MOLECULAR ANALYSIS OF CYANOBACTERIAL RNA POLYMERASE GENES

by

WenQin Xie

Dissertation submitted to the faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY
in
Biochemistry and Nutrition
APPROVED:

Malcolm Potts, Chairman

Thomas W. Keenan

Thomas O. Sitz

Robert H. White

John L. Johnson

August, 1989
Blacksburg, Virginia

MOLECULAR ANALYSIS OF CYANOBACTERIAL RNA POLYMERASE GENES

by

WenQin Xie

Malcolm Potts, Chairman

Department of Biochemistry and Nutrition

(ABSTRACT)

The RNA plymerase genes rpoBC1C2 (the rpoB and rpoC2 are incomplete) of the cyanobacterium Nostoc commune have been cloned, sequenced, and expressed both in vivo and in vitro using E. coli systems. The rpo genes encode large subunits of DNA-dependent RNA polymerase. Two genes in N. commune, rpoC1 and rpoC2, correspond to rpoC of E. coli, which indicates a divergent evolution of RNA polymerase. The rpoBC1C2 genes of Nostoc are linked in the order of rpoB, rpoC1, and rpoC2, and are transcribed differently from the corresponding rpo genes of E. coli. In E. coli the rpoBC genes are co-transcribed, together with two ribosomal protein genes. The Nostoc rpoB gene is transcribed by one promoter while the rpoC1C2 genes are co-transcribed by another promoter. Northern analysis of Nostoc RNA revealed two transcripts (3.1 and 5.6 kb), which were specific for the rpoB and rpoC1C2 genes, respectively. SDS-PAGE, Coomassie staining and immunoanalysis detected two polypeptides with molecular weights of 72 and 94 kd when the cloned rpoBC1C2 fragment was expressed in E. coli systems. These two polypeptides were assigned as products of rpoC1 and rpoC2, respectively. The

transcription of RNA polymerase genes of N. commune is sensitive to water stress (drying). The rpo transcription ceases upon drying and resumes after the dried cells have been rewetted for more than 5 min. The RNA polymerase enzyme itself, however, is stable under the same drying conditions.

To

ACKNOWLEDGMENTS

I would like to thank Dr. Malcolm Potts for giving me the opportunity to work in his laboratory, and for his guidance and encouragement. I am grateful to my committee members: Drs. Tom Keenan, Tom Sitz, Robert White, and John Johnson for their valuable suggestions and help during my PhD study. I also wish to extend my thanks to Dr. Tim Larson and all other faculty members in the Department of Biochemistry and nutrition for their scientific discussion and advice. I give special thanks to for sharing his expertise in using computers.

I also thank all those friends in the Biochemistry Department and in Blacksburg who have made my stay here so pleasant. I would like to express my personal thanks to my mother who always loves me and who is always there, and to my husband for his patience and support. Finally, I would like to thank my dearest little son for being such a good baby.

WenQin Xie

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

List of Abbreviations

A_{600}: absorbance at 600 nm

A: adenine; one letter code for alanine

AMP: Ampicillin

ATP: adenosine 5'-monophosphate

BCIP: 5-bromo-4-chloro-3-indolyl phosphate

BSA: Bovine serum albumin

C: cytosine; one-letter code for cysteine

CTP: cytidine triphosphate

EtBr: ethidium bromide

dATP: deoxyadenosine triphosphate

dCTP: deoxycytidine triphosphate

ddNTP: dideoxynucleotide triphosphate

dGTP: deoxyguanosine triphosphate

dTTP: deoxythymidine triphosphate

DEPC: diethylpyrocarbonate

DNA: deoxyribonucleic acid

DNase: deoxyribonuclease

DTT: dithiothreitol

EDTA: ethylenediaminetetraacetic acid

g: force of gravity

HRP: horse-radish peroxidase

Ig: immunoglobulin

IPTG: isopropyl-1-thio- D-galactoside

l: liter

kb: kilobases

kd: kilodaltons

M_r: relative molecular weight

NBT: nitro-blue tetrazolium

OD: optical absorbance

ORF: open reading frame

rpo: RNA polymerase operon

PAGE: polyacrylamide gel electrophoresis

PEG: polyethylene glycol

pfu: plaque-forming units

RNA: ribonucleic acid

RNase: ribonuclease

mRNA: messenger ribonucleic acid

rRNA: ribosomal ribonucleic acid

tRNA: transfer ribonucleic acid

sarcosyl: N-lauryl sarcosine

SA-AP: streptavidin-alkaline phosphatase

SDS: sodium dodecyl sulfate

SDS-PAGE: SDS polyacrylamide gel eletrophoresis

TAC: transcriptionally active chromosome

TCA: trichloracetic acid

Tris: tri(hydroxymethyl)aminomethane

Tris-HCl: Tris-hydrochloride

U: units

UV: ultraviolet

UTP: uridine triphosphate

LIST OF FIGURES

INTRODUCTION

The transcription of genes is directed through the activity of DNA-dependent RNA polymerase (ribonucleotide triphosphate RNA nucleotidyl transferase, EC 2.7.7.6), together with ancilliary factors. There is a single form of RNA polymerase in prokaryotes, which is responsible for the synthesis of all three kinds of RNA; mRNA, rRNA and tRNA (Burgess, 1969 a, b). The RNA polymerase of E. coli, which has been studied extensively, consists of four different polypeptides, $(alpha)_2$, beta, beta', and sigma, with molecular weights of 40,000, 150,000, 160,000, and 86,000, respectively (Chamberlin, 1976). The sigma factor binds only loosely to the core enzyme and can be dissociated by phosphocellulose chromatography during the process of purification. The protein complexes with or without the sigma factor are classified as the holoenzyme and core enzyme, respectively.

The genes encoding each of the subunits of RNA polymerase of E. coli have been cloned and sequenced (Dennis et al., 1985; Ovchinnikov et al., 1981). The rpoBC genes, coding for the beta and beta' subunits of E. coli, are linked and under the control of the same promotor (PL10), which together with the genes encoding L10 (rplJ) and L12 (rplL) ribosomal proteins, form the L10 operon (Dennis et al., 1985).

RNA polymerases have been purified to homogeneity from eukaryotes (Chambon, 1975; Roeder, 1976, Sentenac, 1985). There are three different kinds of RNA polymerase in eukaryotes, RNA polymerases I, II, and III. They are characterized by their different functions and their resistances to the toxin

alpha-amanitin. RNA polymerase I is resistant to amanitin and is responsible for the synthesis of 18S, 28S and 5.8S rRNA. RNA polymerase II is very sensitive to the toxin and is responsible for the synthesis of mRNA, and most small nuclear RNAs. RNA polymerase III is relatively less sensitive to the toxin and is selective for the synthesis of tRNA, 5S RNA, and other types of small nuclear RNA and cytoplasmic RNA.

Eukaryotic RNA polymerases are more complex enzymes than the prokaryotic forms, and consist of two non-identical large subunits, A and B, and at least seven small subunits (Paul, 1981; Lewis and Burgess, 1982). In general these enzymes are capable of random initiation on DNA and can start transcription at single-strand breaks or "nicks" in the template. The function of each subunit of the eukaryotic enzyme is poorly defined. The nomenclature for different subunits of eukaryotic RNA polymerase also has not been standardized.

Using alpha-amanitin as a genetic indicator, the gene encoding the largest subunit (A, 215,000 daltons) of the eukaryotic RNA polymerase has been cloned from Drosophila melanogaster (Searles et al., 1982). Using this gene as a probe, the corresponding genes have been isolated from yeast (Ingles et al., 1984), from mouse (Cordon et al., 1985), and from human (Cho et al., 1985). The deduced amino acid sequence of this gene has sequence similarity with that of the beta' subunit of RNA polymerase of E. coli.

The genes coding for the 220,000 daltons (A) and 150,000 daltons (B) subunits of polymerase II of yeast have been cloned using antibodies directed against the individual subunits (Young and Davis, 1983). A series of clones, which potentially code for each of the subunits of RNA polymerases I, II and III, have been isolated from yeast DNA using antibodies (Riva et al., 1986).

Using a DNA fragment containing the gene coding for the B subunit (150,000 daltons) of RNA polymerase of yeast, the corresponding gene has been isolated from Drosophila melanogaster (Faust et al., 1986). Studies of the gene encoding the second largest subunit (B, 140,000 M_r) of RNA polymerase II indicate nine regions with striking similarity to regions within the beta subunit of E. coli (Falkenburg et al., 1987).

The RNA polymerases of organelles, plastids and mitochondria, are less well characterized. The exact subunit composition of these organellar RNA polymerases is uncertain. Recently, some progress has been made in the characterization of the RNA polymerases of chloroplasts.

There are two RNA polymerase activities in the chloroplasts of Euglena gracilis (Greenburg et al., 1984). One polymerase activity is tightly bound to chloroplast DNA and this complex is called the transcriptionally active chromosome (TAC). This activity is responsible for the synthesis of rRNA in chloroplasts. The other polymerase activity is found in a soluble extract of the chloroplast, and is selective for the synthesis of mRNA and tRNA in the chloroplasts of this organism.

DNA sequences which are homologous to rpoB and rpoC of E. coli have been found in genomes of chloroplasts of some plants (Ohyama et al., 1988; Ohme et al., 1986). In spinach chloroplasts, sequences corresponding to rpoA, rpoB and rpoC have been cloned and sequenced. Evidence suggests that these genes encode four subunits of the RNA polymerase which is active in the synthesis of tRNA and mRNA in chloroplasts (Little et al., 1988; Hudson et al., 1988). The rpoA gene codes for the alpha subunit. The rpoB gene codes for a 121,000 M_r polypeptide, which is homoloqous to the beta subunit of the RNA polymerase of E. coli.

Whereas two rpoC-like genes (rpoC1 and rpoC2) code for two polypeptides (78,000 and 154,000 M_r) homologous to the N- and C-terminal portions, respectively, of the E. coli beta' subunit. As in spinach, the rpoB, rpoC1C2 genes of the chloroplast are linked and cotranscribed, though there is an intron in the rpoC1 gene (Hudson et al., 1988).

RNA polymerases have been purified from six species of archaebacteria, representing four different orders (Zillig and Stetter, 1985). The RNA polymerase of archaebacteria has one of two basic structures [BAC or B"B'AC]. Archaebacteria possess only one kind of RNA polymerase like eubacteria (Leffers et al., 1989). However, the component patterns of the archaebacterial enzymes on SDS-PAGE are rather complex, resembling eukaryotic RNA polymerases. Seven components have been reproducibly found in the enzyme from Halobacterium sp., a representative of Halobacteria, and 11 components in the enzyme from +Thermoplasma sp., a representative of the sulfur-dependent archaebacteria.

Genes encoding the four largest subunits of the RNA polymerase of archaebacteria, B", B', A, and C, have been cloned and sequenced from Methanobacterium thermoautotrophicum (Schallenberg et al., 1988). The four genes form a cluster in this order: B", B', A, and C, and have a common orientation.

Like other prokaryotes, cyanobacteria have only one kind of RNA polymerase. The cyanobacterial RNA polymerase has been purified from Anacystis nidulans (Herzfeld and Zillig, 1971) and from Anabaena PCC 7120 (Schneider et al., 1987). The component pattern of RNA polymerase of cyanobacteria resembles that of the E. coli RNA polymerase. Recently, a unique subunit, gamma, has been discovered in core enzyme of the RNA polymerase of Anabaena, in addition to

$(alpha)_2$, beta and beta' (Schneider et al., 1987). The same subunit has been reported in all 15 tested species of cyanobacteria (Schneider et al., 1988). Antiserum directed against the gamma subunit showed a cross reactivity with the beta' subunit of RNA polymerase of E. coli. The molecular weights of alpha, beta, beta' and sigma subunits of the RNA polymerase of Anacystis have been reported as 39,000, 147,000, 125,000, and 86,000, respectively (Herzfeld and Zillig, 1971). The molecular weights of alpha, beta, gamma, beta' and sigma subunits of Anabaena have been reported as 41,000, 124,000, 66,000, 171,000 and 52,000, respectively.

The primary amino acid sequence of RNA polymerase is conserved as a consequence of the enzyme's basic function. Similarities between RNA polymerases, especially between the two large subunits, have been shown from archaebacteria to eubacteria and to eukaryotes at both the DNA and protein levels (Allison et al., 1986; Puhler et al., 1989). Comparison of the amino acid sequences of the largest subunit (A), of RNA polymerases II and III of Saccharomyces cerevisiae and the beta' subunit of E. coli RNA polymerase, revealed six regions (I to VI) of marked conservation (Allison et al., 1986). Furthermore, antigenic determinants are conserved among these subunits from different organisms (Zillig and Groupp, 1987). The B subunit of eukaryotic RNA polymerases II & III is serologically related to the B or B" & B'subunits of archaebacterial RNA polymerases and to the beta subunit of eubacterial RNA polymerases; The A subunit of eukaryotic RNA polymerases II & III is serologically related to the A & C subunits of archaebacterial RNA polymerases and to the beta' subunit of

eubacterial RNA polymerases (E. coli). A summary of gene structures and organizations of the rpo genes, encoding the two large subunits of RNA polymerases from different organisms, is presented in Figure 1.

The significance of each subunit of RNA polymerase, as well as how each of the subunits interact with one other in the active enzyme, is not known. Since the two largest subunits of the enzyme are conserved among different organisms, it is believed that these two subunits play a crucial role in the polymerization of ribonucleotides. Most knowledge contributing to the understanding of the function of individual subunits of RNA polymerase comes from studies of the polymerase of E.coli. The core enzyme is capable of polymerizing ribonucleotides, while the sigma factor is required for initiation of RNA synthesis at specific sites. It is believed that the beta subunit is involved in all stages of RNA synthesis including promotor selectivity, substrate binding, catalysis, initiation, stringent response and chain termination (Nomura et al., 1984). The beta' subunit contains the DNA template binding site and also is involved in the recognition of promotors (Nomura et al., 1984). The two alpha subunits assemble the two large subunits to form the core enzyme. The sensitivity of a bacterium to inhibitors of transcription, such as rifampicin, streptolydigin, and streptovaricin, is a consequence of interaction of the inhibitors with the beta subunit.

RNA polymerase is involved in the first step of gene expression, transcription. In order to funcation physiologically, there must be a mechanism which controls where and when a polymerase should work. It is unknown what mechanisms control the specificity of RNA polymerase, but it is clear that sigma factors as well as other protein factors are components which control the selectivity. A novel sigma factor is involved in the heat shock response (Grossman et al., 1985).

Multiple sigma-like proteins have been isolated from Bacillus subtilis, and one of them is related directly to the control of the sporulation process in this organism (Haldenway et al., 1980, 1981).

