atpC*) in that order. Closely linked upstream of the ATP synthase subunit genes is an open reading frame denoted gene 1, which is equivalent to the *uncI* gene of *Escherichia coli*. The *atpI* gene cluster is at least 10 kilobase pairs distant in the genome from *apt2*, a cluster of genes encoding the β (*atpB*) and ε (*atpE*) subunits of the ATP synthase. This two-clustered ATP synthase gene arrangement is intermediate between those found in chloroplasts and *E. coli*. A unique feature of the *Anabaena atpI* cluster is overlap between the coding regions for *atpF* and *atpD*. The *atpI* cluster is transcribed as a single 7-kilobase polycistronic mRNA that initiates 140 base pairs upstream of gene 1. The deduced translation products for the *Anabaena* sp. strain PCC 7120 subunit genes are more similar to chloroplast ATP synthase subunits than to those of *E. coli*.

The proton-translocating ATP synthase is a multimeric membrane protein complex that couples a transmembrane gradient of electrochemical potential energy produced during electron transport to formation of ATP. This ubiquitous enzyme is found in cell membranes of bacteria, in inner mitochondrial membranes, and in thylakoid membranes of plant chloroplasts (reviewed in references 15, 18, and 30). In all examples studied the enzyme consists of two multimeric components: an extrinsic portion, F_1, composed of subunits denoted α, β, γ, δ, and ε, and in integral membrane portion F_0, composed of several subunits which vary depending on the source of the ATP synthase. In *Escherichia coli* and chloroplasts, the two systems used for comparison in this study, there are three (*a* to *c*) and four (I to IV) F_0 subunits, respectively.

In *E. coli*, genes encoding all eight subunits of the ATP synthase are tightly linked and cotranscribed (15). For the ATP synthase of chloroplasts, genes for some subunits are encoded in the nucleus and others are encoded in the organelle genome. The chloroplast ATP synthase genes of higher plants are organized into two separate transcriptional units: the β and ε genes are linked and cotranscribed (35), while a second cluster containing the I, III, IV, and α subunit genes map many kilobase pairs away in a second ATP synthase gene cluster (9, 19). The γ, δ, and subunit II genes are nuclear (34).

The cyanobacteria are procaryotes with an oxygen-evolving photosynthetic system nearly identical to that of plant chloroplasts. The similarity between cyanobacterial and plant photosystems, as well as the procaryotelike features of chloroplasts, lends support to the proposal (28) that plant chloroplasts may have evolved from close relatives of the cyanobacteria. The characterization of ATP synthase genes provides an opportunity to study the evolution of an important set of genes present in bacteria, cyanobacteria, and chloroplasts. In previous work the genes encoding the β (*atpB*) and ε (*atpE*) subunits of the ATP synthase in *Anabaena* sp. strain 7120 were shown to be linked and cotranscribed (10). In this paper the structure and arrangement of the genes encoding the α (*atpA*), γ (*atpC*), δ (*atpD*), *a* (*atpI*), *b* (*atpF*), *b'* (*atpG*), and *c* (*atpH*) subunits of the ATP synthase from *Anabaena* sp. strain 7120 are reported. These genes are tightly linked in the order *atpI-atpH-atpG-atpF-atpD-atpA-atpC* and form a single transcription unit. Preceding *atpI* and within the transcription unit is an eighth gene of unknown function denoted gene 1. The 7-kilobase (kb) polycistronic mRNA is initiated 140 base pairs (bp) from the start of the operon. This large ATP synthase gene cluster is at least 10 kbp from the *atpB-atpE* operon. A similar arrangement of two ATP synthase gene clusters has been reported for the cyanobacterium *Synechococcus* sp. strain PCC 6301 (8).

MATERIALS AND METHODS

Isolation of cyanobacterial, bacteriophage, and bacterial DNAs. *Anabaena* sp. strain PCC 7120 total cellular DNA was isolated from 15-liter cultures as described previously (11). Lambda bacteriophage and *E. coli* strains were grown and phage and plasmid DNA were isolated as described previously (38).

Isolation of genomic clones containing ATP synthase genes. A recombinant phage library of *Sau*3AI partial fragments of *Anabaena* sp. strain PCC 7120 total DNA cloned in the lambda vector λL47.1, obtained from James Golden, was screened by plaque hybridization (38) by using a cloned fragment of *Chlamydomonas reinhardii* chloroplast DNA. The plasmid p86, obtained from J. Boynton, contains a 3.6-kbp *Eco*RI fragment of *C. reinhardii* chloroplast DNA that carries the 5' end of the *atpA* gene. p86 was digested

* Corresponding author.

† Paper no. 11394 of the Journal Series of the North Carolina Agricultural Research Service.

‡ Present address: Department of Biochemistry and Molecular Biology, Louisiana State University Medical Center, New Orleans, LA 70112.

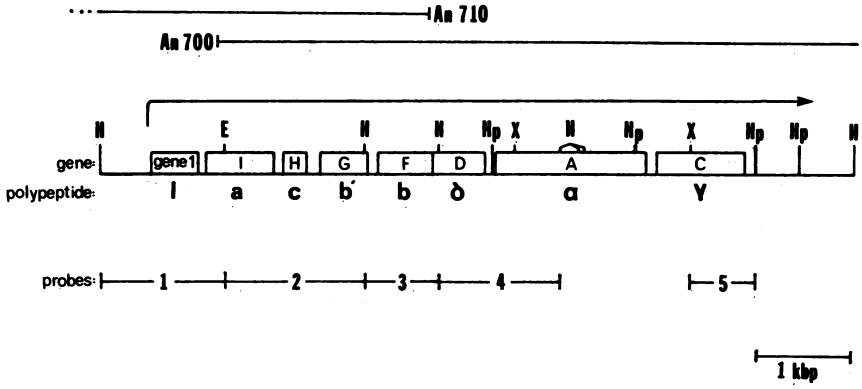


FIG. 1. Physical map of the region containing the *Anabaena* sp. strain PCC 7120 *atpI* gene cluster. The lines denoted An700 and An710 represent the inserts of *Anabaena* DNA from the recombinant clones λAn700 and λAn710, respectively. Symbol: ▭, gene coding regions within which the gene (*atp*) designations are given. Gene product designations are given below the genes. The *atpF* and *atpD* coding regions overlap. *atpI* is transcribed as a polycistronic mRNA that initiates 140 bp upstream of gene 1. The numbered bars indicate specific DNA fragments used in Northern and Southern blot hybridizations. Abbreviations for restriction endonuclease sites: X, *Xba*I; Hp, *Hpa*I; H, *Hin*dIII; E, *Eco*RI.

with *Eco*RI, and the plasmid insert was purified by electroelution from a 0.8% agarose gel. The insert was cloned into an SP6 vector (Promega Biotec, Madison, Wis.), and a ^{32}P-labeled ribonucleotide probe was prepared as specified by the manufacturer. Labeled probes were purified by passage through a Sephadex G-25 spin column (27) and hybridized with plaque lifts as previously described (38). Following hybridization, the filters were washed once in 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.5% sodium dodecyl sulfate at room temperature for 20 min and then a second time in 0.1× SSC–0.5% sodium dodecyl sulfate at 42°C for 2 h. A final 30-min wash was carried out under the same conditions as the second wash.

Positive clones from the screening with the *atpA* probe were rescreened with an *atpH* probe. The *atpH* probe, obtained from John Gray, contained a 140-bp *Hae*III-*Hin*dIII fragment internal to the pea chloroplast *atpH* gene cloned into M13mp9 (22). The double-stranded form of the *atpH* clone was digested with *Hin*dIII and *Bam*HI, and the insert was purified by electroelution from a 1.5% agarose gel. The fragment was cloned into an SP6 vector (Promega Biotec), and a ^{32}P-labeled ribonucleotide probe was prepared as described above. One clone, denoted λAn700, which hybridized with both the heterologous *atpH* and *atpA* probes, was selected from the lambda library of *Anabaena* sp. strain PCC 7120 DNA and characterized further.

Mapping and subcloning of *Anabaena* ATP synthase gene fragments. A restriction endonuclease map of λAn700 was determined by using *Hin*dIII, *Eco*RI, and *Xba*I. Southern blots containing various restriction endonuclease digests of λAn700 were hybridized with the heterologous *atpA* and *atpH* probes. The *Hin*dIII fragments homologous to these two probes and *Hin*dIII fragments which mapped between them in λAn700 were subcloned into the M13 vectors mp18 and mp19 (Fig. 1). An *Xba*I fragment which overlapped the two *aptA*-homologous *Hin*dIII fragments was also subcloned into the same M13 vectors (Fig. 1).

Initial DNA sequence determinations on the cloned *Hin*dIII fragments indicated that λAn700 did not contain the entire gene cluster. A probe from near the 5′ end of the λAn700 insert (Fig. 1, probe 3) was therefore used to screen the λL47.1 recombinant library for an overlapping clone. One clone, denoted λAn710, which was complementary to probe 3 but not to probe 4 (Fig. 1), was selected and mapped. A 1.3-kbp *Eco*RI-*Hin*dIII fragment of λAn710 containing the 5′ end of the gene cluster and upstream sequences was subcloned into M13mp18 and M13mp19.

Analysis of ATP synthase gene copy number. The number of copies of the ATP synthase genes in the *Anabaena* sp. strain PCC 7120 genome was determined by hybridization of ATP synthase gene probes to Southern blots of *Anabaena* cellular DNA digested with *Hin*dIII or *Eco*RI. Probes 1 through 4 consisted of restriction fragments produced by digestion of λAn700 DNA with *Hin*dIII and *Eco*RI. Probe 5 was produced by digesting λAn700 with *Xba*I and *Hpa*I. The fragments were purified by electrophoresis from agarose gels, ^{32}P labeled by nick translation, and hybridized with Southern blots as described previously (38).

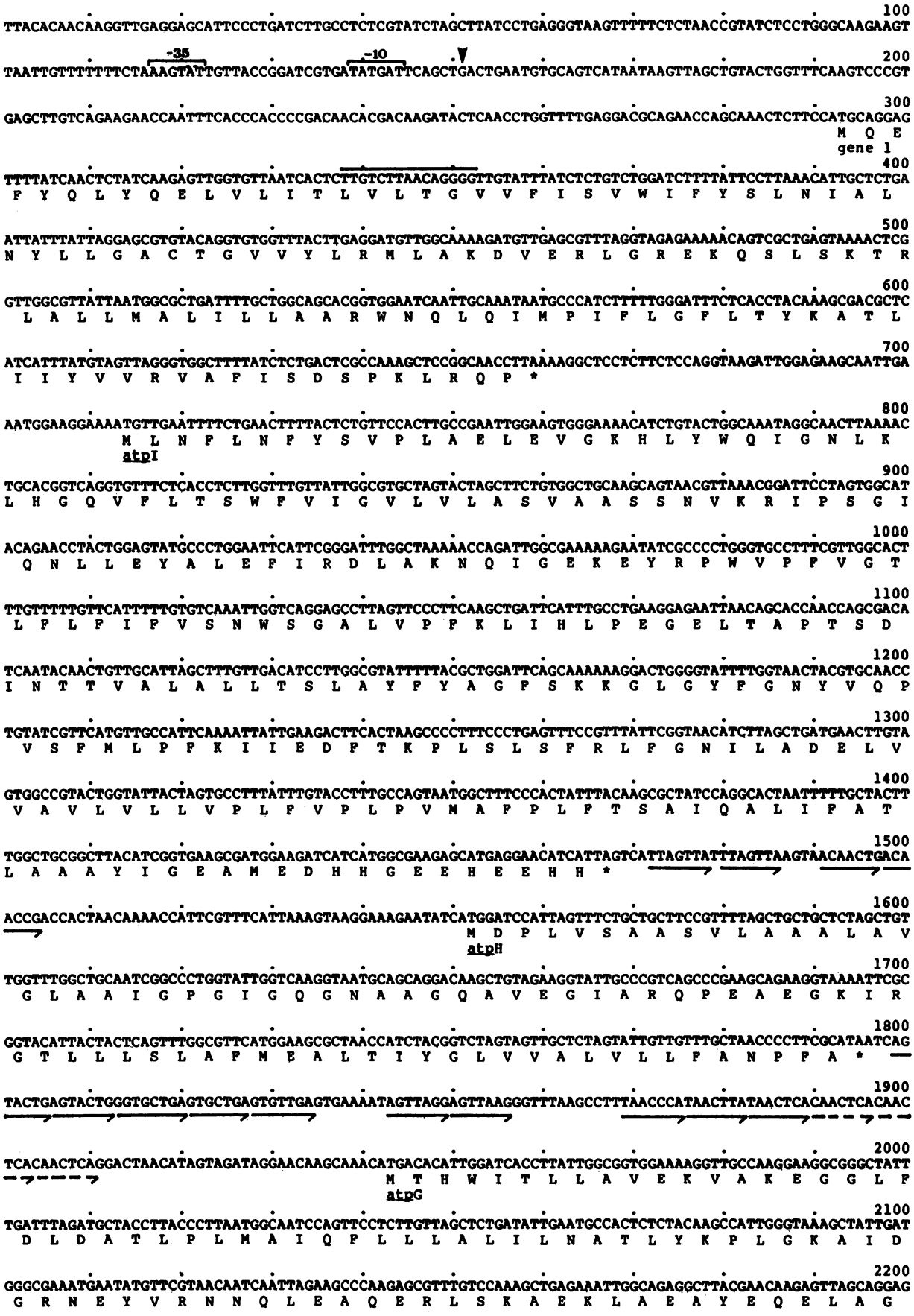
DNA sequence determinations. The sequences of the *Hin*dIII and *Xba*I fragments subcloned in M13 were determined by using the chain termination method (39) and site-directed oligonucleotide primers. The sequences of both strands were determined. Sequence analyses were performed by using the IBI DNA/Protein Sequence Analysis System computer software (International Biotechnologies, Inc., New Haven, Conn.).

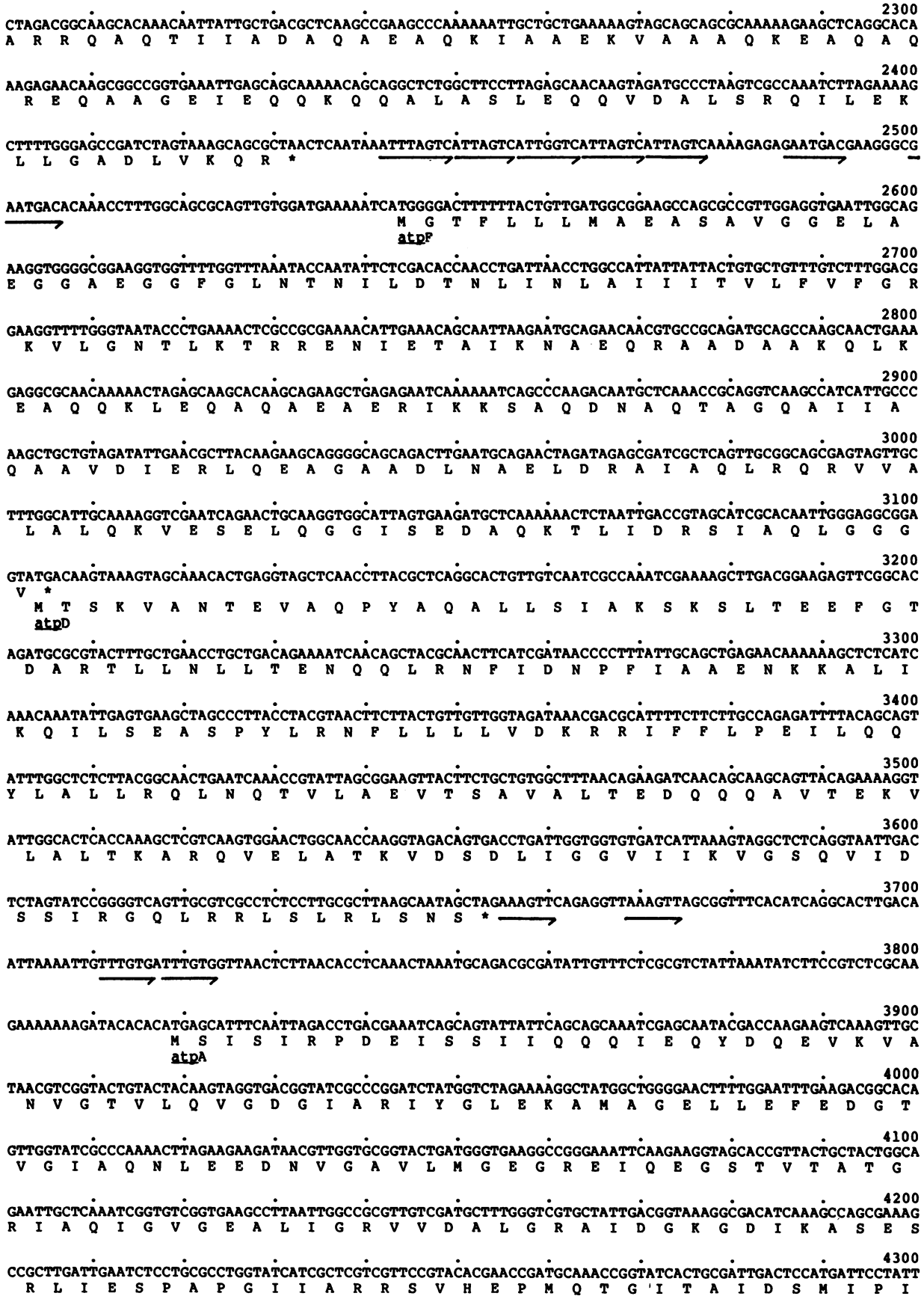
RNA analysis. Total RNA was isolated as described previously (44). RNA was denatured with formaldehyde and formamide, fractionated on 1% agarose gels containing 2.2 M formaldehyde, and then transferred to nitrocellulose as described previously (27). Size markers consisting of DNA fragments of known size (1-kb ladder and lambda DNA digested with *Hin*dIII [Bethesda Research Laboratories, Inc., Gaithersburg, Md.]) were similarly denatured and fractionated. Northern (RNA) blots were hybridized with nick-translated probes 1 through 5 (Fig. 1) and washed as described for Southern blots.