Organisms are classified into three different kingdoms: archaebacteria, eubacteria and eukaryotes, using 16S ribosomal RNA as an evolutionary chronometer. The differences in these organisms' RNA polymerases provide an alternative means of classification due to the fact that these enzymes are ubiquitous, highly-conserved, and a complex macromolecules. Even though it is believed that archaebacteria are a more "primitive" form of life, their RNA polymerases are complex and resemble those of eukaryotes. The immuno-cross reaction of RNA polymerases is marked between archaebacteria and eukaryotes and less so between archaebacteria and eubacteria. The consensus sequence of archaebacterial promoters is more similar to that of eukaryotic RNA polymerase II promoters than to promotors of eubacterial RNA polymerases (Puhler et al., 1989).

Most archaebacteria live in extreme environments, at high temperature, at high ionic strength, at low pH, or at low oxygen concentration. The extreme conditions of high temperature and high ionic strength represent also the intracellular condition. The transcription systems of these organisms must have adapted to these conditions in special ways. How RNA polymerase responds to shifts of temperature, to shifts of salt concentration, or to shifts of pH, is not known. Water stress provides another challenge to the regulation of gene expression. The present

study was aimed at providing some information about what happens to the RNA polymerase, the key control factor of the transcription system, during the processes of desiccation and rewetting of the cyanobacterium Nostoc commune.

Nostoc commune is a filamentous nitrogen-fixing cyanobacterium. Cyanobacteria (blue-green algae) are a diverse group of photosynthetic prokaryotes. With a few exceptions, cyanobacteria are obligate photoautotrophs and use CO_2 as their sole carbon source and light as their energy source. The photosynthetic system of cyanobacteria consists of chlorophyll a and phycobiliproteins (phycoerythrin, phycocyanin and allo-phycocyanin) which absorb light in the range of 560-650 nm. Some cyanobacteria are able to use organic substrates in addition to, or instead of CO_2, while remaining dependent upon light for an energy source. These organisms are classified as facultative photoheterotrophs. Cyanobacteria that are capable of using organic substrates as both carbon and energy sources in the dark are facultative chemoheterotrophs. Cyanobacteria are capable of switching from a chloroplast type oxygen-evolving photosynthesis (PSI and PSII) to anaerobic photosynthesis (PSI only). In the absence of a fixed nitrogen source vegetative cells of Nostoc can undergo differentiation to form structurally- and biochemically-modified cells, the heterocysts, which permit the organisms to fix nitrogen aerobically (Carr and Whitton, 1982). Species of Nostoc are among the most widespread microorganisms, owing to their ability to tolerate water stress, desiccation and extremes of temperature (Whitton and Potts, 1982).

Nostoc (field material) can still recover cellular activities upon rewetting after being dried for over 100 years (Cameron, 1962). Upon rewetting, desiccated cells of Nostoc resume respiration, photosynthesis and nitrogen fixation sequentially (Scherer et al.,1984). Fermentor-grown material of Nostoc UTEX584 has shown sensitivity to water stress (Potts and Bowman, 1985). The drying of cells and rewetting of desiccated cells leads to rapid and significant changes in cell volume and turgor pressure, which must impose considerable stress upon the cells; however, the cell membrane remains intact during these processes (Olie and Potts, 1986). Cells maintain an intact protein synthesis complex during the first two hours of the drying process although there is a decrease in the polysome content of the cells (Angeloni and Potts, 1986). The polysomes recover quickly upon the rewetting of cells. The protein index of Nostoc commune UTEX584 changes when cells are subjected to the drying process (Potts, 1986). Novel proteins have been isolated from N. commune (field material) when the cells were subjected to drying (Scherer and Potts, 1989).

In recent years progress has been made in the molecular biology of cyanobacteria. Nif genes of Anabaena have been cloned and studied by Haselkorn's group (Golden et al., 1985; Mazur et al., 1980; Rice et al., 1981). The genes coding for two members of a thylakoid membrane protein gene family have been cloned, sequenced and expressed in E. coli (Curtis et al., 1984). A recA-like gene of Synechococcus sp. has been cloned and characterized (Murphy et al., 1987). The leuB gene , which codes for an enzyme in the leucine biosynthesis pathway, has been isolated (Cangelosi et al., 1986). The apcA and apcB genes, encoding components of phycobilisomes, have been cloned (Houmard, et al., 1986). The gene, coding for a photosystem II polypeptide, has been cloned,

sequenced, and expressed in E. coli (Golden et al., 1982). In Nostoc commune UTEX584, nifUHD genes and the gene encoding an indole phosphate hydrolase have been isolated (Defrancesco and Potts, 1988; Xie et al., 1989). Further progress includes the development of conjugation and transformation systems (Buzhy et al., 1985; Herrero and Wolk, 1986), and the construction of shuttle vectors and systems for genetic mapping (Gendle et al., 1983).

The transcription and translation systems of cyanobacteria have seldom been studied. The control of gene expression under water stress is also poorly understood. The purpose of this study was to provide fundamental data about one of the key factors of transcription, RNA polymerase, in cyanobacteria. How are these rpo genes, corresponding to the two largest subunits of RNA polymerase of E. coli, organized and regulated in cyanobacteria? How do RNA polymerase and its encoding genes respond to water stress?

Material and Methods

All procedures in this study were performed at room temperature unless otherwise indicated. All centrifugations, depending upon the volume of samples, were performed using a microcentrifuge (Fisher, Model 235B) or a table-top clinical centrifuge (IEC Centra-7R), unless otherwise indicated. Disposable plastic gloves were worn whenever RNA samples and RNA reagents were handled. All nitrocellulose membranes (0.45 um) and Nytran membranes (0.2 um) used in this research were obtained from Schleicher & Schuell and Immobilon membranes were from Millipore. The sonicator used in this study was a Fisher Sonic Dismembrator (Model 300). The compositions of buffers are listed in Appendix 1 unless otherwise listed in the test. The characteristics of bacterial strains, plasmids and phage are listed in Appendix 2.

1. Growth and Maintenance of Cells

Nostoc commune UTEX 584 was grown in continuous axenic culture in a 1.8-l airlift fermenter (BRL/Life Technologies, Inc., Gaithersburg, MD, USA). Cells were grown in BG11 medium (Rippka et al., 1979) at 32 °C, with a photon flux density of 10 umol photons m^{-2} s^{-1} (Jager and Potts, 1988). For DNA isolation the cells were grown to very high density. Exponential cells were obtained as follows: the cells were grown to high cell density, the fermenter then was emptied to 1/3 volume, and refilled with fresh medium, and the growth was continued for an additional six hours. The cells were harvested by centrifugation at 16,000g for 10 min, at 4°C, using the Sorvall GSA rotor in a superspeed centrifuge (Sorvall RC-5B).

2. Preparation and Manipulation of DNA:

2.1 Restriction Enzyme Digestion of DNA

All restriction enzymes and other DNA modification enzymes were used by following the manufacturer's suggestions unless described otherwise.

2.2 Phenol/Chloroform Extraction of DNA and RNA

The DNA or RNA solution to be purified, was extracted respectively, with an equal volume of pre-equilibrated phenol, phenol:chloroform mixture (1:1 v/v), and chloroform:isoamyl alcohol mixture (24:1 v/v). After each extraction the two phaces were separated by centrifugation. The phenol/chloroform extraction was repeated until the interface was clear. Distilled phenol was mixed with m-cresol and 8-hydroxyquinoline (ratio of 1000:140:0.4, w/w/w), and then the mixture was equilibrated with 1 M Tris-HCl, pH 8.0, until the pH of the phenol reached approximately 8.

2.3 Ethanol Precipitation of Nucleic Acids

DNA or RNA was precipitated as follows unless otherwise described. The salt concentration of the DNA or RNA solution to be precipitated was adjusted to 0.3 M with 3 M NaOAc (pH 5.2), and the nucleic acids were precipitated with two volumes of 95% ethanol. The precipitation was allowed to proceed at -70°C until the solution got frozen, or at -20°C overnight.

2.4 Cesium Chloride (CsCl) Density Gradient Ultracentrifugation

Cesium chloride (0.742 mg/ml) was dissolved in TE buffer (10 mM Tris-HCl, pH 8.0 and 1 mM EDTA), supplemented with ethidium bromide (EtBr, 100 ug/ml). The centrifugation was performed in 10-ml polycarbonate tubes at 45,000 rpm in a Beckman Ti 50 rotor, at 18°C for 44 hours, or in a Beckman VTi rotor for

10 to 18 hours. The DNA bands which were usually visible to the naked eye, or which were located by illumination with UV irradiation, were collected with a sterile Pasteur pipette. The DNA solution was desalted by passage through a Sephadex column (NAP-25, Pharmacia), and then dialysed against TE buffer (4 l) overnight with one buffer change.

2.5 Isolation of DNA from Nostoc commune UTEX584

The cell pellet (about 40g) was frozen in liquid nitrogen and pulverized to a powder with a liquid nitrogen-chilled mortar and pestle. The cell powder was then resuspended in buffer containing 15% (w/v) sucrose, 50 mM Tris-HCl (pH 7.5), 10 mM EDTA, and then subjected to five freeze-thaw cycles using liquid nitrogen and warm tap water. The cells were treated with lysozyme (10 mg/ml) for 4.5 hours , at 37°C with shaking until the color of the solution became brown as the result of release and oxidation of phycobiliproteins. Proteinase K (20 ug/ml) and sarcosyl (0.5% w/v) were added, and the solution was kept at 50°C, with mild shaking for four hours. The solution was then extracted with phenol/chloroform. The aqueous solution was adjusted to 0.3 M NaOAc with 3.0 M stock NaOAc (pH 5.2). DNA was precipitated by addition of an equal volume of ice-cold isopropanol. The precipitated DNA was spooled using a glass rod, and washed with 70% ethanol. The spooled DNA was resuspended in TE buffer and subjected to further purification using CsCl density gradient ultracentrifugation.

2.6 Plasmid DNA Mini-preparation

Plasmid DNA was isolated using the alkaline lysis procedure of Bambing and Doly as procedure described by Promega (Promega Catalogue p3-4,1986/1987). Bacterial cells, carrying appropriate plasmid DNA, were grown overnight, with shaking, in 5 ml of SOB medium supplemented with 100 ug/ml Ampicillin (Amp).

Cells from 1.5 ml of cultures were collected by centrifugation in a microcentrifuge tube for 1 min. The cells were resuspended, by vortexing, in 100 ul of an ice-cold solution of 50 mM glucose, 10 mM EDTA, and 25 mM Tris-HCl, pH 8.0, and the mixture was incubated for 5 min. To this mixture, 200 ul of a freshly-prepared solution of 0.2 N NaOH and 1% SDS was added, and the solution was mixed by inversion. The mixture was then incubated for 5 min on ice. The solution was then neutralized by inversion after the addition of 150 ul of ice-cold potassium acetate, pH 4.8 (the solution was 3 M with respect to potassium and 5 M with respect to acetate). This mixture was incubated on ice for 5 min, and the cell debris, chromosomal DNA, and the majority of proteins were removed by centrifugation in a microcentrifuge for 5 min. To the supernatant, RNase A was added to a final concentration of 20 ug/ml, and the solution was incubated for 20 min at 37°C. The remainder of the proteins were removed by the phenol/chloroform extraction procedure. The plasmid DNA was precipitated with 2 volumes of 95% ethanol, and the mixture was kept at -70°C for a minimum of 5 min. The DNA was collected by centrifuging for 5 min in a microcentrifuge. The DNA pellet was resuspended in 16 ul distilled water and 4 ul of 4 M NaCl. The DNA was then re-precipitated with 20 ul of 13% polyethylene glycol (PEG, MW 8,000) by incubating the mixture on ice for 20 min. Again the DNA was collected by centrifugation. The DNA pellet was washed with 70% ethanol, dried under vacuum, and dissolved in 25 ul of distilled water.

2.7 Large Scale Preparation of Plasmid DNA

Bacterial cells were collected from 250-ml overnight cultures in Amp-supplemented (100 ug/ml) SOB media. The plasmid DNA was purified by following the same procedure for mini-preparation (see 2.3 above) with

proportional increases of all reagents. The DNA was used as purified or was subjected to further purification (CsCl density gradient ultracentrifugation) depending on the purpose of the experiment.

2.8 Recovery of DNA from Agarose Gel

DNA fragments to be isolated were separated electrophoretically in an 0.8% (w/v) agarose gel or in an 0.8% (w/v) low-melting agarose gel. For regular agarose gels the desired DNA fragments were recovered through electroelution (Young et al., 1979). When low-melting agarose gel was used, the gel slice containing the desired DNA fragments were excised, weighed, and melted at 70°C for 5-10 min after being mixed with two volumes (2 times the weight of the agarose gel slice) of TE buffer. The DNA was then purified and recovered through the phenol/chloroform extraction and the ethanol precipitation procedures.

2.9 Preparation of 3-7 Kb DNA EcoRI-EcoRI Fragments of Nostoc Genomic DNA

The 3-7 Kb DNA fragments were purified through a sucrose gradient (Maniatis, 1982). A 8-35% (w/v) sucrose gradient was made with a gradient maker (Bio-Rad) in a polyethylene tube (38 ml). The EcoRI-digested Nostoc genomic DNA was added to the top of the gradient and the gradient was centrifuged for 20 hours at 24,000 rpm in a Beckman SW 27 rotor at 18°C. Every 0.5 ml fraction of the gradient was collected with a fraction collector, and every third fraction was checked for its size of DNA fragments by using agarose gel electrophoresis. The fractions containing the 3-7 Kb fragments were pooled, and the sucrose was removed on a Sephadex column (NAP-25, Pharmacia) and with dialysis. The DNA was then recovered by ethanol precipitation.

2.10 Preparation of Plasmid Vector DNA

The plasmids pGEM-4 and pUC18 (1 ug each), were completely digested, respectively, with restriction enzyme EcoRI (5 units), and the fragments were de-phosphorylated with alkaline phosphatase using a standard procedure (Maniatis et al., 1982). The vector DNA was purified by the phenol/chloroform extraction and recovered by the ethanol precipitation.

2.11 Deionization of Formamide

The formamide used in this study was obtained from Sigma and was deionized as suggested by the manufacturer. Formamide (10 ml) was mixed with 1 g of ion-exchange resin [AG 502-X8 (D), Bio-Rad], and was shaken for 1 hour. The formamide was separated from the resin by centrifugation and stored at -20°C until used.

2.12 Preparation of Salmon Sperm DNA

Salmon sperm DNA (Sigma Type-III, sodium salt) was used in pre-hybridization and hybridizations during Northern analysis. The salmon sperm DNA was sheared as described by Maniatis et al. (1982). The salmon sperm DNA was dissolved in 0.1 x SSC buffer at a concentration of 10 mg/ml and the solution was passed through an 18-gauge hypodermic needle several times. The sheared DNA was boiled for 10 min and stored at -20°C in small aliquots.