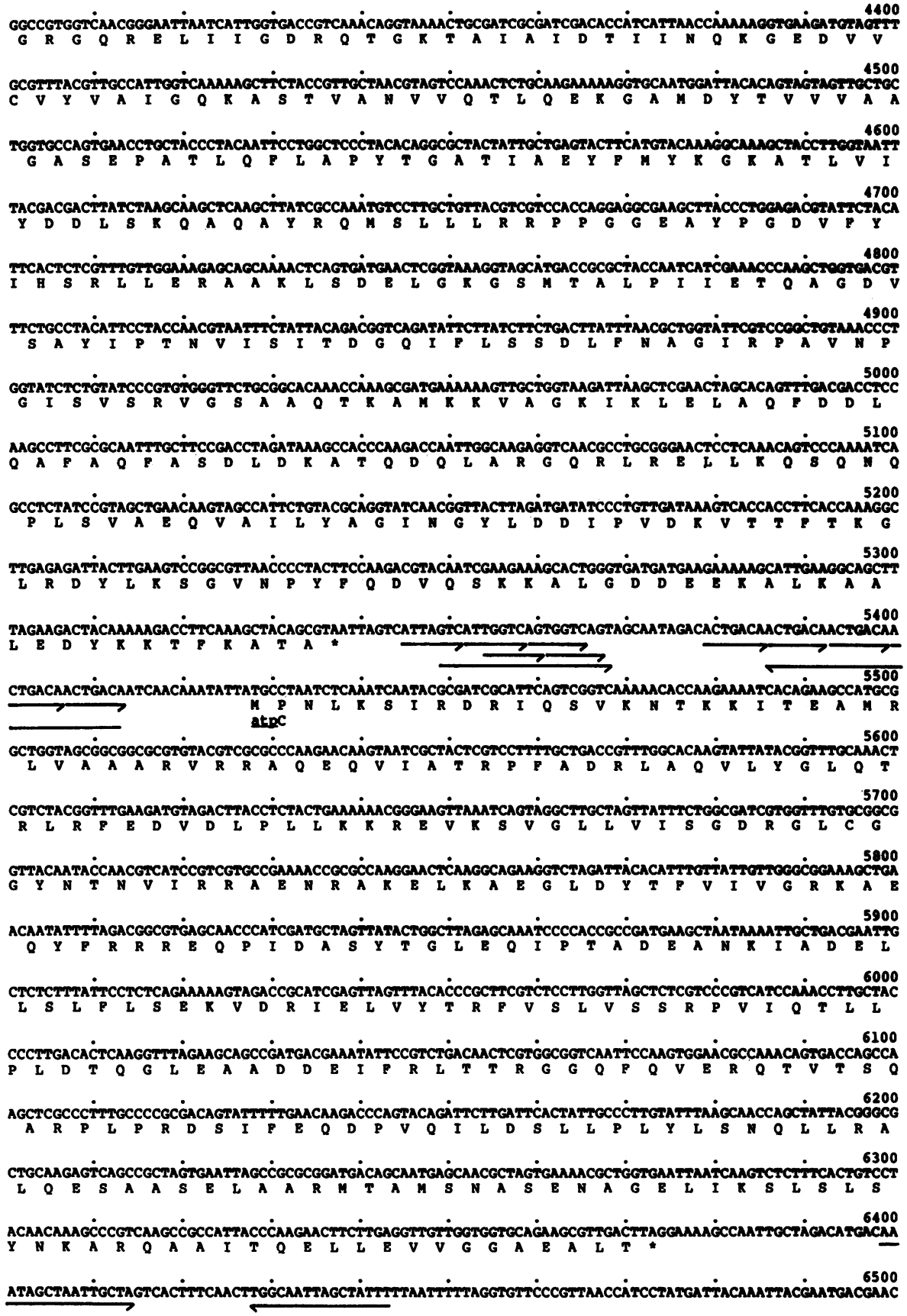
Oligonucleotide synthesis. All oligonucleotides were synthesized in an automated DNA synthesizer (model 380A; Applied Biosystems, Inc., Foster City, Calif.). Oligonucleotides were prepared for DNA sequencing or primer extension as described previously (10).

Primer extension and S1 nuclease assays. Primer extension assays were performed as described previously (10) by using an oligonucleotide primer of 15 bases specific to the gene 1 coding region (Fig. 2).

A probe for the S1 nuclease protection assay was prepared by cloning the 1.3-kbp *Hin*dIII-*Eco*RI fragment (Fig. 1,


-35
-10
gene 1
atpI
atpH
atpG


2300
CTAGACGGCAAGCACAAACAATTATTGCTGACGCTCAAGCCGAAGCCCAAAAAATTGCTGCTGAAAAAGTAGCAGCAGCGCAAAAAGAAGCTCAGGCACA
A R R Q A Q T I I A D A Q A E A Q K I A A E K V A A A Q K E A Q A Q
2400
AAGAGAACAAGCGGCCGGTGAAATTGAGCAGCAAAAACAGCAGGCTCTGGCTTCCTTAGAGCAACAAGTAGATGCCCTAAGTCGCCAAATCTTAGAAAAG
R E Q A A G E I E Q Q K Q Q A L A S L E Q Q V D A L S R Q I L E K
2500
CTTTTGGGAGCCGATCTAGTAAAGCAGCGCTAACTCAATAAATTTAGTCATTAGTCATTGGTCATTAGTCATTAGTCAAAAGAGAGAATGACGAAGGGCG
L L G A D L V K Q R *
2600
AATGACACAAACCTTTGGCAGCGCAGTTGTGGATGAAAAATCATGGGGACTTTTTTACTGTTGATGGCGGAAGCCAGCGCCGTTGGAGGTGAATTGGCAG
M G T F L L L M A E A S A V G G E L A
atpF
2700
AAGGTGGGGCGGAAGGTGGTTTTGGTTTAAATACCAATATTCTCGACACCAACCTGATTAACCTGGCCATTATTATTACTGTGCTGTTTGTCTTTGGACG
E G G A E G G F G L N T N I L D T N L I N L A I I I T V L F V F G R
2800
GAAGGTTTTGGGTAATACCCTGAAAACTCGCCGCGAAAACATTGAAACAGCAATTAAGAATGCAGAACAACGTGCCGCAGATGCAGCCAAGCAACTGAAA
K V L G N T L K T R R E N I E T A I K N A E Q R A A D A A K Q L K
2900
GAGGCGCAACAAAAACTAGAGCAAGCACAAGCAGAAGCTGAGAGAATCAAAAAATCAGCCCAAGACAATGCTCAAACCGCAGGTCAAGCCATCATTGCCC
E A Q Q K L E Q A Q A E A E R I K K S A Q D N A Q T A G Q A I I A
3000
AAGCTGCTGTAGATATTGAACGCTTACAAGAAGCAGGGGCAGCAGACTTGAATGCAGAACTAGATAGAGCGATCGCTCAGTTGCGGCAGCGAGTAGTTGC
Q A A V D I E R L Q E A G A A D L N A E L D R A I A Q L R Q R V V A
3100
TTTGGCATTGCAAAAGGTCGAATCAGAACTGCAAGGTGGCATTAGTGAAGATGCTCAAAAAACTCTAATTGACCGTAGCATCGCACAATTGGGAGGCGGA
L A L Q K V E S E L Q G G I S E D A Q K T L I D R S I A Q L G G G
3200
GTATGACAAGTAAAGTAGCAAACACTGAGGTAGCTCAACCTTACGCTCAGGCACTGTTGTCAATCGCCAAATCGAAAAGCTTGACGGAAGAGTTCGGCAC
V *
M T S K V A N T E V A Q P Y A Q A L L S I A K S K S L T E E F G T
atpD
3300
AGATGCGCGTACTTTGCTGAACCTGCTGACAGAAAATCAACAGCTACGCAACTTCATCGATAACCCCTTTATTGCAGCTGAGAACAAAAAAGCTCTCATC
D A R T L L N L L T E N Q Q L R N F I D N P F I A A E N K K A L I
3400
AAACAAATATTGAGTGAAGCTAGCCCTTACCTACGTAACTTCTTACTGTTGTTGGTAGATAAACGACGCATTTTCTTCTTGCCAGAGATTTTACAGCAGT
K Q I L S E A S P Y L R N F L L L L V D K R R I F F L P E I L Q Q
3500
ATTTGGCTCTCTTACGGCAACTGAATCAAACCGTATTAGCGGAAGTTACTTCTGCTGTGGCTTTAACAGAAGATCAACAGCAAGCAGTTACAGAAAAGGT
Y L A L L R Q L N Q T V L A E V T S A V A L T E D Q Q Q A V T E K V
3600
ATTGGCACTCACCAAAGCTCGTCAAGTGGAACTGGCAACCAAGGTAGACAGTGACCTGATTGGTGGTGTGATCATTAAAGTAGGCTCTCAGGTAATTGAC
L A L T K A R Q V E L A T K V D S D L I G G V I I K V G S Q V I D
3700
TCTAGTATCCGGGGTCAGTTGCGTCGCCTCTCCTTGCGCTTAAGCAATAGCTAGAAAGTTCAGAGGTTAAAGTTAGCGGTTTCACATCAGGCACTTGACA
S S I R G Q L R R L S L R L S N S *
3800
ATTAAAATTGTTTGTGATTTGTGGTTAACTCTTAACACCTCAAACTAAATGCAGACGCGATATTGTTTCTCGCGTCTATTAAATATCTTCCGTCTCGCAA
3900
GAAAAAAAGATACACACATGAGCATTTCAATTAGACCTGACGAAATCAGCAGTATTATTCAGCAGCAAATCGAGCAATACGACCAAGAAGTCAAAGTTGC
M S I S I R P D E I S S I I Q Q Q I E Q Y D Q E V K V A
atpA
4000
TAACGTCGGTACTGTACTACAAGTAGGTGACGGTATCGCCCGGATCTATGGTCTAGAAAAGGCTATGGCTGGGGAACTTTTGGAATTTGAAGACGGCACA
N V G T V L Q V G D G I A R I Y G L E K A M A G E L L E F E D G T
4100
GTTGGTATCGCCCAAAACTTAGAAGAAGATAACGTTGGTGCGGTACTGATGGGTGAAGGCCGGGAAATTCAAGAAGGTAGCACCGTTACTGCTACTGGCA
V G I A Q N L E E D N V G A V L M G E G R E I Q E G S T V T A T G
4200
GAATTGCTCAAATCGGTGTCGGTGAAGCCTTAATTGGCCGCGTTGTCGATGCTTTGGGTCGTGCTATTGACGGTAAAGGCGACATCAAAGCCAGCGAAAG
R I A Q I G V G E A L I G R V V D A L G R A I D G K G D I K A S E S
4300
CCGCTTGATTGAATCTCCTGCGCCTGGTATCATCGCTCGTCGTTCCGTACACGAACCGATGCAAACCGGTATCACTGCGATTGACTCCATGATTCCTATT
R L I E S P A P G I I A R R S V H E P M Q T G I T A I D S M I P I


4400
GGCCGTGGTCAACGGGAATTAATCATTGGTGACCGTCAAACAGGTAAAACTGCGATCGCGATCGACACCATCATTAACCAAAAAGGTGAAGATGTAGTTT
G R G Q R E L I I G D R Q T G K T A I A I D T I I N Q K G E D V V
4500
GCGTTTACGTTGCCATTGGTCAAAAAGCTTCTACCGTTGCTAACGTAGTCCAAACTCTGCAAGAAAAAGGTGCAATGGATTACACAGTAGTAGTTGCTGC
C V Y V A I G Q K A S T V A N V V Q T L Q E K G A M D Y T V V V A A
4600
TGGTGCCAGTGAACCTGCTACCCTACAATTCCTGGCTCCCTACACAGGCGCTACTATTGCTGAGTACTTCATGTACAAAGGCAAAGCTACCTTGGTAATT
G A S E P A T L Q F L A P Y T G A T I A E Y F M Y K G K A T L V I
4700
TACGACGACTTATCTAAGCAAGCTCAAGCTTATCGCCAAATGTCCTTGCTGTTACGTCGTCCACCAGGAGGCGAAGCTTACCCTGGAGACGTATTCTACA
Y D D L S K Q A Q A Y R Q M S L L L R R P P G G E A Y P G D V F Y
4800
TTCACTCTCGTTTGTTGGAAAGAGCAGCAAAACTCAGTGATGAACTCGGTAAAGGTAGCATGACCGCGCTACCAATCATCGAAACCCAAGCTGGTGACGT
I H S R L L E R A A K L S D E L G K G S M T A L P I I E T Q A G D V
4900
TTCTGCCTACATTCCTACCAACGTAATTTCTATTACAGACGGTCAGATATTCTTATCTTCTGACTTATTTAACGCTGGTATTCGTCCGGCTGTAAACCCT
S A Y I P T N V I S I T D G Q I F L S S D L F N A G I R P A V N P
5000
GGTATCTCTGTATCCCGTGTGGGTTCTGCGGCACAAACCAAAGCGATGAAAAAAGTTGCTGGTAAGATTAAGCTCGAACTAGCACAGTTTGACGACCTCC
G I S V S R V G S A A Q T K A M K K V A G K I K L E L A Q F D D L
5100
AAGCCTTCGCGCAATTTGCTTCCGACCTAGATAAAGCCACCCAAGACCAATTGGCAAGAGGTCAACGCCTGCGGGAACTCCTCAAACAGTCCCAAAATCA
Q A F A Q F A S D L D K A T Q D Q L A R G Q R L R E L L K Q S Q N Q
5200
GCCTCTATCCGTAGCTGAACAAGTAGCCATTCTGTACGCAGGTATCAACGGTTACTTAGATGATATCCCTGTTGATAAAGTCACCACCTTCACCAAAGGC
P L S V A E Q V A I L Y A G I N G Y L D D I P V D K V T T F T K G
5300
TTGAGAGATTACTTGAAGTCCGGCGTTAACCCCTACTTCCAAGACGTACAATCGAAGAAAGCACTGGGTGATGATGAAGAAAAAGCATTGAAGGCAGCTT
L R D Y L K S G V N P Y F Q D V Q S K K A L G D D E E K A L K A A
5400
TAGAAGACTACAAAAAGACCTTCAAAGCTACAGCGTAATTAGTCATTAGTCATTGGTCAGTGGTCAGTAGCAATAGACACTGACAACTGACAACTGACAA
L E D Y K K T F K A T A *
5500
CTGACAACTGACAATCAACAAATATTATGCCTAATCTCAAATCAATACGCGATCGCATTCAGTCGGTCAAAAACACCAAGAAAATCACAGAAGCCATGCG
M P N L K S I R D R I Q S V K N T K K I T E A M R
atpC
5600
GCTGGTAGCGGCGGCGCGTGTACGTCGCGCCCAAGAACAAGTAATCGCTACTCGTCCTTTTGCTGACCGTTTGGCACAAGTATTATACGGTTTGCAAACT
L V A A A R V R R A Q E Q V I A T R P F A D R L A Q V L Y G L Q T
5700
CGTCTACGGTTTGAAGATGTAGACTTACCTCTACTGAAAAAACGGGAAGTTAAATCAGTAGGCTTGCTAGTTATTTCTGGCGATCGTGGTTTGTGCGGCG
R L R P E D V D L P L L K K R E V K S V G L L V I S G D R G L C G
5800
GTTACAATACCAACGTCATCCGTCGTGCCGAAAACCGCGCCAAGGAACTCAAGGCAGAAGGTCTAGATTACACATTTGTTATTGTTGGGCGGAAAGCTGA
G Y N T N V I R R A E N R A K E L K A E G L D Y T F V I V G R K A E
5900
ACAATATTTTAGACGGCGTGAGCAACCCATCGATGCTAGTTATACTGGCTTAGAGCAAATCCCCACCGCCGATGAAGCTAATAAAATTGCTGACGAATTG
Q Y F R R R E Q P I D A S Y T G L E Q I P T A D E A N K I A D E L
6000
CTCTCTTTATTCCTCTCAGAAAAAGTAGACCGCATCGAGTTAGTTTACACCCGCTTCGTCTCCTTGGTTAGCTCTCGTCCCGTCATCCAAACCTTGCTAC
L S L F L S E K V D R I E L V Y T R F V S L V S S R P V I Q T L L
6100
CCCTTGACACTCAAGGTTTAGAAGCAGCCGATGACGAAATATTCCGTCTGACAACTCGTGGCGGTCAATTCCAAGTGGAACGCCAAACAGTGACCAGCCA
P L D T Q G L E A A D D E I F R L T T R G G Q F Q V E R Q T V T S Q
6200
AGCTCGCCCTTTGCCCCGCGACAGTATTTTTGAACAAGACCCAGTACAGATTCTTGATTCACTATTGCCCTTGTATTTAAGCAACCAGCTATTACGGGCG
A R P L P R D S I F E Q D P V Q I L D S L L P L Y L S N Q L L R A
6300
CTGCAAGAGTCAGCCGCTAGTGAATTAGCCGCGCGGATGACAGCAATGAGCAACGCTAGTGAAAACGCTGGTGAATTAATCAAGTCTCTTTCACTGTCCT
L Q E S A A S E L A A R M T A M S N A S E N A G E L I K S L S L S
6400
ACAACAAAGCCCGTCAAGCCGCCATTACCCAAGAACTTCTTGAGGTTGTTGGTGGTGCAGAAGCGTTGACTTAGGAAAAGCCAATTGCTAGACATGACAA
Y N K A R Q A A I T Q E L L E V V G G A E A L T *
6500
ATAGCTAATTGCTAGTCACTTTCAACTTGGCAATTAGCTATTTTTAATTTTTAGGTGTTCCCGTTAACCATCCTATGATTACAAATTACGAATGACGAAC

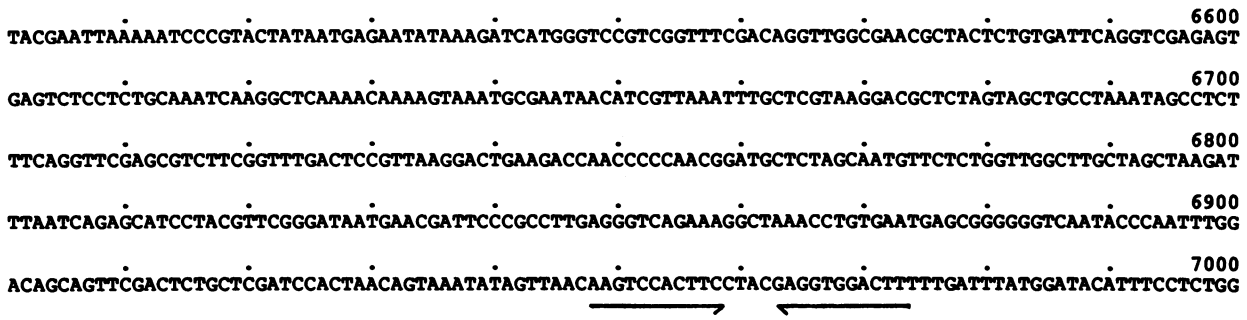


FIG. 2. Nucleotide sequence of the noncoding strand of the *apt1* gene cluster and flanking regions. Translation products are given below the nucleotide sequence. The complement to the sequence of the primer used in the primer extension assay is overlined. The arrowhead indicates the 5′ end of the mRNA transcript as determined by primer extension. Clusters of direct repeat sequences are underlined by arrows. Convergent arrows mark inverted repeat sequences capable of forming stable stem-loop structures in mRNA.

probe 1) of λAn710 into the vector pBS M13$^+$ (Stratagene Cloning Systems, San Diego, Calif.). A full-length ^{32}P-labeled ribonucleotide probe complementary to the mRNA was prepared as specified by the manufacturer. S1 nuclease assays were performed as described previously (27) with the ribonucleotide probe and total RNA.

RESULTS

Cloning of the ATP synthase genes. A lambda recombinant library of total *Anabaena* sp. strain PCC 7120 DNA was screened with a probe containing a portion of the gene for the α subunit (*atpA*) of the ATP synthase from *C. reinhardtii* chloroplasts. Positive clones were rescreened with a *c* subunit gene (*atpH*) from pea chloroplasts. A clone denoted λAn700, which hybridized with both the *atpA* and *atpH* probes, was selected and characterized by restriction endonuclease mapping (Fig. 1).