3. Prepararion of DNA and RNA Probes

3.1 DNA Probe

Plasmids pPD489 and pPD490 were generous gifts of Dr. P. Dennis (Dennis et al., 1985). These two plasmids are derivatives of pBR322 and contain, respectively EcoR1-EcoR1 DNA fragments from within the rpoC (fragment size of 2.6 kb), and

rpoB (fragment size of 2.8 kb) genes of E. coli. These two inserts, after excision from the vectors, were used as probes (Fig.2) to detect the corresponding sequences in Nostoc.

Each plasmid DNA (100 ug) was digested with EcoRI (125 units) for 5 hours at 37°C. The enzyme was denatured by heating at 65 °C for 10 min. The DNA fragments were separated by electrophoresis on a 0.8% (w/v) agarose gel, and the insert DNA's were recovered by electroelution. The eluted DNAs were precipitated with ethanol.

The probe DNAs were labeled with biotin-11-dUTP, an analogue of dTTP, by nick-translation using a BRL nick-translation kit (Xie et al., 1989c). The labelling was performed by mixing 1 ug of DNA, 5 ul of solution A [dATP, dGTP and dTTP mixture, 0.2 mM each in 500 mM Tris-HCl, pH 7.8, 50 mM $MgCl_2$, 100 mM 2-mercaptoethanol, and 100 ug/ml nuclease-free bovine serum albumin (BSA)], 2.5 ul of biotin-11-dUTP (0.4 mM in 100 mM Tris-HCl, pH 7.5 and 0.1 mM EDTA), and 5 ul of solution C (0.4 U/ul DNA polymerase and 40 pg/ul DNAse I). The volume was brought to 50 ul with distilled water and the solution was incubated for 90 min at 15°C. The reaction was stopped by adding 5 ul of 300 mM EDTA. The prepared probes were stored at -20 °C until used.

3.2 RNA Ribo-probe

EcoRI-linearised pWQX039, KpnI-digested pWQX026, and PvuI-digested pWQX001 were used as templates to make ribo-probes for the rpoC2, rpoC1 and rpoB genes of N. commune, respectively. The plasmid DNAs used here were isolated by a large scale preparation without CsCl centrifugation. Each plasmid DNA (10 ug) was digested completely with the indicated restriction enzyme (20 units to 40 units), and the efficiency of cleavage was checked through agarose gel

electrophoresis. After digestion, the plasmid DNAs were treated with 1 ul of DNAse-free RNAse (10 mg/ml) for 30 min. Each template DNA was purified through standard phenol/chloroform extraction and ethanol precipitation. Finally, each template DNA was resuspended in 20 ul of diethylpyrocarbonate (DEPC)-treated distilled water and kept at 4°C until used.

The RNA probe was labeled with alpha-^{35}S-UTP (New England Nuclear, 1284 Ci/mmol, 12.5 mCi/ml, and 9.74 nmol/ml). The single strand RNA probes (anti-sense strand) were made with T7 RNA polymerase using a Promega ribo-probe kit. Template DNA (1-2 ug) was mixed with 4 ul of 5x transcription buffer (200 mM Tris-HCl, pH 7.5, 30 mM $MgCl_2$, 10 mM spermidine and 50 mM NaCl), 4 ul of mixture of ATP, CTP, GTP (2.5 mM each), and 2.6 ul of UTP (100 uM), and the volume of the mixture was dried down in a speed vac and adjusted to 8.2 ul with DEPC-treated water. The reaction was started by adding 2 ul of 1 mg/ml BSA (acetylated), 2 ul of 100 mM DTT, 0.8 ul RNAsin (Promega, 65 units/ul), 0.5 ul T7 RNA polymerase (10 units), and 6.5 ul of ^{35}S-UTP, in a total volume of 20 ul. The reaction was performed for one hour at 40°C. The labelled RNA probe was kept at -70°C until use.

4. Color Detection of Biotinylated DNA

The membrane-bound biotinylated DNA was visualized using a BRL BlueGene system. A nitrocellulose membrane (0.45 um, Schleicher & Schuell) supporting biotinylated DNA was wetted in Buffer 1, and blocked for 30 min using 3% (w/v) BSA (fraction V) in Buffer 1 at 50°C. The membrane then was treated with streptavidin-alkaline phosphatase conjugate (SA-AP, 1:1000 dilution with Buffer 1) for 10 min at room temperature. The non-specifically bound SA-AP was removed by washing with Buffer 1 for 30 min, with 2 buffer changes. The membrane was

then rinsed in Buffer 3. Finally, the membrane was placed in 5-bromo-4-chloro-3-indolyl phosphate (BCIP, 166 ug/ml) and nitro-blue tetrazolium (NBT, 330 ug/ml) solution in Buffer 3. Positive signals were seen within 10 min.

5. Southern Blotting

DNA to be blotted was separated on a 0.8% (w/v) agarose gel by electrophoresis and was transferred onto nitrocellulose membrane through a capillary reaction using a standard procedure (Maniatis et al., 1982). The DNA was denatured by soaking the agarose gel in 0.5 M NaOH and 1.5 M NaCl solution for 30 min, and the gel was neutralized with 0.5 M Tris-HCl, pH 7.5, and 1.5 M NaCl solution for 30 min with one buffer change. Nitrocellulose membrane of a desired size was cut out and wetted with 2xSSC buffer. The DNA then was transferred (overnight) onto the membrane through capillary action with 20xSSC buffer. The DNA was immobilized on the membrane by baking, for two hours, at 80°C in a vacuum oven.

6. DNA-DNA Hybridization

6.1 Hybridization

The stringency used was determined through dot blot hybridization using different hybridization and post-hybridization washing conditions. The hybridizations were performed in sealed plastic bags. Membranes were prehybridized at 42°C for 2-4 hours using a DNA hybridization buffer. The buffer was discarded after prehybridization. Biotinylated probe DNA was denatured by heating for 5 min at 95°C. The hybridization usually was performed overnight at 42°C.

6.2 Post-hybridization Washing

Non-specifically bound probe was washed off as follows: the membrane was rinsed in 1xSSC + 0.1% w/v SDS buffer for about 10 min with one buffer change. The membrane was then washed at 50°C for 90 min in 0.5xSSC + 0.1% w/v SDS with two buffer changes, and rinsed again in 1xSSC + 0.1% w/v SDS.

7. Construction of Genomic Library

The 3-7 Kb EcoR1-EcoR1 DNA fragments of Nostoc genomic DNA were used as insert DNA; dephosphorylated, EcoR1 digested lambda gt10 DNA (Promega) was used as vector DNA. The insert DNA and vector DNA were mixed in a ratio of 1:2, and were ligated with T4 DNA ligase (Maniatis et al., 1982) with the following modifications. The ligation was performed at room temperature for 2 hours in the presence of 5% $PEG_{8,000}$. The ligated DNA was packaged using packaging extracts prepared as described by Holn (1979).

8. Preparation of Packaging Extracts

E. coli strains BHB2688 and BHB2690, both carrying phage lambda as a lysogen (Holn, 1979), were used to prepare packaging extracts. The lambda genome of BHB2688 has a mutation in the gene for protein E which is essential for phage packaging; the strain BHB2690 carries a mutation in its D gene of the phage genome whose product is required for the formation of phage particles. Neither the prophage of BHB2688 nor the prophage of BHB2690 can package lambda DNA and form viable phage particles after induction of the prophage by a temperature shift from 32°C to 42°C. The packaging extracts were prepared by sonication (Maniatis et al., 1982) using a Fisher Sonic Dismembrator Model 300.

Cultures of BHB2688 and BHB2690 were grown in LB medium (100 ml), at 32°C, overnight. Fresh medium (2 x 1000 ml in two Fernbach flasks) was inoculated separately with the overnight cultures having an initial OD_{600} of 0.08-0.1. The growth of the bacteria was continued at 32°C until the OD_{600} reached 0.3, then the prophages were induced by shaking mildly at 42°C for 15 min. The induced cells were continuously incubated at 38°C for 3 hours, with mild shaking, until the bacteria were ready to lyse. The cells were collected by centrifugation (4,000g, 10 min) and were resuspended in 2.5 ml of a cold sonication buffer. The cells were lysed by sonicating at a medium power (60-65 Watts) with +short bursts until the solution of cells became clear and was no longer viscous. The solutions were chilled on ice between each burst. The cell debris was pelleted by centrifugation (10,000 rpm in SS34 rotor, 10 min at 4°C) and was discarded. To each supernatant, an equal volume of cold sonication buffer and 1/6 volume of packaging buffer were added and mixed; the mixture then was divided into 30-ul aliquots. The aliquots were frozen in liquid nitrogen immediately and kept at -70°C until needed.

9. Packaging Phage Particles

Recombinant phage DNA solution (5 ul) was added to an Eppendorf tube containing 30 ul of prohead donor (BHB2690), and mixed with a sealed capillary tube. The mixture was transferred to another Eppendorf tube containing 30 ul of protein donor (BHB2688) and was mixed again. The incubation was for 2.5 hours at 37°C.

10. Screening of Phage Genomic Library Using Biotinylated DNA Probe

E. coli strain C600hfl was used as a host for propagating phage lambda. An overnight culture of C600hfl in LB medium (100 ml) was prepared, and the cells were collected by centrifugation (4,000g, 10 min at 4°C). The pellet was resuspended in fresh LB medium to an OD_{600} of 2. This cell suspension was used during the screening procedures. For each screening, 2,000 pfu of recombinant phage were plated on one LB agar (1.5% w/v) plate (82 cm^2). The infection of phage was performed by adding the phage solution to 100 ul of the cell suspension and the mixture was incubated at 37°C for 20 min. The infected cells were poured over bottom agar after being mixed with 3 ml of top agar solution (0.7% agar in LB, 50°C). The phage particles were propagated for 12 hours at 37°C.

The phage plaques were then transferred, in situ, onto nitrocellulose membrane circles by placing the membrane on the top of each plate and lifting gently. The first, the second, and the third lifts were placed on the top of the plates for 30 sec, 2 min, and 5 min, respectively. The membranes were dried in air with phage plaques facing up. The phage DNA was released and denatured by soaking the membranes in 0.1 M NaOH/0.5 M NaCl for 40 sec, in 0.2 M Tri-HCl (pH 7.5)/1.5 M NaCl for 1 min, and in buffer 6xSSC for 1 min. The membranes were air-dried again. Finally the membranes were processed for hybridization using the biotinylated rpoC E.coli DNA probe. Possible positive plaques were subjected to secondary screening which was the same as the primary screening except that only 100 phage particles were plated on each plate.

11. Deletion Cloning

Deletion clones of inserts of plasmid pWQX001 and pWQX005 were constructed as described (Xie and Potts, 1989a), using E. coli strain DH5-alpha ($r_k^- m_k^+$) as the host cells. Four micrograms of each subcloned plasmid DNA were

double digested with restriction enzymes SphI and BamHI. The digested DNA was precipitated with ethanol and redissolved in 100 ul of exonuclease III buffer. Exonuclease III (3 ul, 65 units/ul, BRL) was added to the DNA solution, and the solution was mixed immediately and then incubated at 37°C. During the incubation, aliquots of the reaction solution (3 ul) were removed every 30 seconds and transferred to ice-chilled 0.5-ml Eppendorf tubes, each containing 2 ul of 10x exonuclease VII buffer to terminate the digestion of exonuclease III. At the end each of the Eppendorf tubes contained six aliquots (6x 3 ul, a pool) that had been removed consecutively from the reaction mixture. Thus six pools of exonuclease III digested DNA were generated. Each pool of DNA was then treated with 1 ul of exonuclease VII (1 u/ul, BRL) at 37°C for 1 hour, and then the enzymes were inactivated by heating the solutions at 70°C for 5 min. To each of the tubes, 1.5 ul of 0.2 M $MgCl_2$, 4 ul of 2 mM solution of dATP, dGTP, dCTP, and dTTP, and 2 units of Klenow fragment (Promega) were added; the mixtures were incubated at room temperature for 30 min to achieve DNA fragments with blunt ends. The DNA was then ligated with T4 DNA ligase using a standard procedure (Maniatis et al, 1982). Finally, the DNA was used to transform competent cells of E. coli DH5-alpha (Hanahan, 1985).

12. Competent Cells

Competent cells of DH5-alpha were prepared as described by Hanahan (1985).

13. Preliminary Screening of Deletion Clones

Deletion clones of pWQX001 and pWQX005 were screened without any DNA purification steps as described (Xie and Potts, 1989a). Using sterile toothpicks, transformants derived from each separate pool of the treated DNA were transferred one by one from master plates and streaked on SOB-agar (Amp, 100

ug/ml) plates. The plates were incubated at 37°C overnight. Plates were used directly for further analysis or were kept at 4°C for up to three weeks before use. About 5-10% of the cell mass from each streak was removed using sterile toothpicks and transferred to a 0.5-ml Eppendorf tube by mixing vigorously in 10 ul of protoplasting buffer. The buffer became turbid immediately upon mixing. The suspensions were incubated at room temperature for 20-30 minutes. The protoplast suspensions and DNA molecular size markers were loaded into gel slots after preloading of each well with 5 ul of lysis buffer. The samples were then analyzed by agarose gel electrophoresis. The agarose gel contained 0.05% (w/v) SDS.

14. Sequencing of DNA

The selected DNA templates were sequenced using a Promega sequencing kit with ^{35}S-dATP (>400 Ci/mmol; Du Pont-New England Nuclear) and the Klenow fragment of DNA polymerase I, following the dideoxy chain-termination method (Sanger et al., 1977).

15. Drying And Rewetting Nostoc Cells

Cells of Nostoc (about 20 g wet cells) in exponential growth were collected as described earlier. Cells (5 g) were used for RNA isolation and the remainder of the cells were divided into about 0.7-g portions, which were spread evenly on nylon mesh supports. For the drying process, aliquots of cells were air-dried for 1 day and 5 days, respectively. For the rewetting process, aliquots of cells, after 5 days of drying, were rewetted, respectively, for 5 min, 10 min, 30 min, 1 hr, 5 hr, 24 hr, and 4 days by adding BG-11 media to them. The cells then were processed for isolation of total RNA or total protein.