Hybridization of the *atpA* and *atpH* probes to restriction endonuclease digests of λAn700 DNA indicated *Hin*dIII fragments of 1.3 and 2.8 kbp with homology to *atpA* and a 3.6-kbp *Hin*dIII fragment with homology to *atpH*. The *atpH*-homologous fragment contained one end of the *Anabaena* DNA insert in λAn700 and approximately 2.0 kbp of vector DNA. The three identified fragments and the *Hin*dIII fragments which mapped between them in λAn700 were subcloned into M13 vectors. An *Xba*I fragment which overlapped the two *atpA* homologous *Hin*dIII fragments was also subcloned into M13.

DNA sequence analysis of the ATP synthase genes. The sequences of the subcloned *Hin*dIII and *Xba*I fragments of λAn700 were determined. Translation of the DNA sequence revealed that one end of the insert in λAn700 fell within an open reading frame (ORF) which was followed by six closely linked ORFs on the same DNA strand (Fig. 2). To obtain the sequence of the incomplete ORF and upstream sequences, we isolated an overlapping clone, denoted λAn710, from the lambda recombinant library (Fig. 1). A 1.3-kbp *Hin*dIII-*Eco*RI fragment of λAn710 containing the incomplete ORF was subcloned, and its sequence was determined. This fragment contained the 5′, end of the partial ORF from λAn700, which was preceded by an additional closely linked ORF (Fig. 1). Thus, the entire gene cluster contained eight ORFs and was denoted *atp1* in accordance with the nomenclature proposed by Cozens and Walker (8).

There are no other ORFs on either strand of the DNA in close proximity to the *atp1* cluster. The 1.3 kbp of sequence characterized downstream from *atpC* does not contain any ORFs, nor does the 700 bp of sequence upstream from gene 1.

Identification of F_1 subunit genes. Cyanobacterial ATP synthases are similar in size to those of *E. coli* and chloroplasts (20, 26), and the F_1 component is composed of five subunits (α, β, γ, δ, and ε), which are similar in size to those of other organisms. To identify the ATP synthase F_1 subunit genes in *atp1*, we compared the ORFs with the deduced translation products of F_1 subunit genes from other organisms. The α subunit gene (*atpA*) was identified on the basis of the high similarity of the deduced translation product to the tobacco chloroplast and *E. coli* α subunits (Table 1). The *atpA* gene encodes a protein of M_r 57,181 or 54,301, depending on which of two in-frame methionines is used to initiate translation.

The ORFs upstream and downstream of *atpA* showed homology to the *E. coli* δ and γ subunits, respectively (Fig. 1; Table 1). Their placement relative to *atpA*, which was the same as in the *E. coli* ATP synthase operon, further supported the identification of these as δ (*atpD*) and γ (*atpC*) subunit genes. The *atpC* and *atpD* genes encode proteins of M_r 35,213 and 20,345, respectively, in good agreement with the molecular weights determined by denaturing sodium dodecyl sulfate-gel electrophoresis for the δ and γ subunits from the cyanobacterium *Spirulina platensis* (20).

TABLE 1. Similarity with *Anabaena* proteins[a]

ATP synthase subunit	% Similarity with proteins from:		
	Synechococcus sp.[b]	Chloroplasts[c]	*E. coli*[d]
α	78	69	55
β	87	81	68
γ	80	N[e]	35
δ	49	N	26
ε	70	41	33
a (IV)	80	68	17
b (I)	41	26	26
b′ (II)	58	N	NE[f]
c (III)	96	89	33

[a] *Anaebaena* sp. strain PCC 7120; percent identical amino acids after alignment for maximum similarity.
[b] *Synechococcus* sp. strain PCC 6301 (8).
[c] Amino acid sequences for all subunits are from spinach chloroplasts (19, 47), except for the α subunit, which is from wheat chloroplasts (23).
[d] Amino acid sequences from references 16, 17, and 40.
[e] N, Nucleus encoded in plants; complete amino acid sequences not available.
[f] NE, No equivalent subunit.

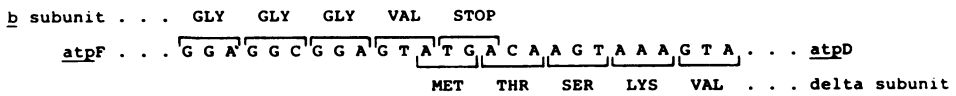


FIG. 3. Nucleotide and amino acid sequence showing the overlap between the *atpF* and *atpD* genes.

Identification of F_0 subunit genes. Little information is available on the amino acid sequences of cyanobacterial F_0 subunits or the subunit composition of the F_0 component. Upstream from the three F_1 genes in *atpI* are four ORFs (Fig. 1) with various degrees of similarity to F_0 subunits from different sources. The third ORF in *atpI* is very similar to subunits *c* and III from *E. coli* and chloroplasts, respectively, and is identified as a *c* subunit gene (*atpH*). The second ORF downstream from *atpH* has sequence similarity to chloroplast I subunits and is therefore identified as a *b* subunit gene (*atpF*). The ORF preceding *atpF* has similarity with the partial amino acid sequence available for subunit II of chloroplasts (4). An equivalent subunit is lacking in the *E. coli* F_0. However, an analogous ORF has been characterized in *Synechococcus* sp. strain PCC 6301 and denoted *b'* (45) on the basis of the similarity of its structure to those of *b* subunits. In accordance with this suggested nomenclature, the fourth ORF in the *Anabaena atpI* cluster is identified as a *b'* (II) subunit gene (*atpG*). The *atpF*, *atpG*, and *atpH* gene sequences encode proteins of M_r 19,754, 17,953, and 7,894, respectively.

Two additional ORFs map upstream of *aptH* in the *atpI* cluster. The first of these does not appear to encode an ATP synthase subunit, but has a similar hydropathy profile to the *uncI* gene product of *E. coli* (data not shown). *uncI*, which has a similar position in the *E. coli* ATP synthase gene cluster, encodes a hydrophobic polypeptide of unknown function. An analogous gene in the *Synechococcus* sp. strain PCC 6301 *atpI* cluster has been denoted gene 1 (9), and this nomenclature is also used for the first ORF in the *Anabaena* sp. strain PCC 7120 gene cluster. The gene 1 product has two potential initiator methionines and therefore has M_r 18,766 or 14,028. Between gene 1 and *atpH* is an ORF with homology to subunits *a* and IV of *E. coli* and chloroplasts, respectively (Table 1); it is identified as an *a* subunit gene (*atpI*). The *atpI* gene encodes a protein of M_r 27,974. The order of the genes in the *Anabaena atpI* gene cluster is thus gene *1-atpI-atpH-atpG-atpF-atpD-atpA-atpC*.

Features of the *atpI* DNA sequence. Within the *atpI* cluster, the spacer regions between genes range from 60 to 170 bp in length, except for the *atpF* and *atpD* arrangement. These genes are fused such that the last two codons of *atpF* overlap the first two codons of *atpD* (Fig. 3). The other intergenic spacer regions of the *atpI* cluster contain numerous short direct repeat sequences (Fig. 2). In the *atpA-atpC* intergenic region, the two sets of direct repeats are complementary to each other (i.e., an inverted repeat), such that a stable stem-loop structure (43) could be formed in the mRNA (Fig. 4). Two other inverted repeat sequences which could form stable stem-loop structures are observed at 29 and 548 bp downstream from the *atpI* cluster (Fig. 2, positions 6399 and 6948; Fig. 4).

Analysis of ATP synthase gene copy number. Radiolabeled DNA probes were prepared from various regions of the ATP synthase gene cluster (Fig. 1, probes 1 to 5) and hybridized to Southern blots of total *Anabaena* sp. strain PCC 7120 DNA digested with *Hin*dIII or *Eco*RI (data not shown). All probes hybridized to single bands of the expected size in each digest, indicating that these genes occur as single copies in the genome.

Analysis of transcripts. Nick-translated probes produced from various regions of *atpI* (Fig. 1) were hybridized with Northern blots of total *Anabaena* sp. strain PCC 7120 RNA to assess the number and sizes of mRNA species derived from the gene cluster. Each probe hybridized with an mRNA species of approximately 7.0 kb (Fig. 5), suggesting that the genes are part of a large transcriptional unit. All probes showed some hybridization to RNA species in the size ranges of the rRNA bands. Probe 1 (Fig. 1) hybridized with two additional bands of approximately 2.1 and 1.2 kb. Probe 2 also hybridized with a 2.1-kb transcript (Fig. 5).

Primer extension assays performed with total RNA and a primer specific to the gene 1 coding region (Fig. 2) produced a single primer-extended product which mapped 140 bp upstream from the predicted translation start for gene 1 (Fig. 6). S1 nuclease mapping with a probe that included the 5' end of *atpI*, all of gene 1, and 500 bp 5' to gene 1 (Fig. 1, probe 1) confirmed the single mRNA endpoint identified in the primer extension assay and demonstrated that no additional RNA endpoints map between gene 1 and *atpI* (data not shown).