16. RNA Preparations

16.1 Isolation of Total RNA from Fresh Nostoc Cells

Cells of Nostoc (wet weight 5 g) in exponential growth were used to purify total RNA. The cells were frozen immediately using liquid nitrogen and pulverized into a fine powder using a pre-chilled mortar and pestle. The powder was transferred to a 50-ml glass tube and the tube was kept on ice. A solution (1.5 ml) of 1 M NaOAc/1%SDS (pH 5.2) was added and the mixture was vortexed at top speed for three 2 min intervals. The tube was chilled on ice between each vortexing. Five ml of cold buffered phenol and 200 ul of 2-mercaptoethanol were added and the suspension was mixed by vortexing at medium speed for a few minutes. During these processes the cells were partially frozen. Nine ml of 1 M NaOAc/1%SDS buffer was added to dilute the solution, and the mixture then was incubated at 50°C for about 10 min. Five ml of chloroform/isoamyl alcohol (24:1) was added and the solution was mixed by vortexing at a low speed for 30 seconds. The mixture then was transferred to a 40-ml polyallomer tube and the two phases were separated by centrifugation (10,000 rpm, SS34 rotor, 10 min at 4°C). The aqueous phase was extracted again with phenol/chloroform mixture until the interface was clean, and with chloroform/isoamyl alcohol. Nucleic acids were precipitated with ethanol. The pellet was dissolved in 1.7 ml of diethylpyrocarbonate (DEPC, Sigma)-treated water, and the RNA from a 1.5-ml solution was precipitated with 0.5 ml 8 M LiCl at 4°C overnight. The RNA was collected by centrifugation at 9,000 rpm for 50 min in a SS34 rotor at 4°C. The pellet was redissolved in 2 ml of water and divided into 0.5 ml portions. RNA from each of the aliquots was precipitated and washed with ethanol. The RNA aliquots were stored in 70% ethanol at -70°C as they were, or were further treated to

remove any DNA. The pellet was dried in a speed vac and was dissolved in 30 ul of RQ DNase buffer (40 mM Tris-HCl, pH 7.9, 10 mM NaCl and 6 mM $MgCl_2$. The RNA sample was then treated with 0.5 unit of RQ DNase (Promega) at 37°C for 30 min.

16.2 Isolation of Total RNA from Dried and Rewetted Nostoc Cells

Cells from each preparation were frozen in liquid nitrogen, ground to a fine powder, and transferred to 15-ml tubes. Buffered phenol (1.5 ml) and 1.5 ml of extraction buffer were added immediately to each of the tubes. The mixture was sonicated at a power level of 30 for 3x15 seconds, with cooling between each burst. The mixture then was vortexed at a top speed for about 20 min. The two phases were separated by centrifugation, and the aqueous phase was extracted again with phenol/chloroform (three times) and with chloroform/isoamyl alcohol (once). The nucleic acids were precipitated and washed with ethanol. The pellet was dissolved in 100 ul water and RNA was precipitated with LiCl as described above. Finally the RNA was dissolved in 20 ul of DEPC-treated water and was treated with RQ DNase (as above) and the treated RNA was kept at -70°C until used.

17. Northern Blotting of RNA

17.1 Electrophoresis of RNA under Denaturation Condition by Formaldehyde

The RNA samples (10 ug each) and RNA markers (3 ug, BRL) were denatured and separated on a 1% w/v agarose-formaldehyde gel essentially by following the procedure of Fourney et al. (1988) with some modifications. The modifications were that no EtBr was added to RNA samples before loading them on the gel and that the gel was only rinsed in water for a few min after electrophoresis. The RNA

was then transferred onto Nytran membrane (0.2 um, Schleicher & Schuell) through capillary action (Maniatis et al., 1982). The RNA samples were immobilized on the membrane by baking at 80°C for 1-2 hour in a vacuum oven.

17.2 Electrophoresis of RNA under Denaturation Condition by Glyoxal

The RNA samples (10 ug each) and RNA markers (3 ug, BRL) were denatured with glyoxal and separated on a 1% (w/v) agarose gel as described (Williams and Mason, 1985). The separated RNA's were then transferred and immobilized onto a Nytran membrane as before.

18. RNA-RNA Hybridization

18.1 Prehybridization and Hybridization

Northern-transferred membranes were prehybridized in a plastic bag with RNA hybridization buffer at 50°C for 4-5 hours. The prehybridization buffer was discarded and the bag was re-filled with 1-3 ml of fresh buffer depending on the size of the membrane. The ^{35}S labelled ribo-probe (3-8 ul) was added to the bag and the hybridization was performed at 50°C, overnight, with gentle shaking.

18.2 Post Washing

The hybridized membrane was rinsed in 1xSSC/0.1% SDS buffer for a few minutes with one buffer change. The filter was washed in 0.1xSSC/0.1% SDS buffer at 60°C for 50 min with two buffer changes. Finally, the membrane was rinsed again with 1xSSC/0.1% SDS buffer. The membrane was then air-dried.

18.3 Detection of RNA-RNA Hybrid

RNA-RNA hybrids were detected through autoradiography.

19. Protein Preparation

19.1 Isolation of Total Soluble Proteins from E. coli

Total soluble proteins were isolated from untransformed E. coli strain JM109 cells and JM109 cells harboring plasmid pX1 (Fig.19) through an SDS boiling method (Silhavy et al., 1984), using a cracking buffer.

Cells were collected from 1.5-ml cell culture by centrifugation and the pellet was resuspended in 100 ul of cracking buffer. The suspended cells were boiled for 5 min and kept at -20°C until used.

19.2 Isolation of Total Proteins from Fresh Nostoc Cells

Three grams of wet Nostoc cells were frozen in liquid nitrogen, ground into a powder in a chilled mortar and pestle, and transferred to a 15-ml tube. Six ml of cracking buffer were added to the powder and the mixture was sonicated for 3x30 seconds at a power level of 30 Watts using the sonicator. The cell suspension was cooled on ice between each burst. The breakage of the cells was monitored using a microscope or by illuminating one drop of the cell suspension on a UV box. The release of phycobiliproteins from the cell gave an orange fluorescence upon cell lysis. The suspension became clear as the cells broke. The cell debris was removed by ultracentrifugation at 35,000 rpm in a Beckman Ti50 rotor for 1 hour at 4°C. The supernatant was then aliquoted into small portions and kept at -70°C until used.

19.3 Isolation of Total Proteins from Dried And Rewetted Nostoc Cells

The same procedure as described above was used to isolate total proteins from the treated cells, except that only 0.7 g of cells were used.

20. Expression of Gamma And Beta' (truncated) Subunits of RNA Polymerase of Nostoc in E. coli Cells

The 6.5-kb fragment containing the rpoBC1C2 genes (rpoB and rpoC2 genes are incomplete) of Nostoc was subcloned in an expression vector pUC18 at the EcoR1 site. The plasmid DNA was then used to transform competent cells of E. coli JM109. A colony of transformed cells was picked and grown overnight in 5 ml SOB/Amp (100 ug/ml) medium. The next day two 50-ml cultures of SOB/Amp (100 ug/ml) medium were inoculated, respectively, with 500 ul of the overnight culture and incubated at 37°C with shaking. At the same time, 50-ml of SOB medium was inoculated with stock JM109 cells, which was used as the control sample. After the OD_{600} of the culture reached 0.5, the expression of the Rpo proteins in one culture was induced using isopropylthiogalactoside (IPTG, final concentration of 10 mM). The cultures were continously incubated at 37°C. Portions (1.5 ml) of the cultures were removed from each flask every 30 min after induction, and total soluble proteins were isolated as described above.

21. In vitro Expression of Gamma And Beta' Subunits of RNA Polymerase of Nostoc

The N. commune rpoBC1C2 genes were translated in vitro using a cell-free coupled transcription-translation system obtained from Du Pont-New England Nuclear, following the manufacturer's suggestions. Proteins were synthesized in the presence of ^{35}S-methionine (NEN, 1134 Ci/mmol, 9.07 mCi/ml). Plasmid pWQX035 (1.5 ug) was used as DNA template in this experiment. The plasmid was purified by using large-scale preparation and further purified by two rounds of cesium chloride ultracentrifugation (see Section 2.2). About 60 uCi of

^{35}S-methionine was used in each sample. Two controls were performed with the same condition, one without the addition of plasmid DNA, and the other with unlabelled methionine.

The plasmid DNA (1.5 ug in 1 ul) or 1 ul of water was mixed sequentially with 6 ul of cocktail B, 6.5 ul of ^{35}-S(9 uCi/ul) or non-labeled methionine, and 6 ul of S-30 extract solution. The reagents were mixed by pipetting up and down for a few times. The reactions were performed for 1 hour at 37 °C.

The incorporation of ^{35}S in the sample where plasmid was used, was assayed as follows: 1 ul of the reaction solution was removed from the sample and spotted on a circle of Whatman grade 3 filter paper. The circle was air-dried for 10 min. The proteins were precipitated onto the membrane by placing the filter into a boiling solution containing 10% trichloride acetic acid (TCA) and 1 mg/ml casein, for 10 min. The circle was then washed three times using a 5% (w/v) TCA solution, each for about 10 min. Radioactivity on the sample filter was measured in a scintillation counter.

22. SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-Polyacrylamide gels were prepared by a standard procedure (Ausubel et al., 1987) using a mini-apparatus (Bio-Rad). The separation gel used was 8% (w/v) acrylamide, pH 8.8, and the stacking gel was 4% (w/v) acrylamide, pH 6.8. Electrophoresis was performed at 20 mA per gel until the dye front reached the bottom of the gel. Protein samples were heated at 50°C for 10 min before application to the gels.

23. Immunoanalysis

Antibodies used in this study were a generous gift of Dr. Robert Haselkorn (Dept. of Cellular Biology and Human Genetics, University of Chicago). The antibodies (rabbit IgG) were directed against the core enzyme of the RNA polymerase of Anabeana PCC7120.

23.1 Western Blotting

After SDS-PAGE, the gel was soaked in transfer buffer for 10 min, and proteins on the gel were transferred by electrophoresis to Immobilon membranes (Millipore). A piece of Immobilon membrane of desired size was placed for a few seconds in 100 % methanol and was transferred to water for 10 min. The membrane was then equilibrated with transfer buffer for at least 10 min before use. The transfer was performed for two hours at 24 volts using a GENIE transfer apparatus (Idea Scientific Company). About 1.8 liters of transfer buffer were used for each transfer. The buffer contained 25 mM Tris-base, 192 mM glycine, 20% (v/v) methanol, and 0.01% (w/v) SDS, and was de-gassed for 15-20 min using a aspirator before the SDS was added.

23.2 Antigen and Antibody Interaction

Western-blotted membranes were blocked by treatment with 1% (w/v) BSA made up with TTBS buffer (TBS: 20 mM Tris-HCl, pH 7.5, and 500 mM NaCl; TTBS: TBS + 0.05% Tween-20) for 30 min at room temperature. The membrane was then treated with a dilute antibody solution (1:2000 dilution with TTBS) for 4 hours at room temperature. The non-specific binding of antibodies was removed by washing with TTBS buffer three times, at room temperature, each wash for about 5-10 min.

23.3 Detection of Antigen and Antiboby Conjugate

a) Detection by Protein A-Horseradish Peroxidase (HRP)

After the non-specifically bound antibodies were removed, the membrane was treated for 1 hour with diluted protein A-HRP conjugate (Bio-Rad, 1:10,000 dilution with antibody buffer (1% gelatin in TTBS)) with gentle agitation. The non-specifically bound conjugate was removed as described above. Just before use, 60 mg of HRP color development reagent (Bio-Rad) was dissolved in 20 ml ice-cold methanol (A), and 60 ul of cold 30% hydrogen peroxide (H_2O_2) were added to 100 ml room temperature TBS buffer (B), and A and B were mixed gently. The membrane then was placed in this solution, and the reaction was performed for up to 40 min in the dark with gentle agitation.

b) Detection by Anti-IgG-Alkaline Phosphatase

After washing, the membrane was treated with diluted anti(rabbit)-IgG-alkaline phosphatase (1 to 2500 dilution with TTBS) for 30 min. The blot was washed with TTBS buffer three times for 10 min each. Bound enzyme complex was visualized as described earlier for biotinylated DNA.

23.4 Immunoprecipitation

The in vitro translated proteins (20 ul) were mixed with 1 ul diluted antibody (1 to 1,000 dilution with antibody buffer) and the mixture was incubated for 15 min. Diluted Horseradish Peroxidase (HRP)-protein A conjugate (3 ul, 1 to 7500, with TTBS) was added, the solution was mixed, and then incubated for 10 min. The treated solution was centrifuged for 10 min at 4°C in a microcentrifuge. The supernatant was transferred to another 0.5-ml tube, and the pellet was resuspended in 20 ul of cracking buffer. The proteins in the supernatant were precipitated with acetone.

To the supernatant an equal volume of ice-cold acetone was added, the solution was mixed, and then incubated on ice for 10 min. The proteins were then pelleted by centrifugation for 10 min in a microcentrifuge at 4°C. The pellet was resuspended in 40 ul of cracking buffer.

23.5 Visualization of Proteins Immobilized on Immobilon Membrane

The immobilized protein samples and size markers were visualized using Amido-black (Sigma). The membrane to be stained was wetted for a few seconds in 100% methanol and then placed in diluted Amido-black solution (1:10 dilution). The dilution was made as follows: acetic acid (7%) and ethanol (20%). The staining was stopped by rinsing several times with water.

RESULTS

Isolation of rpoBC1C2 of Nostoc commune

Southern analysis of EcoRI-digested Nostoc genomic DNA with biotinylated probe of a rpoC of E. coli detected a single fragment of about 6.5 kb (Fig. 3). This positive result justified the use of this probe for the isolation of the corresponding genes in the Nostoc commune UTEX 584 library.

A phage genomic library was constructed using 3-7-kb EcoRI-EcoRI fragments of Nostoc DNA as insert and lambda gt10 as the vector. The library was screened using the same probe. After screening about 40,000 recombinant phage, a phage clone was isolated from a positive-hybridizing plaque after secondary screening (Fig. 4). The positive plaques showed a reaction 2 min after being placed in the color detection system. The clone (lambda gt10X3, Fig. 5) was further confirmed through Southern analysis, using the same probe after the DNA was purified and linearized with EcoRI (data not shown). The cloned DNA also hybridized with a biotinylated probe of E. coli rpoB (Fig. 11C). This indicated that the cloned DNA contained both rpoBC genes of Nostoc commune.

Sequencing of the rpoBC1C2 genes of Nostoc commune

In order to sequence the cloned DNA fragment, the insert of lambda gt10X3 was subcloned into pGEM-4 (Promega) at its EcoRI site in both orientations, and the resulting plasmid DNA's were named pWQX001 and pWQX005, respectively (Fig. 6, 8, and 9). Restriction enzyme mapping of the insert was conducted using restriction enzymes which cut the vector DNA only once at its multiple cloning site (Fig. 7). These enzymes were HindIII, SphI, PstI, SalI, AccI, HincII, XbaI, BamHI,

AvaI, SmaI, KpnI, and SacI. Among them BamHI, SacI, AvaI, XbaI, PstI, SalI, and SphI did not cut the cloned fragment. The map is presented in Figure 8 and Figure 9. A complete restriction enzyme map is presented in Appendix 3.

Deletion clones of pWQX001 and pWQX005 were constructed as described in the methods section. Exonuclease III can only digest double-stranded DNA with ends of 3'-underhang. BamHI digestion generated ends with 3'-underhang; SphI digestion generated ends with 3'-overhang. Therefore, these two plasmid DNAs, after being double-digested with these two enzymes, could only be digested counter-clockwise by the exonuclease III (Fig. 8 and 9). The generated single strand DNA was removed with exonuclease VII. The use of Klenow fragment here was to insure that the ends of the DNA fragments, after exonuclease digestion, were blunt-ended. The colonies carrying plasmid DNA with different deletion inserts were screened in a preliminary test without purification steps using agarose gel electrophoresis (Fig. 10). HindIII-digested lambda DNA, a colony carrying plasmid DNA without deletion (pWQX001), and a colony carrying plasmid DNA without insert DNA (pGEM-4) were used as markers during electrophoresis. About 100 colonies were selected through this procedure and their plasmid DNA were purified through a mini-procedure. The size of each plasmid DNA was estimated more accurately after the DNA was purified and linearized with EcoRI, or EcoRI and HindIII. One ul out of 25 ul of the purified DNA was used for this purpose.

Deletion derivatives of pWQX001: pWQX035, pWQX031, pWQX025, pWQX022, pWQX015, and pWQX012, contained insert DNA of sizes of 6297 bp, 4500 bp, 4300 bp, 2686 bp, 2218 bp, and 1238 bp, respectively. These clones were analyzed by using Southern analysis, after digestion with restriction enzyme

HindIII and EcoRI (Fig. 11). The probes used here were biotinylated rpoB and rpoC of E. coli. Probe rpoB only hybridized with DNA from clones without a deletion (pWQX001) or with a small deletion (pWQX035) (Fig. 11C); probe rpoC hybridized with all DNA except DNA from pWQX012 (Fig. 11D). This indicated that the deletion occurred first within the rpoB gene of pWQX001, then extended to the rpoC genes of Noctoc commune. The HindIII and EcoRI double digestion of pWQX001 generated four fragments with sizes of about 5350, 700, 1050, and 2875 bp. These four fragments were arranged in the same order as listed here. The unidirectional deletion of this plasmid DNA first occurred on the DNA fragment of 5350 bp, then on the 700 bp fragment, and last on the 1050 bp fragment (Fig. 11A and 11B). The last fragment is from the vector DNA. Therefore, the orientations of insert DNA in plasmid pWQX001 and pWQX005 are as shown in Fiqures 7 and 8. Sequence analysis was completed using 52 overlapping deletion clones (Fig. 12) , which covered almost complete two strands of the 6.5-kb fragment. The sequence is shown in Fiqures 13 and 15.

Analysis of sequence data of rpoBC1C2

Analysis of the sequence through the IBI DNA/protein software system revealed three open reading frames (ORF). The first ORF (incomplete) had a potential to code for a polypeptide of M_r value of 88,000 (C-terminal); the second ORF has a potential to encode a polypeptide of M_r value of 70,200 (complete); the third one (incomplete) has a potential for a polypeptide of M_r value of 86,000 (N-terminal). The three possible coding regions are separated by relatively short AT rich sequences (Figs. 13, 15, and 16). The derived amino acid sequence from the first coding region (2217 bp, incomplete) showed extensive sequence similarity with that of the Beta subunit of E. coli (Fig. 17). The second coding region of 622

codons (1866 bp, complete) is homologous to that of the N-terminal portion (from residues 1 to 600) of the Beta' subunit of RNA polymerase of E. coli (Fig. 18). The derived amino acid sequence of the third coding region (2151 bp, incomplete) shows sequence similarity with that of C-terminal (from residue 615 to 1110) of beta' subunit of the enzyme (Fig. 19). The three coding regions are referred to as rpoB, rpoC1, and rpoC2, respectively.

Three distinct redundancies were detected in the rpoB gene of N. commune. The redundancies were apparent in both the DNA and deduced amino acid sequence (Fig. 14). The intergenic regions (about 52 bp) separating rpoB and rpoC1, and rpoC1 and rpoC2, of Nostoc commune, are AT rich sequences and have extensive sequence similarity (Fig. 15). Both of these sequences have the ability to form stable hair-pin like structures (Figs. 20 and 21) with ΔG values of -43.5 and -36.7 Kcal/mol respectively. Several distinct components were recognized in the intergenic sequences of the N. commune genes through visual alignment (Fig. 15). The components (TTAG repeat, TTAATT and CAAAC sequences) were separated at equivalent distances upstream from the presumed translational initiation codons of rpoC1 and rpoC2. Potential ribosomal binding sites (GGA and AGGA) were located 15 and 10 bases upstream from the translation initiation codons of rpoC1 and rpoC2, respectively.

Expression of Nostoc rpo genes in-vivo

The insert of lambda gt10X3 was subcloned into an expression vector pUC18 at its EcoRI site and was named pX1(Fig. 22). The cloned rpoBC1C2 fragment was joined after the lacZ promoter, out of frame with the lacZ gene. The plasmid then was used to transform competent cells of JM109. Total proteins were isolated from the transformed cells, with or without IPTG induction, and from

non-transformed cells (control). For the transformed JM109 cells, IPTG induction should induce the transcription of a fusion transcript of lacZ, interrupted by the rpo mRNA, using the lacZ promoter. The protein samples, together with size markers, were separated on SDS-PAGE (8% w/v), and visualized through Coomassie staining. There was no difference in the banding patterns between the sample isolated from induced transformed cells and the sample isolated from non-induced transformed cells (data not shown). The expression of the rpo genes was checked by SDS-PAGE, Western blot and immunoanalysis. The binding of antibodies to the antigen was detected using protein A - HRP and its color detection reagent.

Protein samples were transferred onto a piece of Immobilon membrane (Millipore). The transfer of proteins onto the membrane was checked by staining the transferred membrane using Amido-black (Sigma) (Fig. 25B).

Immunoanalysis of a Western blot of the total proteins isolated from the transformed cells, revealed two protein bands of 94,000 M_r and 71,000 M_r, which showed cross reaction with the antibodies directed against the core proteins of the RNA polymerase of Anabeana PCC7120 (Fig. 23, lane 5). These two bands were absent in a control protein sample which was isolated from untransformed JM 109 cells (Fig. 23, lane 6).

Expression of Nostoc rpo genes in vitro

The rpoBC1C2 fragment was cloned behind a T7 promoter, and a polycistronic transcript was made with T7 RNA polymerase. The transcript was further translated using its own ribosomal binding sites. The in vitro expression of the rpo genes was performed using an E. coli cell-free transcription-translation system in the presence of ^{35}S-methionine. Plasmid pWQX035 was used as template DNA,

which was a deletion derivative of pWQX001, containing a deletion of 153 bp at the 5' site (Fig. 12). Control experiments were conducted one without plasmid DNA and one with non-labelled methionine. The products of the expression were analyzed through SDS-PAGE, immunoprecipitation, Western blotting and autoradiography. Autoradiographic analysis of the translation products detected a strong band, corresponding to the polypeptide with a M_r of 94,000, and minor bands with lower M_r values (Fig. 24B, lanes 5 and 6; Fig. 23B, lane 3). All these bands of ^{35}S-labeled polypeptides were absent in the control assay where DNA was omitted from the reaction (Fig. 23B, lane 2), and all of them could be precipitated with the antiserum (Fig. 24A and B, lane 5). Of the bands detected after SDS-PAGE and Coomassie staining of the immunoprecipitate, two polypeptides (M_r of 94,000 and 71,000, respectively) precipitated, and they were ^{35}S-labeled and co-migrated with the two polypeptides produced by in-vivo expression (Fig. 24A and B, lane 5). The immunoprecipitated protein sample had the same autoradiographic pattern as the protein sample without immuno-precipitation (Fig. 24B, lane 5 and lane 6). The in-vitro translated polypeptides, however, showed almost non-detectable cross reaction with the antiserum (Fig. 23, lanes 3 and 4). The major proteins present in the translation system were not labeled and were not precipitated by the antiserum, and, as a consequence, remained in the supernatant after immunoprecipitation (Fig. 24, lane 4).

Total proteins isolated from fresh Nostoc cells were used as a control for SDS-PAGE, Western blotting, and immunoanalysis of the samples from the in vivo and in vitro expression experiments. Only the alpha subunit (M_r of 43,000) of RNA polymerase of Nostoc showed a strong cross reaction with the antiserum (Fig. 23A,

lane 1). The gamma subunit of Nostoc showed very weak reaction with the antibodies and the beta and Beta' subunits showed weak reactions with the antiserum. A band of about 180 kd was present in all in vitro translation samples with or without exogenous plasmid DNA (Fig. 23A, lanes 2-4; Fig. 24A, lanes 4-6).

Immunoanalysis of the total protein samples isolated from dried and rewetted N. commune cells

Total proteins were isolated from differently-treated N. commune cells. The treatments were 1 day dry; 5 days dry; and 10 min, 30 min, 1 hour, 5 hour, 24 hour, and four days rewetting, respectively. The rewetted cells were dried for 5 days before being subjected to rewetting. Total proteins isolated from fresh cells was used as a control. These protein samples were analysed by SDS-PAGE (8% w/v), Western blotting and immuno-reaction (Fig. 25A and 26). The antigen-antibody complexes were detected using anti-IgG alkaline phosphatase. Approximately the same amount of proteins of each sample was used in this assay.

Bands of polypeptide (M_r of about 43,000), corresponding to the alpha subunit of RNA polymerase, were present in all samples in almost equal amounts (Fig. 26). A polypeptide (M_r of about 30,000) was present in all samples except in the protein extract from fresh cells (Fig. 26, lane 9).

Northern blotting analysis

RNA isolated from fresh Nostoc cells, before and after LiCl precipitation, was separated by agarose gel electrophoresis and was visualized using EtBr (Fig. 27). The LiCl precipitation of RNA removed much of the DNA.

EcoRI-linearized plasmid pWQX039, KpnI-digested pWQX026, and PvuI-digested pWQX001 were used as templates for making ribo-probes for the rpoC2, the rpoC1, and the rpoB genes. The anti-sense strand RNA probes were

made using T7 RNA polymerase. Using pWQX039 as a template, a ribo-probe was made from nucleotide 5578 to 6456 within the rpoC2 gene. Using KnpI-digested pWQX026 as a template, a ribo-probe was made from nucleotide 2733 to 4084 within the rpoC1 gene. The third template gave a ribo-probe from nucleotide 1 to 1985, within the rpoB gene. The digestions of template DNA with different restriction enzymes were to provide run-off sites for the polymerase during the transcription of ribo-probes.

The transcription analysis of rpoB, rpoC1, and rpoC2 was conducted using a RNA sample isolated from fresh Nostoc cells. RNA samples (10 ug each lane) and RNA size-markers were denatured by the formaldehyde method, separated on a formaldehyde-agarose (1% w/v) gel, and were blotted onto a piece of Nytran membrane. The membrane was cut into strips, each containing one RNA sample.

The rpoC2 ribo-probe hybridized with a single fragment of about 5.6 kb (Fig. 28A); the rpoC1 ribo-probe hybridized with the same fragment (Fig. 28B), but the rpoB ribo-probe hybridized with a single fragment of about 3.1 kb(Fig. 28C and D).

Northern analysis of RNA samples isolated from fresh, dried, and rewetted N. commune cells

Total RNA was isolated from differently-treated N. commune cells. The treatments were drying (1 day and 5 days) and rewetting (5 min, 10 min, 30 min, and 1 hour). The rewetting process was conducted with cells which had been dried for 5 days. RNA isolated from fresh cells was used as a control. About 10 ug of total RNA from each sample was used to analyze their contents for rpo mRNA. The RNA samples used here were denatured by the glyoxyl method. The samples were analyzed by using agarose gel eletrophoresis, Northern blotting, RNA-RNA

hybridization, and autoradiography. The probe used was ^{35}S-labeled ribo-probe for the rpoC2 gene. The hybridized membrane was exposed to Kodak film for 82 hours.

There were no detectable amounts of rpo mRNA in the samples isolated from dried and 5-min rewetted cells (Fig. 29, lanes 1-3). The mRNA re-appeared and gradually increased in amount in samples isolated from cells which had been rewetted for more than 5 min (Fig. 29, lane 4-7). The amount of rpo mRNA present in each sample was estimated by scanning the photograph (Fig. 29) using a densitometer (Shimadzu CS9000 U, reflection scanning). Using RNA isolated from fresh cells as the standard, the 10-min, 30-min, and 60-min rewetted cells contained relatively 28%, 53% and 66% of the amount of mRNA as fresh cells, respectively.

Figure 1. **Comparasion of Gene Structure and Organization of rpoBC Genes Between Archaebacteria, Eukaryotes (RNA polymerase II), Eubacteria, Cyanobacteria, and Chloroplasts.**

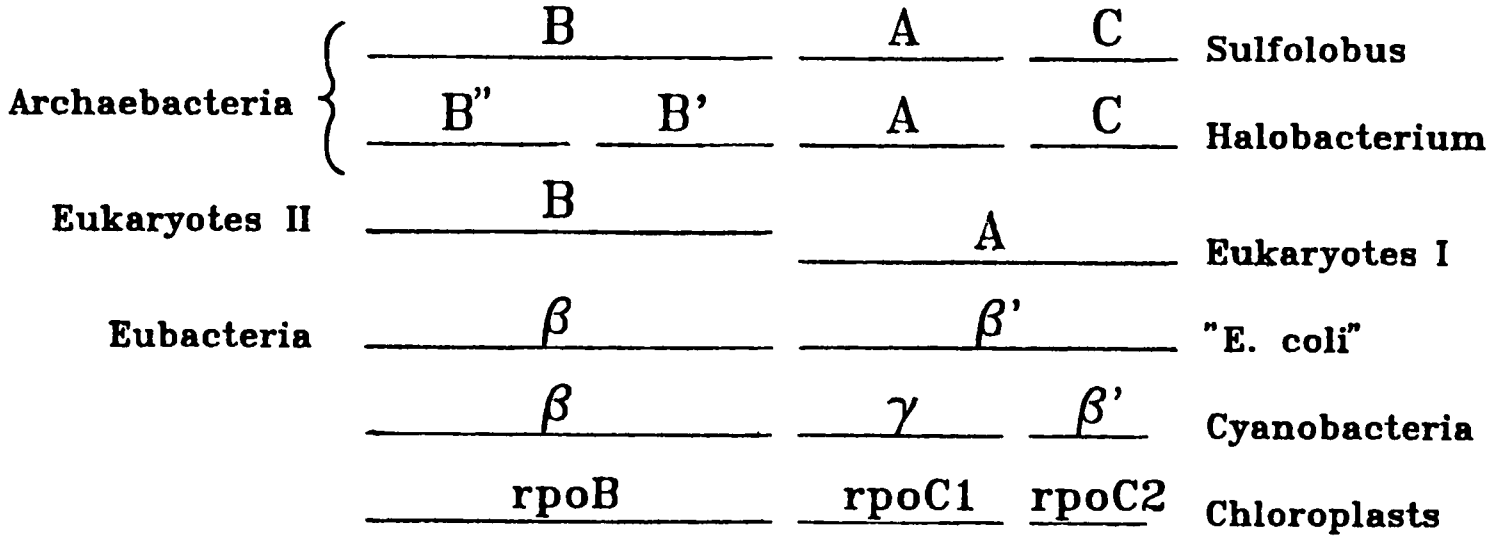

Archaebacteria
B
A
C
Sulfolobus
B"
B'
A
C
Halobacterium
Eukaryotes II
B
A
Eukaryotes I
Eubacteria
β
β'
"E. coli"
β
γ
β'
Cyanobacteria
rpoB
rpoC1
rpoC2
Chloroplasts