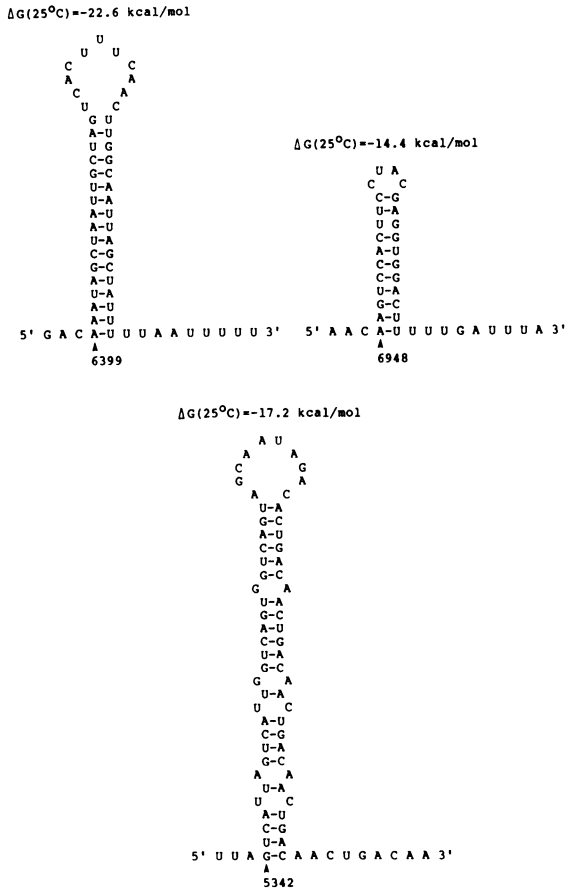


FIG. 4. Nucleotide sequences of regions theoretically capable of forming stable stem-loop structures in mRNA. Nucleotides are numbered as in Fig. 2.

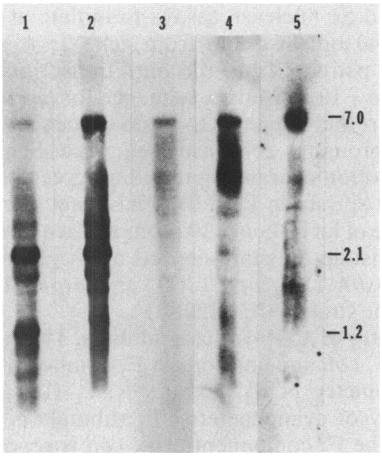


FIG. 5. Identification of *atpI* mRNA species. Total *Anabaena* sp. strain PCC 7120 RNA (20 μg) was denatured, electrophoresed, transferred to nitrocellulose paper, and hybridized with specific probes marked in Fig. 1. Approximate sizes of hybridizing species are shown.

DISCUSSION

Organization of ATP synthase genes in *E. coli*, cyanobacteria, and chloroplasts. In previous work (10) the genes encoding the β (*atpB*) and ε (*atpE*) subunits of the ATP synthase from the filamentous cyanobacterium *Anabaena* sp. strain PCC 7120 were shown to form an operon (*atp2*) and to be unlinked to other ATP synthase genes. A cluster of the remaining eight genes encoding subunits of the ATP synthase has been isolated and characterized by nucleotide sequencing. The genes of this cluster, denoted *atpI*, were identified by comparison of their predicted translation products with amino acid sequences of ATP synthase subunits from other organisms. The organization of the *Anabaena* sp. strain PCC 7120 genes is *atpI-atpH-atpG-atpF-atpD-atpA-atpC*; in addition, a closely linked gene denoted gene 1 is found upstream of *atpI* in this cluster. Gene 1 appears to be equivalent to *E. coli uncI*, which encodes a polypeptide of unknown function. Each of the *Anabaena* sp. strain PCC 7120 ATP synthase genes is single copy. The *atpI* gene cluster is at least 10 kbp from the *atp2* cluster. The same ATP synthase gene organization is observed in the unicellular cyanobacterium *Synechococcus* sp. strain PCC 6301 (8).

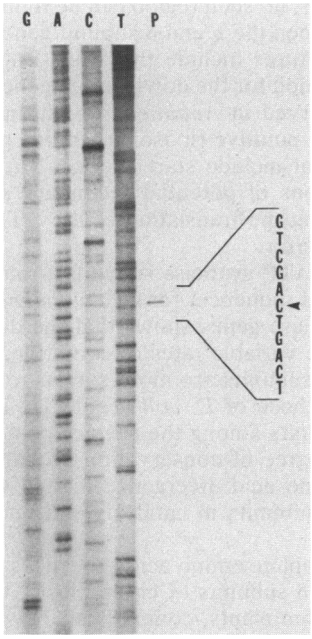


FIG. 6. Identification of the start site for transcription of the *aptI* gene cluster. A 15-nucleotide primer specific to the gene 1 coding region was used in a primer extension assay with total *Anabaena* sp. strain PCC 7120 RNA. The extended product was electrophoresed in parallel with a sequence ladder (G, A, C, T) generated by using the same primer on the noncoding strand of the DNA.

In *E. coli*, eight genes for the ATP synthase subunits and *uncI* are clustered and cotranscribed (15). The two clusters of ATP synthase genes thus characterized from cyanobacteria are ordered in the same manner as in *E. coli*, but both *Synechococcus* sp. strain 6301 and *Anabaena* sp. strain PCC 7120 have an extra gene (*atpG*) relative to *E. coli*. *atpG* encodes a protein similar to the partial amino acid sequence of spinach chloroplast subunit II (4). The *atpG* product has been designated *b'* (45) on the basis of the similarity of its hydropathy profile with those of the *b* subunit of *E. coli* and subunit I of chloroplasts. It has been hypothesized that *atpG* represents a duplicated and diverged copy of *atpF* (9). The characterization of four genes encoding F_0 subunits from two cyanobacterial species suggests that cyanobacterial ATP synthases, like those of chloroplasts (36, 46), contain an F_0 composed of four polypeptides.

The cyanobacterial ATP synthase gene arrangement is reminiscent of the organization in higher-plant chloroplasts. Chloroplast ATP synthase genes are found in two clusters: genes encoding subunits IV (*atpI*), III (*atpH*), I (*atpF*), and α (*atpA*) are found in tandem in that order (9, 18), whereas the *atpB* and *atpE* genes are found in tandem but map elsewhere in the genome (34). Thus, chloroplasts and cyanobacteria have two clusters of ATP synthase genes, with genes within each cluster in the same order. Genes for the remaining chloroplast subunits, II (*atpG*), δ (*atpD*), and γ (*atpC*), however, are encoded in the nucleus (34). The cyanobacterial ATP synthase gene organization is thus intermediate between those observed in *E. coli* and higher-plant chloroplasts.

Features of the *aptI* DNA sequence. A unique feature of the *Anabaena* sp. strain PCC 7120 *atpI* cluster is overlap between the *atpF* and *atpD* genes: nucleotides of the last amino acid codon and stop codons of *atpF* make up the first and part of the second codon of *atpD* (Fig. 3). A similar overlap is not observed in the cyanobacterium *Synechococcus* sp. strain PCC 6301 or in *E. coli*. The *Anabaena* sp. strain PCC 7120 *atpF-atpD* gene overlap cannot be unequivocally established until amino acid sequences for the *b* and δ subunits are available. Two pieces of information, however, support the idea that these genes overlap. (i) All of the *Anabaena* sp. strain PCC 7120 ATP synthase gene coding regions, as well as other *Anabaena* sp. strain PCC 7120 genes thus far characterized, initiate with a methionine codon; the only methionine codon in the *atpD* reading frame is that which overlaps with the *atpF* gene. (ii) The first three amino acids of the *Synechococcus* sp. strain PCC 6301 and *Anabaena* sp. strain PCC 7120 deduced *atpD* gene translation products are identical, and the protein homologies are colinear.

The *Anabaena* sp. strain PCC 7120 *atpF-atpD* gene overlap is reminiscent of the situation found in many plant chloroplasts in which the *atpB* and *atpE* genes overlap (23, 25, 47). The *atpB-atpE* overlap in chloroplasts is similar to

that observed for *atpF* and *atpD* in that the last two codons of *atpB* overlap with the first two codons of *atpE*. *atpB-atpE* overlaps are not observed in *Anabaena* sp. strain PCC 7120 (12), *Synechococcus* sp. strain PCC 6301 (9), or *E. coli* (35). Another feature of chloroplast ATP synthase genes which is absent in *E. coli* and the cyanobacteria studied is observed in the structure of the *atpF* gene. The chloroplast *atpF* genes contain an intron (5, 19, 33, 42), whereas those of *Anabaena* sp. strain PCC 7120 and *Synechococcus* sp. strain 6301 (8) do not.

Most, but not all, of the genes of the *Anabaena* sp. strain PCC 7120 *atpI* cluster are preceded by regions similar to that of *E. coli* ribosome-binding (Shine-Dalgarno) sequences (41). The sequence of the 16S rRNA from *Anabaena* sp. strain PCC 7120 has not been determined, but the 16S rRNA from the cyanobacterium *Synechococcus* sp. strain PCC 6301 has the same sequence as that of *E. coli* in the mRNA binding region (6). Thus, cyanobacterial ribosome-binding sites are expected to be similar to those of *E. coli*. The absence of Shine-Dalgarno sequences has also been observed with the *rbcS* (32) and *psbA* (12) genes of *Anabaena* sp. strain PCC 7120.

The ORFs for gene 1 and *atpA* contain two potential initiator methionines, and since no protein sequence data are available for these polypeptides from *Anabaena* sp. strain PCC 7120, it is not possible to ascertain where translation initiates. The presence of a putative ribosome-binding site near the downstream methionine in gene 1 would suggest that this is the translation start point, and comparison of homologies with the gene 1 products from *E. coli* and *Synechococcus* sp. strain PCC 6301 would support this. The *atpA* gene, however, contains no identifiable ribosome-binding site near either potential initiator methionine codon. As with the gene 1 product, the *Anabaena* sp. strain PCC 7120 α subunit amino acid sequence is homologous with the α subunits from other organisms starting with the downstream methionine. The α subunit from *S. platensis* was estimated by denaturing sodium dodecyl sulfate-gel electrophoresis to be of M_r 53,400 (20). This is in closer agreement to the predicted size (M_r 54,301) of the deduced translation product of *Anabaena atpA* derived from the downstream in-frame methionine.