Figure 2. **DNA Probes of rpoB and rpoC of E. coli.**

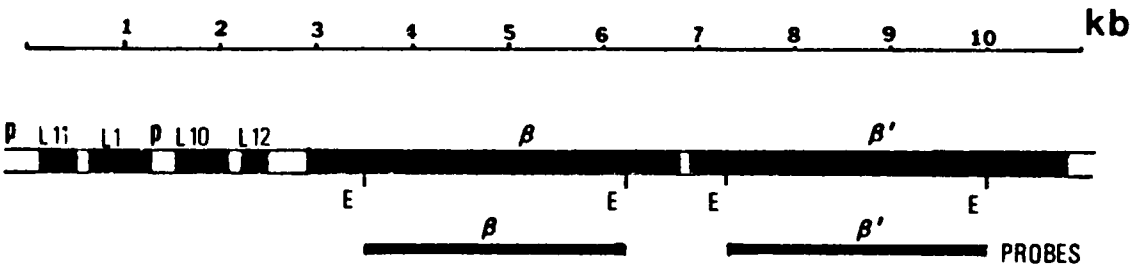
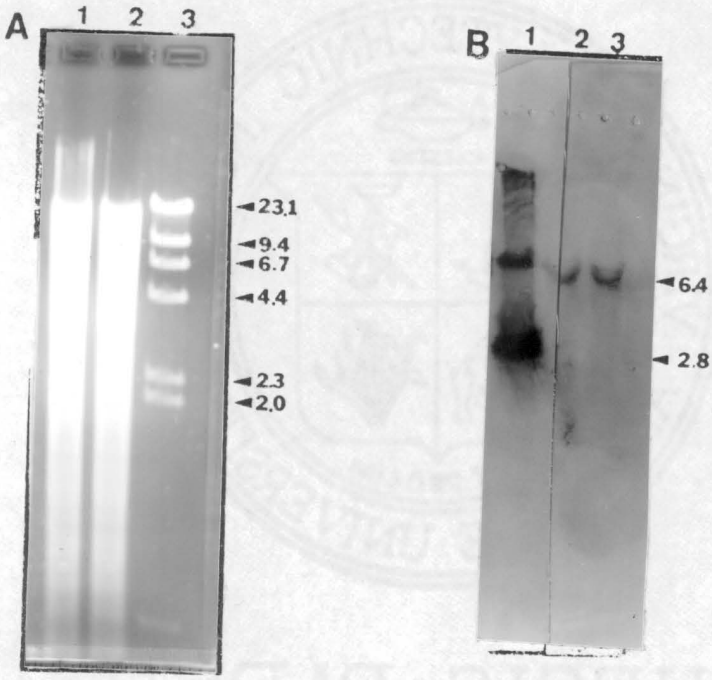

1
2
3
4
5
6
7
8
9
10
kb
p
L11
L1
p
L10
L12
β
β'
E
E
E
E
β
β'
PROBES

Figure 3. **Southern Blot Hybridization of rpoC Probe DNA of E. coli to Nostoc commune UTEX 584 Genomic.** The genomic DNA was digested with restriction enzyme EcoRI. A: Agarose gel electrophoresis: lane 1, EcoRI digested N. commune DNA; lane 2, Same as lane 1; lane 3, lambda DNA digested with restriction enzyme HindIII as size maker; B: lane 1, pND489 as positive control; lane 2, N. commune DNA; lane 3, same as lane 2.


A
1
2
3
23.1
9.4
6.7
4.4
2.3
2.0
B
1
2
3
6.4
2.8

Fiqure 4. **Screening Lambda Phage Genomic Library of N. commune UTEX584 DNA with rpoC Probe of E. coli (secondary screening)**: the dark circles were positive signals; the shadow circles were negative signals.

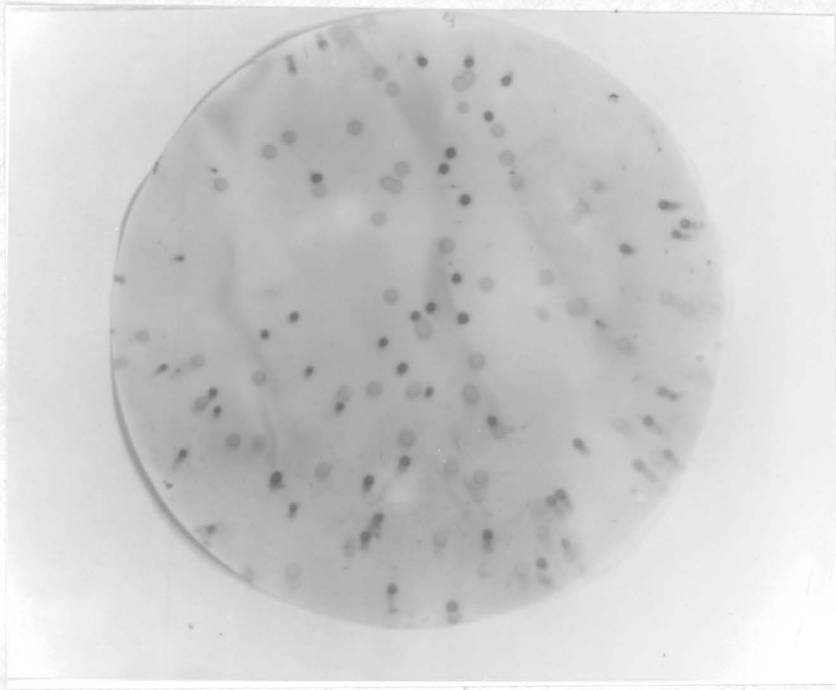
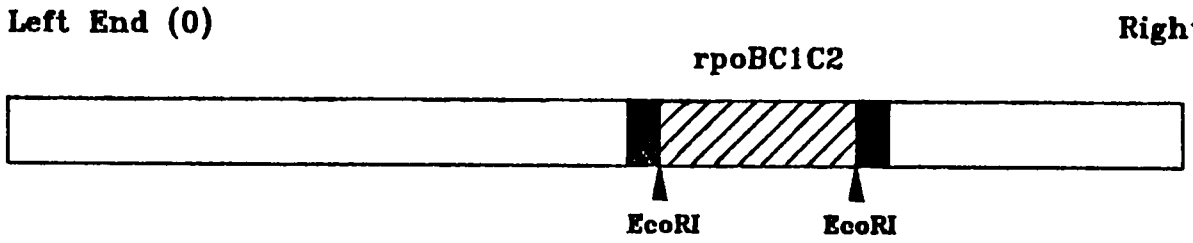


Figure 5. **Lambda Phage Subclone Lambda gt10X3**



λgt10X3 (λgt10+rpoBC1C2)

■ cI gene

Figure 6. **Agarose Gel Electrophoresis.** Plasmid DNA subclones pWQX001 (lane 2), pWQX002 (lane 3), pWQX003 (lane 4), pWQX004 (lane 5), and pWQX005 (lane 6) digested with restriction enzyme HindIII. Lane 1, lambda phage DNA digested with restriction enzyme HindIII (size-markers).

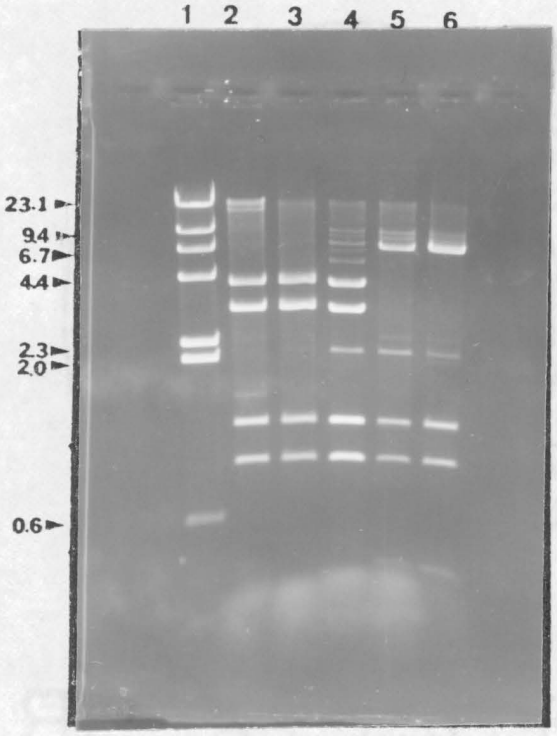

1 2 3 4 5 6
23.1
9.4
6.7
4.4
2.3
2.0
0.6

Figure 7. **Agarose Gel Electrophoresis: Restriction Digestion of pWQX001 and pWQX005**: Lanes 1, 14 and 26, HindIII-lambda DNA markers; lanes 2-13 are pWQX005 digested, respectively, with restriction enzymes HindIII, KpnI, SacI, SmaI, AvaI, BamHI, XbaI, HincII, AccI, SalI, PstI and SphI. Lanes 15-25 are pWQX001 digested, respectively, with restriction enzymes KpnI, SacI, SmaI, AvaI, BamHI, XabI, HincII, AccI, SalI, PstI and SphI.

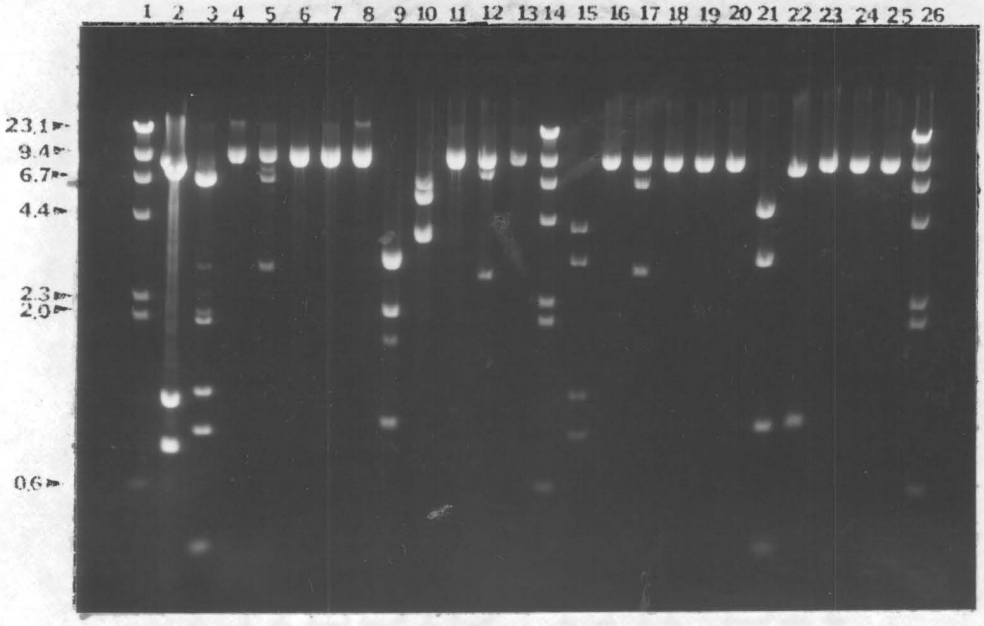

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
23.1
9.4
6.7
4.4
2.3
2.0
0.6

Figure 8. **Restriction Enzyme Map of pWQX001.**

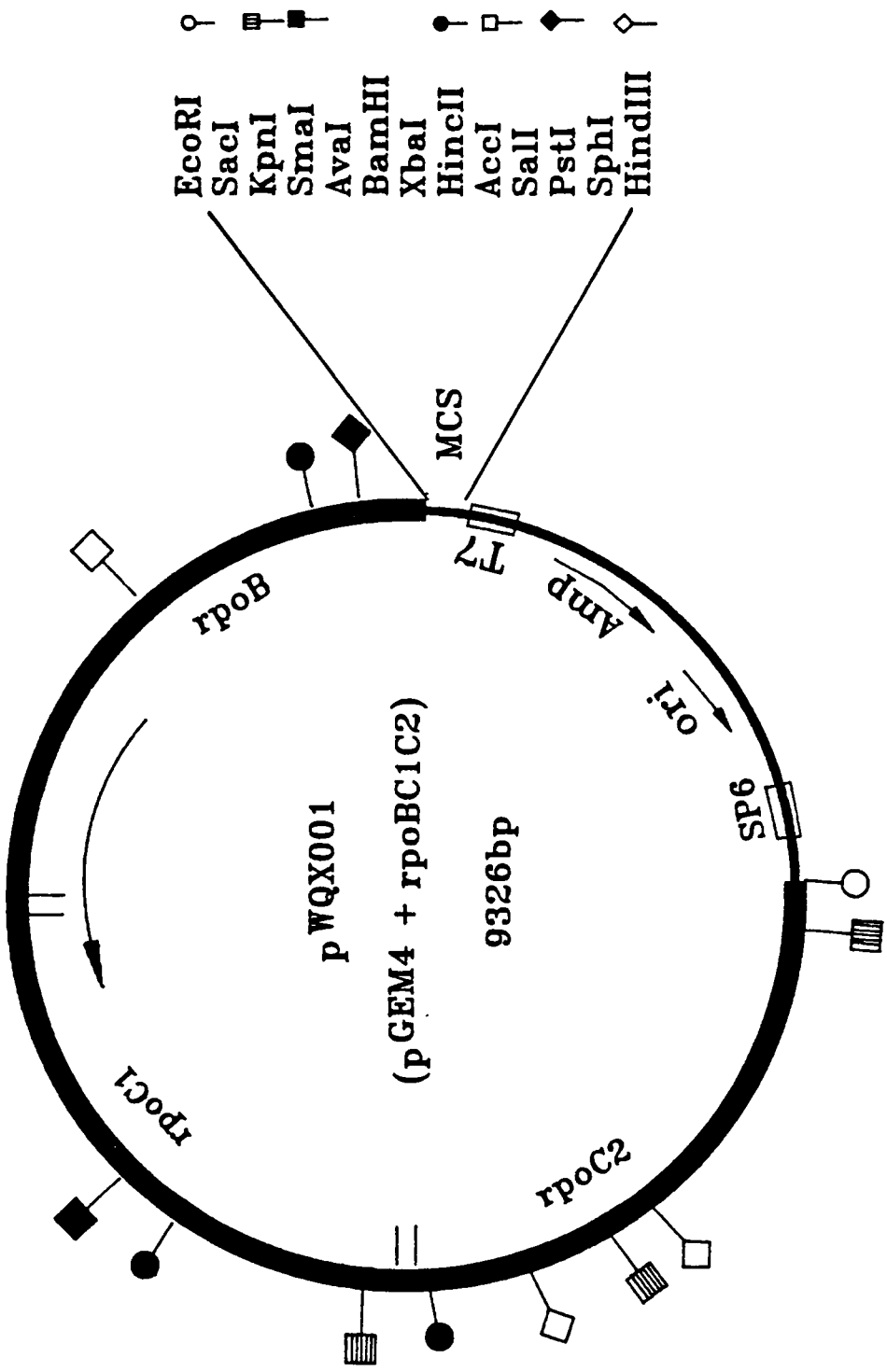

EcoRI
SacI
KpnI
SmaI
AvaI
BamHI
XbaI
HincII
AccI
SalI
PstI
SphI
HindIII
MCS
T7
Amp
ori
SP6
rpoB
rpoC1
rpoC2
pWQX001
(pGEM4 + rpoBC1C2)
9326bp