Many of the intergenic spacer regions of *atpI* are characterized by short direct repeats (Fig. 2). Such repeats are also observed in the *atpB-atpE* spacer of the *atp2* cluster (12). The two sets of repeats in the *atpA-atpC* spacer are complementary; interestingly, these same two sets of repeats are found in the *atpB-atpE* spacer in the same position relative to each other and to the upstream gene. They are also observed downstream of the phycocyanin operon in *Anabaena* sp. strain PCC 7120 (3). The significance of these repeat sequences is not known.

Transcription of *atpI*. All genes of *atpI* hybridized with an mRNA transcript approximately 7.0 kb in length, indicating that genes of this cluster are cotranscribed and thus constitute an operon. Two regions at the 5′ end of *atpI* hybridized to additional RNA species. Probe 1 (Fig. 1) hybridized with a 1.2-kb mRNA species that most probably originates from a gene upstream of *atpI*, whose transcription unit overlaps the 5′ end of probe 1 (D. F. McCarn and S. E. Curtis, unpublished results). Probes 1 and 2 also hybridized with a 2.1-kb species. Since only one RNA endpoint was mapped within probe 1 (see below), this 2.1-kb mRNA must have the same 5′ end as the 7.0-kb mRNA or originate from the opposite strand.

The polycistronic mRNA for *atpI* was shown by primer extension and S1 nuclease assays to initiate at a single site that maps 140 bp upstream from gene 1. A sequence 10 nucleotides upstream from the *atpI* transcription start site resembles the −10 consensus sequence for *E. coli* promoters (37); however, there appears to be no sequence similar to the *E. coli* −35 promoter consensus sequence. Sequences used for the promotion of transcription during vegetative growth in *Anabaena* sp. strain PCC 7120 have not been identified. The presence of an *E. coli* −10 promoter sequence upstream from the transcription start site has also been noted for the *rbcL* (11), *psbA* (12), *petF*1 (1), and *atpB* (10) genes of *Anabaena* sp. strain PCC 7120.

Stoichiometry of ATP synthase subunits. The ATP synthase subunits of *E. coli* and chloroplast F_1 components are found in a stoichiometry of α_3, β_3, γ_1, δ_1, ε_1 (14, 30), and the stoichiometry of cyanobacterial F_1 subunits is expected to be similar. The F_0 component of *E. coli* is reported to have a stoichiometry of a_1, b_2, c_{10-12} (14). Chloroplast and cyanobacterial F_0 components have not been characterized with regard to subunit ratios, but if the stoichiometries are similar to those in *E. coli*, the F_0 may contain one *b* (I) and one *b′* (II) rather than two identical *b* subunits (8). The presence in *E. coli* and *Anabaena* sp. strain PCC 7120 of a single mRNA encoding proteins present in different stoichiometries in the enzyme, together with the absence of evidence for suboperon-length transcripts, suggests that a posttranscriptional mechanism(s) is required to regulate the relative stoichiometries of the subunits. There is evidence that translational control of the operon in *E. coli* is mediated through regions of potential stem-loop structures that include the ribosome-binding sites upstream of gene encoding the *b*, δ, and γ subunits (7, 24).

Analysis of the *aptI* nucleotide sequence revealed several regions of possible secondary structure. As in *E. coli*, there is potential for formation of stem-loop structures in the mRNA between the α and γ subunit genes (Fig. 4). Unlike in *E. coli*, however, no such region can be found between the *c* and *b* and between the *b* and δ subunit genes. In *E. coli* the stem-loop structures include the Shine-Dalgarno sequence and the start codon for the downstream gene. The stem-loop structures observed in *Anabaena* sp. strain PCC 7120 are upstream of the putative ribosome-binding regions and consequently do not include start codons. Thus, it is unlikely that these regions of potential secondary structure play a role in regulating the translation of the ATP synthase subunits in this operon.

Evolution of ATP synthase subunits. Analysis of the derived amino acid sequences for the *Anabaena* sp. strain PCC 7120 ATP synthase genes shows that the different subunits have undergone variable rates of evolution. In general, the cyanobacterial subunits are more similar to those of higher plants than to those of *E. coli* (Table 1). The most highly conserved subunits among the species compared are α and β. This high degree of conservation probably reflects constraints on amino acid divergence related to the proposed roles of these subunits in catalysis and nucleotide binding (30).

Although complete amino acid sequences for the nucleus-encoded γ and δ subunits of chloroplast ATP synthase are not available from plants, comparisons between *E. coli* and cyanobacteria (Table 1) suggest that these subunits are not highly conserved. In *E. coli* the δ subunit is essential for the binding of F_1 to F_0 in membranes (13), and in chloroplasts it has been shown to prevent proton leakage through CF_0 (2). Comparison of the δ subunit sequences from *E. coli*, cyanobacteria, and plant chloroplasts, when available, may pro-

vide information on regions of the polypeptide that are important for subunit function.

The sequence data for the γ subunit poses questions regarding the regulation of cyanobacterial ATP synthases. McCarty and Racker showed that Mg^{2+}-dependent ATP hydrolysis in spinach chloroplasts and Ca^{2+}-dependent ATP hydrolysis in soluble spinach CF_1 can be elicited by light and/or dithiothreitol (29). Nalin and McCarty (31) found a correlation between activation of the latent CF_1 Ca^{2+}-ATPase potential and reduction of a disulfide bond between two cysteine residues in the CF_1 γ subunit. Two additional sulfhydryl groups were identified in the spinach γ subunit protein. In *Anabaena variabilis*, latent Mg^{2+}-ATPase activity in smooth vesicles can be activated by light, a -SH reductant, and an electron donor (35). Thus, the ATP synthase of an *Anabaena* species can undergo a type of activation which is similar to that found in spinach chloroplasts and which contrasts with nonoxygenic photosynthetic and heterotrophic bacteria, in which ATPase activity is not latent (35). Similar light activation is found in the *S. platensis* ATP synthase (21). Analysis of the derived amino acid sequence for the *Anabaena* sp. strain PCC 7120 γ subunit, as well as that for *Synechococcus* sp. strain PCC 6301, however, reveals only a single cysteine residue. Thus, the activation of the latent Ca^{2+}-dependent ATPase activity must occur by a mechanism that is different from the disulfide reduction reported for spinach CF_1.

None of the F_0 subunit amino acid sequences are well conserved between cyanobacteria and *E. coli*; however, the *c* and *a* subunits are well conserved between cyanobacteria and chloroplasts (Table 1). The analysis of cyanobacterial ATP synthase genes suggests that the ATP synthases of cyanobacteria, like those of chloroplasts, have two *b*-type subunits (*b* and *b'*). Neither of these subunits is particularly well conserved between the two cyanobacterial species. Although the F_0 subunits have various degrees of sequence similarity, comparisons of the hydropathy profiles for each of the F_0 subunits from *E. coli*, cyanobacteria, and chloroplasts (8) suggest that their structures are similar. Thus, the maintenance of subunit conformation may be important for the proposed roles of the F_0 subunits in binding of the F_0 to the F_1 portion of the complex and formation of the proton channel (15).

In recent cross-reconstitution studies, Hicks and Yocum (20) showed that light-driven ATP synthesis could be restored by F_1 from *S. platensis* in spinach thylakoid membranes depleted of CF_1 with NaBr; likewise, spinach CF_1 was able to restore ATP synthesis in depleted *S. platensis* membranes. The similarity of cyanobacterial and chloroplast ATP synthases has been confirmed by analysis of cyanobacterial ATP synthase genes. These studies show that the ATP synthases of cyanobacteria and chloroplasts have a similar subunit composition and that many of the subunits have a high degree of sequence similarity. In addition, the ATP synthase genes in chloroplasts and both cyanobacterial species studied display a two-clustered arrangement. The similarity between cyanobacterial and chloroplast ATP synthase structure, as well as ATP synthase gene organization, supports the hypothesis (28) that chloroplasts and cyanobacteria evolved from common ancestors.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant GM32766 from the National Institutes of Health, grant 84-3352 from the North Carolina Biotechnology Center, and grant DMB8614424 from the National Science Foundation. Jawed Alam and Jacqueline Vrba were supported by a postdoctoral fellowship and graduate fellowship, respectively, provided by the North Carolina Consortium for Research and Education in Plant Molecular Biology.

We thank John Boynton and John Gray for providing chloroplast ATP synthase probes.

LITERATURE CITED

1. **Alam, J., R. A. Whitaker, D. W. Krogmann, and S. E. Curtis.** 1986. Isolation and sequence of the gene for ferredoxin I from the cyanobacterium *Anabaena* sp. strain PCC 7120. J. Bacteriol. **168**:1265–1271.
2. **Andreo C, S., W. J. Partie, and R. E. McCarty.** 1982. Effect of ATPase activation and the δ subunit of coupling factor 1 on reconstitution of photophosphorylation. J. Biol. Chem. **257**: 9968–9975.
3. **Belknap, W. R., and R. Haselkorn.** 1987. Cloning and light regulation of expression of the phycocyanin operon of the cyanobacterium *Anabaena*. EMBO J. **6**:871–884.
4. **Berzborn, R. J., J. Otto, W. Finke, H. E. Meyer, and J. Block.** 1987. Conclusions from N-terminal amino acid sequences of subunits delta from spinach and maize CF_1 and of subunits I and II from spinach CF_0. Biol. Chem. Hoppe-Seyler **368**:551–552.
5. **Bird, C. R., B. Koller, A. D. Auffret, A. K. Huttley, C. J. Howe, T. A. Dyer, and J. C. Gray.** 1985. The wheat chloroplast gene for CF_0 subunit I of ATP synthase contains a large intron. EMBO J. **4**:1381–1388.
6. **Borbely, G., and A. Simoncsits.** 1981. 3'-Terminal conserved loops of 16S rRNAs from the cyanobacterium *Synechococcus* PCC 6301 and maize chloroplast differ only in two bases. Biochem. Biophys. Res. Commun. **101**:846–852.
7. **Brusilow, W. S. A., D. J. Klionsky, and R. D. Simoni.** 1982. Differential polypeptide synthesis of the proton-translocating ATPase of *Escherichia coli*. J. Bacteriol. **151**:1363–1371.
8. **Cozens, A. L., and J. E. Walker.** 1987. The organization and sequence of the genes for ATP synthase subunits in the cyanobacterium *Synechococcus* 6301. J. Mol. Biol. **194**:359–383.
9. **Cozens, A. L., J. E. Walker, A. L. Phillips, A. K. Huttly, and J. C. Gray.** 1986. A sixth subunit of ATP synthase, an F_0 component, is encoded in the pea chloroplast genome. EMBO J. **5**:217–222.
10. **Curtis, S. E.** 1987. Genes encoding the beta and epsilon subunits of the proton-translocating ATPase from *Anabaena* sp. strain PCC 7120. J. Bacteriol. **169**:80–86.
11. **Curtis, S. E., and R. Haselkorn.** 1983. Isolation and sequence of the gene for the large subunit of ribulose-1,5-bisphosphate carboxylase from the cyanobacterium *Anabaena* 7120. Proc. Natl. Acad. Sci. USA **80**:1835–1839.
12. **Curtis, S. E., and R. Haselkorn.** 1984. Isolation, sequence and expression of two members of the 32kd thylakoid membrane protein gene family from the cyanobacterium *Anabaena* 7120. Plant Mol. Biol. **3**:249–258.
13. **Dunn, S. D., and L. A. Heppel.** 1981. Properties and functions of the subunits of the *Escherichia coli* coupling factor ATPase. Arch. Biochem. Biophys. **210**:421–436.
14. **Foster, D. L., and R. H. Fillingame.** 1982. Stoichiometry of subunits in the H^+-ATPase of *Escherichia coli*. J. Biol. Chem. **257**:2009–2015.
15. **Futai, M., and H. Kanazawa.** 1983. Strucutre and function of proton-translocating adenosine triphosphatase (F_0F_1): biochemical and molecular biological approaches. Microbiol. Rev. **47**: 285–312.
16. **Gay, N. J., and J. E. Walker.** 1981. The *atp* operon: nucleotide sequence of the region encoding the α subunit of *Escherichia coli* ATP-synthase. Nucleic Acids Res. **9**:2187–2194.
17. **Gay, N. J., and J. E. Walker.** 1981. The *atp* operon: nucleotide sequence of the promoter and the genes for the membrane proteins, and the α subunit of *Escherichia coli* ATP synthase. Nucleic Acids Res. **9**:3919–3926.
18. **Hatefi, Y.** 1985. The mitochondrial electron transport and oxidative phosphorylation system. Annu. Rev. Biochem. **54**:1015–1069.
19. **Hennig, J., and R. G. Herrmann.** 1986. Chloroplast ATP synthase of spinach contains nine nonidentical subunit species, six

of which are encoded by plastid chromosomes in two operons in a phylogenetically conserved arrangement. Mol. Gen. Genet. **8:** 543–549.

20. **Hicks, D. B., and C. F. Yocum.** 1986. Properties of the cyanobacterial coupling factor ATPase from *Spirulina platensis*. I. Electrophoretic characterization and reconstitution of photophosphorylation. Arch. Biochem. Biophys. **245:**220–229.
21. **Hicks, D. B., and C. F. Yocum.** 1986. Properties of the cyanobacterial coupling factor ATPase from *Spirulina platensis*. II. Activity of the purified and membrane-bound enzyme. Arch. Biochem. Biophys. **245:**230–237.
22. **Howe, C. J., A. D. Auffret, A. Doherty, C. M. Bowman, T. A. Dyer, and J. C. Gray.** 1982. Location and nucleotide sequence of the gene for the proton-translocating subunit of wheat chloroplast ATP synthase. Proc. Natl. Acad. Sci. USA **79:**6903–6907.
23. **Howe, C. J., I. M. Fearnley, J. E. Walker, T. A. Dyer, and J. C. Gray.** 1985. Nucleotide sequences of the genes for the alpha, beta and epsilon subunits of wheat chloroplast ATP synthase. Plant Mol. Biol. **4:**333–345.
24. **Klionsky, D. J., D. G. Skalnik, and R. D. Simoni.** 1984. Differential translation of the genes encoding the proton-translocating ATPase of *Escherichia coli*. J. Biol. Chem. **18:**8096–8099.
25. **Krebbers, E. T., I. M. Larrinua, L. McIntosh, and L. Bogorad.** 1982. The maize chloroplast genes for the β and ε subunits of the photosynthetic coupling factor CF_1 are fused. Nucleic Acids Res. **10:**4985–5002.
26. **Lubberding, H. J., G. Zimmer, H. S. Van Walraven, J. Schrickz, and R. Kraayenhof.** 1983. Isolation, purification and characterization of the ATPase complex from the thermophilic cyanobacterium *Synechococcus* 6716. Eur. J. Biochem. **137:**95–99.
27. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
28. **Margulis, L.** 1981. Symbiosis in cell evolution, p. 37–62. W. H. Freeman and Co., San Francisco.
29. **McCarty, R. E., and E. Racker.** 1968. Partial resolution of the enzymes catalyzing photophosphorylation. III. Activation of adenosine triphosphatase and ^{32}P-labeled orthophosphate adenosine triphosphate exchange in chloroplasts. J. Biol. Chem. **243:** 129–137.
30. **Merchant, S., and B. R. Selman.** 1985. Photosynthetic ATPases: purification, properties, subunit isolation and function. Photosynth. Res. **6:**3–31.
31. **Nalin, C. M., and R. E. McCarty.** 1984. Characterization of the cysteinyl-containing peptides of the γ subunit of coupling factor 1. J. Biol. Chem. **259:**7281–7285.
32. **Nierzwicki-Bauer, S. A., S. E. Curtis, and R. Haselkorn.** 1984. Cotranscription of genes encoding the small and large subunits of ribulose-1,5-biphosphate carboxylase in the cyanobacterium *Anabaena* 7120. Proc. Natl. Acad. Sci. USA **81:**5961–5965.
33. **Ohyama, K., H. Fukuzawa, T. Kohchi, H. Shirai, T. Sano, K. Umesono, Y. Shiki, M. Takeuschi, Z. Chang, S. Aota, H. Inokuchi, and H. Ozeki.** 1986. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature (London) **322:**572–574.
34. **Palmer, J. D.** 1985. Comparative organization of chloroplast genomes. Annu. Rev. Genet. **19:**325–354.
35. **Petrack, B., and F. Lipmann.** 1961. Photophosphorylation and photohydrolysis in cell-free preparations of blue-green algae, p. 621–630. *In* W. D. McElroy and B. Glass (ed.), Light and life. The Johns Hopkins Press, Baltimore.
36. **Pick, U., and E. Racker.** 1979. Purification and reconstitution of the N,N′-dicyclohexycarbodiimide-sensitive ATPase complex from spinach chloroplasts. J. Biol. Chem. **254:**2793–2799.
37. **Reznikoff, W. S., D. A. Siegele, D.W. Cowing, and C. A. Gross.** 1985. The regulation of transcription initiation in bacteria. Annu. Rev. Genet. **19:**355–388.
38. **Rice, D., B. J. Mazur, and R. Haselkorn.** 1982. Isolation and physical mapping of nitrogen fixation genes from the cyanobacterium *Anabaena* 7120. J. Biol. Chem. **257:**13157–13167.
39. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
40. **Saraste, M., N. J. Gay, A. Eberle, M. J. Runswick, and J. E. Walker.** 1981. The *atp* operon: nucleotide sequence of the genes for the γ, β, and ε subunits of *Escherichia coli* ATP synthase. Nucleic Acids Res. **9:**5286–5296.
41. **Shine, J., and L. Dalgarno.** 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71:**1342–1346.
42. **Shinozaki, K., H. Deno, T. Wakasugi, and M. Sugiura.** 1986. Tobacco chloroplast gene coding for subunit I of proton-translocating ATPase: comparison with the wheat subunit I and *E. coli* subunit *b*. Curr. Genet. **10:**421–423.
43. **Tinoco, I., Jr., P. N. Borer, B. Dengler, M. Levine, O. C. Uhlenbeck, D. M. Crothers, and J. Gralla.** 1973. Improved estimation of secondary structure in ribonucleic acids. Nature (London) New Biol. **246:**40–41.
44. **Tumer, N. E., S. J. Robinson, and R. Haselkorn.** 1983. Different promoters for the *Anabaena* glutamine synthetase gene during growth using molecular or fixed nitrogen. Nature (London) **306:** 337–341.
45. **Walker, J. E., and V. L. J. Tybulewicz.** 1985. Comparative genetics and biochemistry of light-driven ATP synthases, p. 141–153. *In* C. Arntzen, L. Bogorad, S. Bonitz, and K. Steinback (ed.), Molecular biology of the photosynthetic apparatus. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
46. **Westhoff, P., J. Alt, N. Nelson, and R. G. Herrmann.** 1985. Genes and transcripts for the ATP synthase CF_0 subunit I and II from spinach thylakoid membranes. Mol. Gen. Genet. **199:**290–299.
47. **Zurawski, G., W. Bottomley, and P. R. Whitfeld.** 1982. Structures of the genes for the β and ε subunits of spinach chloroplast ATPase indicate a dicistronic mRNA and an overlapping translation stop/start signal. Proc. Natl. Acad. Sci. USA **79:**6260–6264.