Figure 9. **Restriction Enzyme Map of pWQX005.**

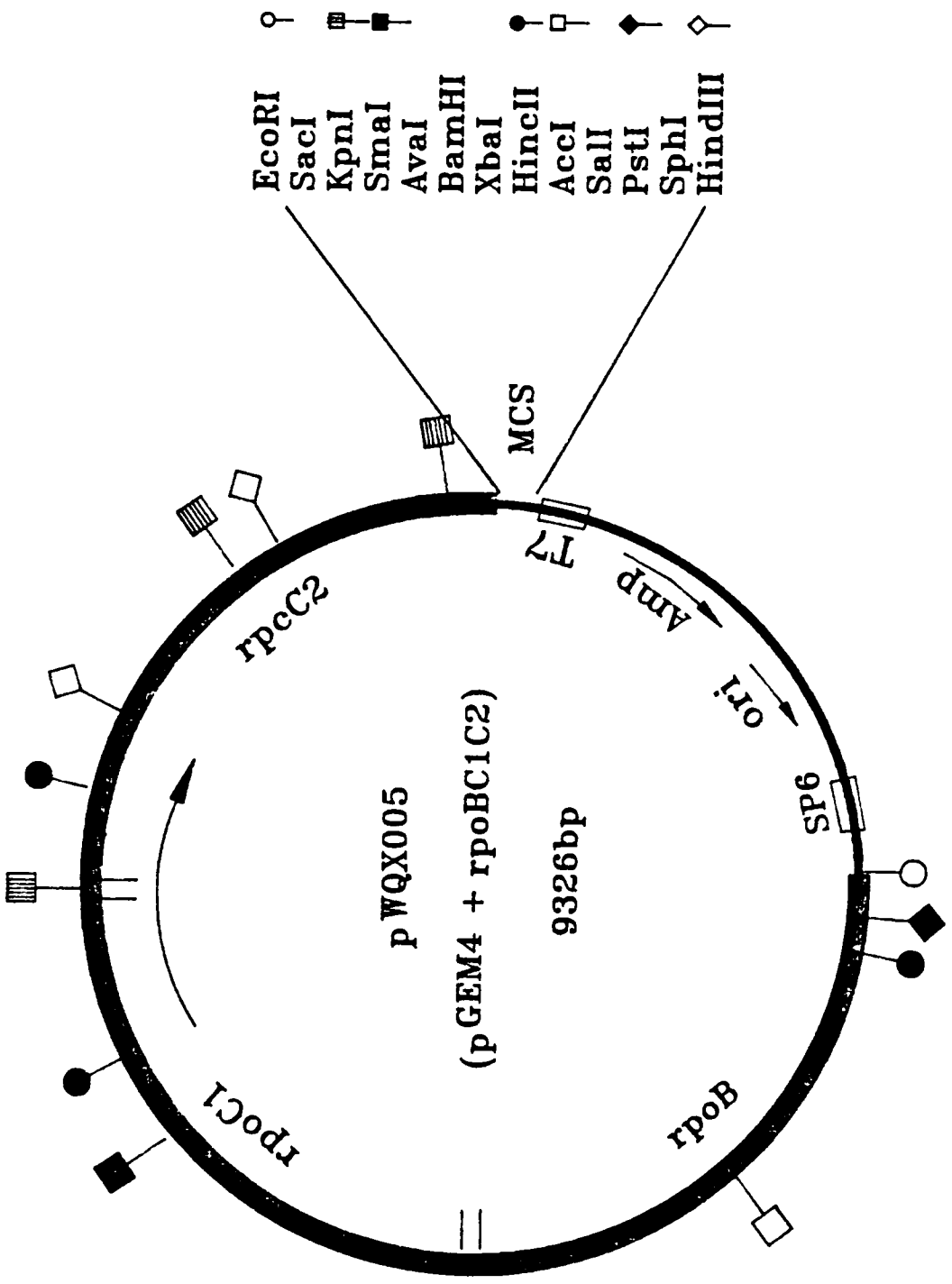

EcoRI
SacI
KpnI
SmaI
AvaI
BamHI
XbaI
HincII
AccI
SalI
PstI
SphI
HindIII
MCS
T7
Amp
ori
SP6
rpoB
rpoC1
rpoC2
pWQX005
(pGEM4 + rpoBC1C2)
9326bp

Figure 10. **Agarose Gel Electrophoresis of Unpurified Plasmid DNA from Deletion Clones.** Lane 1, HindIII-lambda DNA markers (numbers were in kb); lane 2-19, covalently closed circular (ccc) plasmid DNA from deletion clones; lane 20, cccpGEM-4 DNA (2.8 kb); CHR, chromosomal DNA.

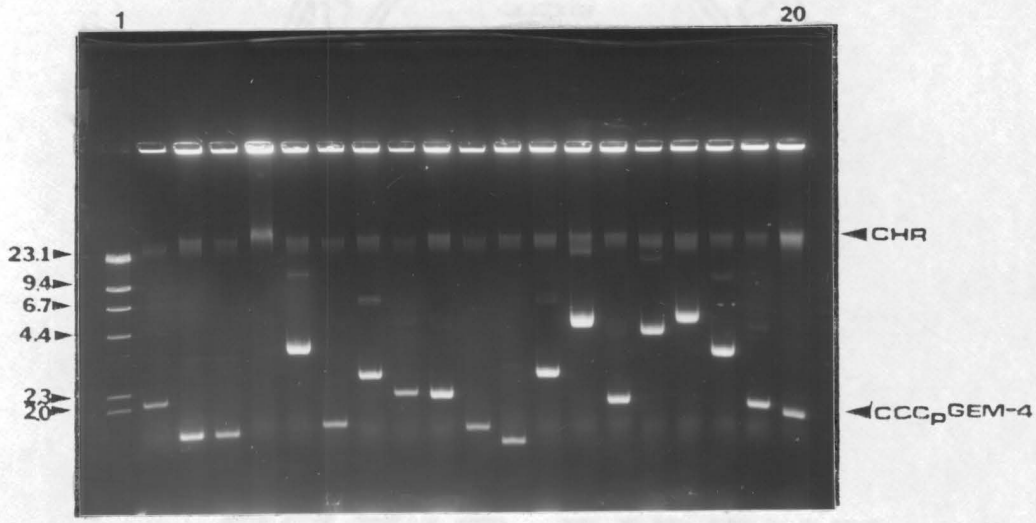
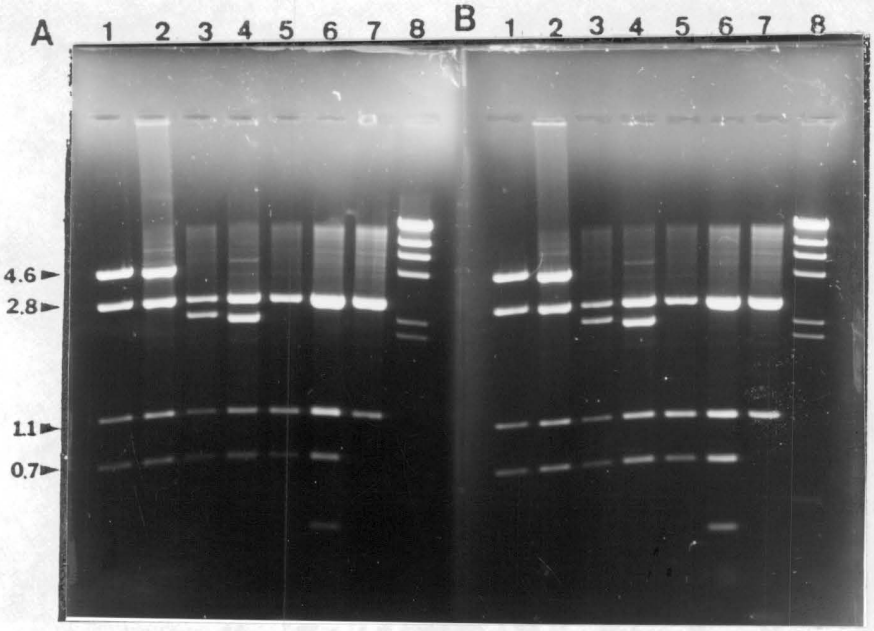

1
20
23.1
9.4
6.7
4.4
2.3
2.0
CHR
CCC pGEM-4

Figure 11. **Electrophoresis (A and B) and Southern Analysis (C and D) of Deletion Derivatives of pWQX001**: All plasmid DNA samples were digested with restriction enzymes EcoRI and HindIII. Lane 1, pWQX001; lane 2, pWQX035; lane 3, pWQX031; lane 4, pWQX025; lane 5, pWQX022; lane 6, pWQX015; lane 7, pWQX012; and lane 8, HindIII digested lambda DNA (size-markers). C was hybridized with biotinylated rpoB probe of E. coli; D was hybridized with biotinylated rpoC probe of E. coli.


A
1 2 3 4 5 6 7 8
B
1 2 3 4 5 6 7 8
4.6
2.8
1.1
0.7

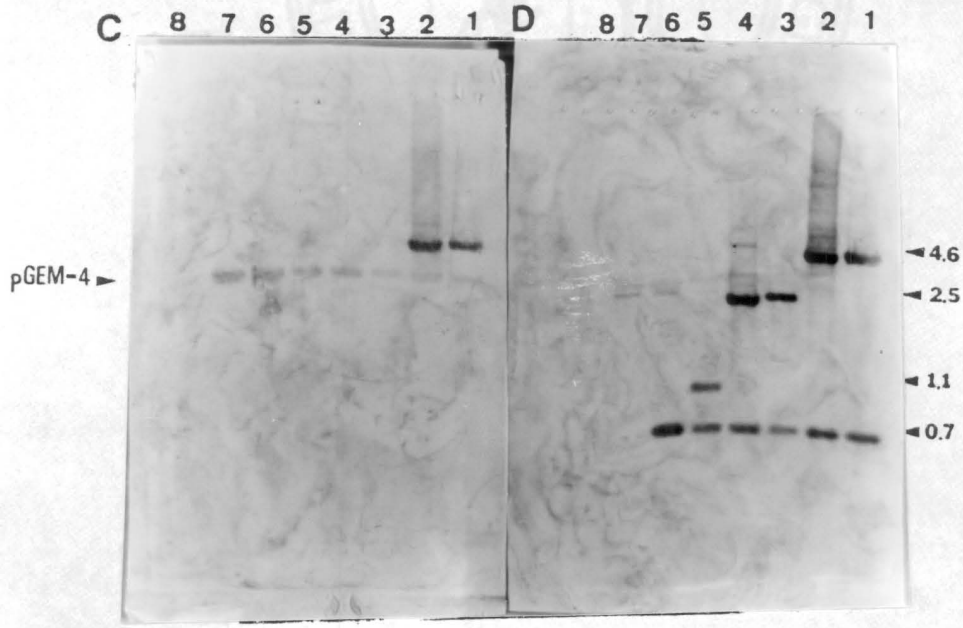

C
8 7 6 5 4 3 2 1
D
8 7 6 5 4 3 2 1
pGEM-4
4.6
2.5
1.1
0.7

Figure 12. **Deletion Clones for DNA Sequencing:** All deletion clones were sequenced with T7 primer.

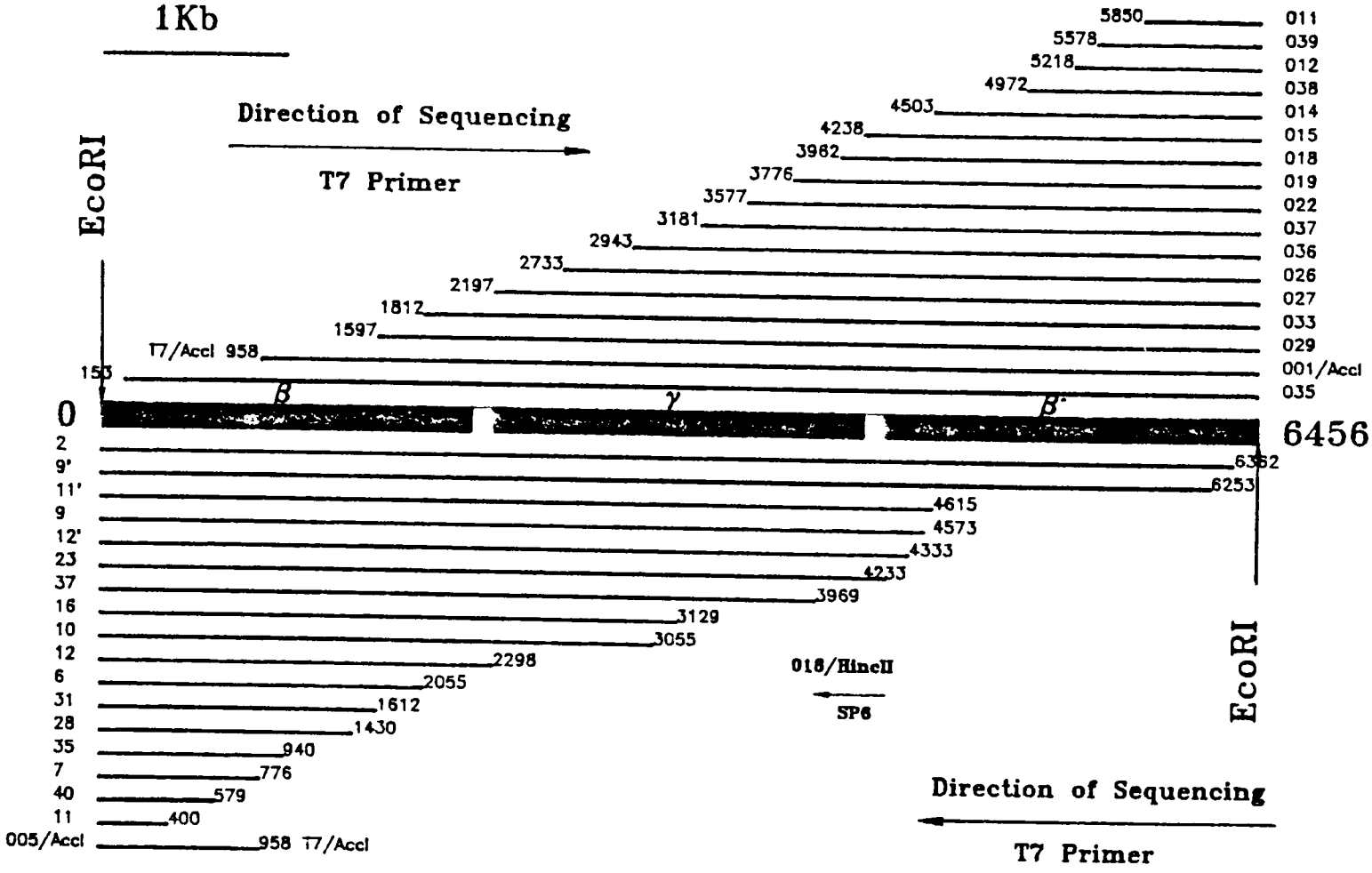

1Kb
Direction of Sequencing
T7 Primer
EcoRI
5850 011
5578 039
5218 012
4972 038
4503 014
4238 015
3962 018
3776 019
3577 022
3181 037
2943 036
2733 026
2197 027
1812 033
1597 029
T7/AccI 958 001/AccI
153 035
β
γ
β'
0
6456
2
9'
11'
9
12'
23
37
16
10
12
6
31
28
35
7
40
11
005/AccI
6362
6253
4615
4573
4333
4233
3969
3129
3055
2298
2055
1612
1430
940
776
579
400
958 T7/AccI
018/HincII
SP6
EcoRI
Direction of Sequencing
T7 Primer

Figure 13. **Nucleotide Sequence of the 3' Portion of rpoB of UTEX 584:** Regions of sequence redundancy are underlined and numbered. This portion of DNA contains three sequence redundancies.

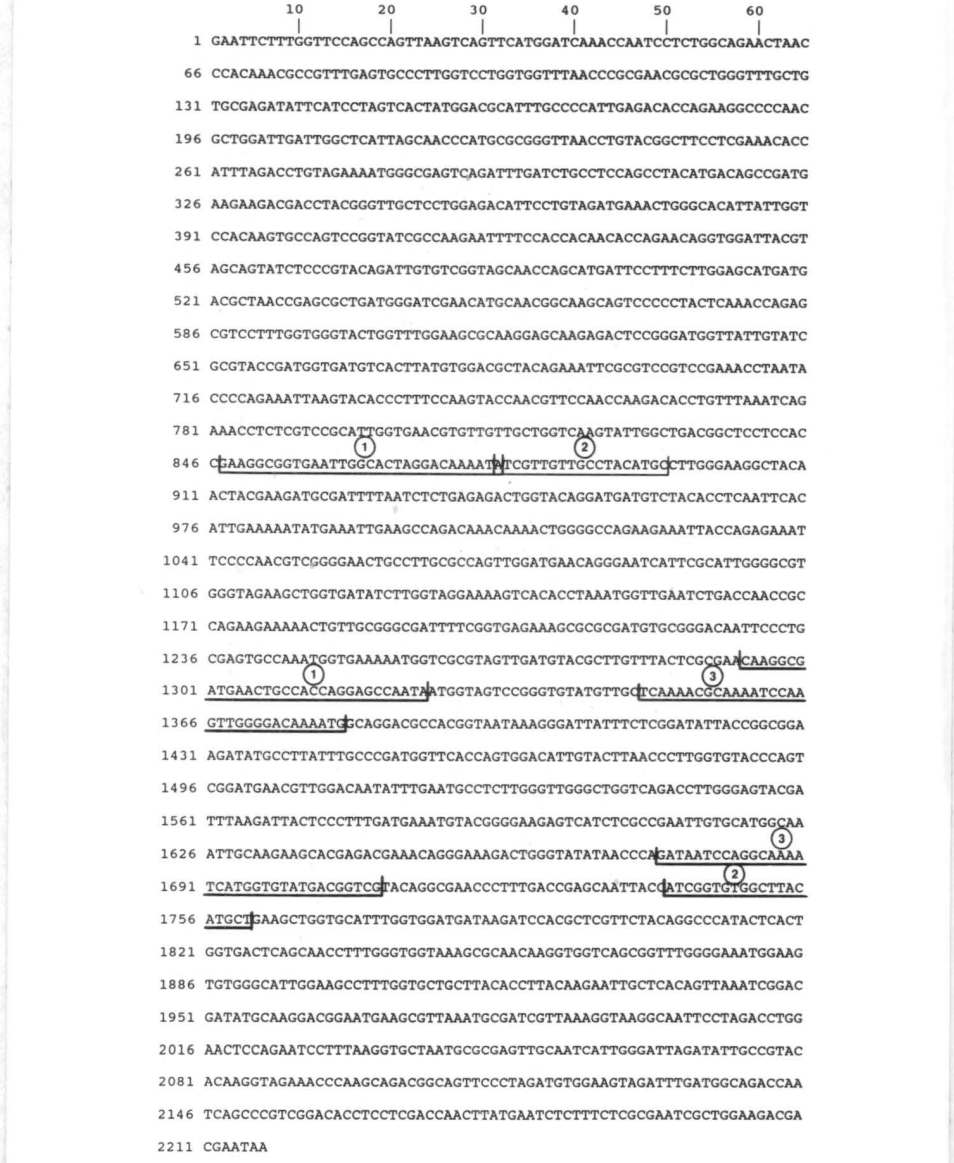

10 20 30 40 50 60
1 GAATTCTTTGGTTCCAGCCAGTTAAGTCAGTTCATGGATCAAACCAATCCTCTGGCAGAACTAAC
66 CCACAAACGCCGTTTGAGTGCCCTTGGTCCTGGTGGTTTAACCCGCGAACGCGCTGGGTTTGCTG
131 TGCGAGATATTCATCCTAGTCACTATGGACGCATTTGCCCCATTGAGACACCAGAAGGCCCCAAC
196 GCTGGATTGATTGGCTCATTAGCAACCCATGCGCGGGTTAACCTGTACGGCTTCCTCGAAACACC
261 ATTTAGACCTGTAGAAAATGGGCGAGTCAGATTTGATCTGCCTCCAGCCTACATGACAGCCGATG
326 AAGAAGACGACCTACGGGTTGCTCCTGGAGACATTCCTGTAGATGAAACTGGGCACATTATTGGT
391 CCACAAGTGCCAGTCCGGTATCGCCAAGAATTTTCCACCACAACACCAGAACAGGTGGATTACGT
456 AGCAGTATCTCCCGTACAGATTGTGTCGGTAGCAACCAGCATGATTCCTTTCTTGGAGCATGATG
521 ACGCTAACCGAGCGCTGATGGGATCGAACATGCAACGGCAAGCAGTCCCCCTACTCAAACCAGAG
586 CGTCCTTTGGTGGGTACTGGTTTGGAAGCGCAAGGAGCAAGAGACTCCGGGATGGTTATTGTATC
651 GCGTACCGATGGTGATGTCACTTATGTGGACGCTACAGAAATTCGCGTCCGTCCGAAACCTAATA
716 CCCCAGAAATTAAGTACACCCTTTCCAAGTACCAACGTTCCAACCAAGACACCTGTTTAAATCAG
781 AAACCTCTCGTCCGCATTGGTGAACGTGTTGTTGCTGGTCAAGTATTGGCTGACGGCTCCTCCAC
846 CGAAGGCGGTGAATTGGCACTAGGACAAAATATCGTTGTTGCCTACATGCCTTGGGAAGGCTACA
911 ACTACGAAGATGCGATTTTAATCTCTGAGAGACTGGTACAGGATGATGTCTACACCTCAATTCAC
976 ATTGAAAAATATGAAATTGAAGCCAGACAAACAAAACTGGGGCCAGAAGAAATTACCAGAGAAAT
1041 TCCCCAACGTCGGGGAACTGCCTTGCGCCAGTTGGATGAACAGGGAATCATTCGCATTGGGGCGT
1106 GGGTAGAAGCTGGTGATATCTTGGTAGGAAAAGTCACACCTAAATGGTTGAATCTGACCAACCGC
1171 CAGAAGAAAAACTGTTGCGGGCGATTTTCGGTGAGAAAGCGCGCGATGTGCGGGACAATTCCCTG
1236 CGAGTGCCAAATGGTGAAAAATGGTCGCGTAGTTGATGTACGCTTGTTTACTCGCGAACAAGGCG
1301 ATGAACTGCCACCAGGAGCCAATAATGGTAGTCCGGGTGTATGTTGCTCAAAACGCAAAATCCAA
1366 GTTGGGGACAAAATGCAGGACGCCACGGTAATAAAGGGATTATTTCTCGGATATTACCGGCGGA
1431 AGATATGCCTTATTTGCCCGATGGTTCACCAGTGGACATTGTACTTAACCCTTGGTGTACCCAGT
1496 CGGATGAACGTTGGACAATATTTGAATGCCTCTTGGGTTGGGCTGGTCAGACCTTGGGAGTACGA
1561 TTTAAGATTACTCCCTTTGATGAAATGTACGGGGAAGAGTCATCTCGCCGAATTGTGCATGGCAA
1626 ATTGCAAGAAGCACGAGACGAAACAGGGAAAGACTGGGTATATAACCCAGATAATCCAGGCAAAA
1691 TCATGGTGTATGACGGTCGTACAGGCGAACCCTTTGACCGAGCAATTACGATCGGTGTGGCTTAC
1756 ATGCTGAAGCTGGTGCATTTGGTGGATGATAAGATCCACGCTCGTTCTACAGGCCCATACTCACT
1821 GGTGACTCAGCAACCTTTGGGTGGTAAAGCGCAACAAGGTGGTCAGCGGTTTGGGGAAATGGAAG
1886 TGTGGGCATTGGAAGCCTTTGGTGCTGCTTACACCTTACAAGAATTGCTCACAGTTAAATCGGAC
1951 GATATGCAAGGACGGAATGAAGCGTTAAATGCGATCGTTAAAGGTAAGGCAATTCCTAGACCTGG
2016 AACTCCAGAATCCTTTAAGGTGCTAATGCGCGAGTTGCAATCATTGGGATTAGATATTGCCGTAC
2081 ACAAGGTAGAAACCCAAGCAGACGGCAGTTCCCTAGATGTGGAAGTAGATTTGATGGCAGACCAA
2146 TCAGCCCGTCGGACACCTCCTCGACCAACTTATGAATCTCTTTCTCGCGAATCGCTGGAAGACGA
2211 CGAATAA

Figure 14. **The Redundancies in rpoB Fragment**: The redundancies are apparent in both the DNA and derived amino acid sequence of rpoB. Numbers refer to the nucleotide or amino acid position in cloned fragment of N. commune rpoB or beta, respectively. The two components of repeat 3 (Nc1, Nc2) are aligned with the repeats in E. coli rpoB (Ec1,Ec2).

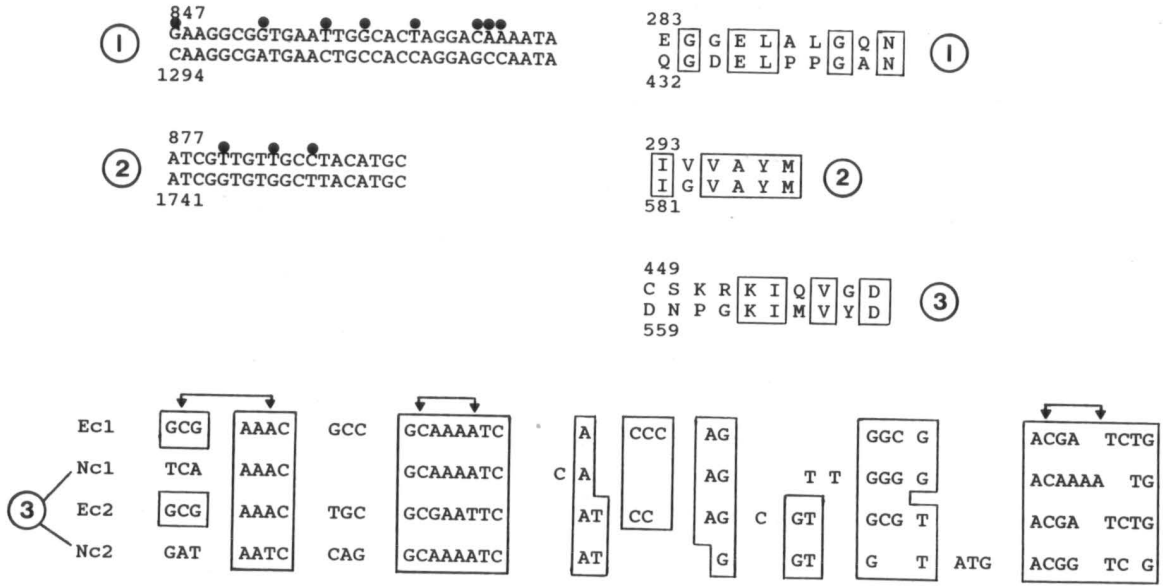

847
GAAGGCGGTGAATTGGCACTAGGACAAAATA
CAAGGCGATGAACTGCCACCAGGAGCCAATA
1294
283
E G G E L A L G Q N
Q G D E L P P G A N
432
I
877
ATCGTTGTTGCCTACATGC
ATCGGTGTGGCTTACATGC
1741
293
I V V A Y M
I G V A Y M
581
2
449
C S K R K I Q V G D
D N P G K I M V Y D
559
3
Ec1
Nc1
Ec2
Nc2
GCG AAAC GCC GCAAAATC
TCA AAAC GCAAAATC
GCG AAAC TGC GCGAATTC
GAT AATC CAG GCAAAATC
A CCC AG GGC G ACGA TCTG
C A AG T T GGG G ACAAAA TG
AT CC AG C GT GCG T ACGA TCTG
AT G GT G T ATG ACGG TC G

Figure 15. **Nucleic Acid Sequence of rpoC1 and N-terminal Portion of rpoC2**: The sequence presented begins with the translational termination codon (Trm) of rpoB. Translational initiation codons for rpoC1 and rpoC2 are indicated (Met). In the sequences between the coding regions of rpoB, rpoC1 and rpoC2, light underlining indicates direct repeats and bold indicates two regions of sequence similarity (see Figures). Putative Shine-Dalgarno sequences are indicated with overlining.

```
                    10         20         30         40         50         60
                Trm |          |          |          |          |          |
   1 TAAGGAGTGCTGAGTGCTAAATGCTAAATGCTGAGTATTAAAAAACTCAGTAATTAGCATTTAGC

  66 ACTTAGAAATCTCTTAATTCAAACCTTAAAACTAAAATATACTCAAAACTCGGAACTCAGCACTT

         Met
 131 AAGTATGAGACCTGCCCAAACTAATCAGTTTGACTACGTAAAAATCGGCTTGGCTTCCCCTGAAC

 196 GTATTCGTCAGTGGGGCGAGAGAACATTGCCCAATGGTCAAGTAGTCGGTGAAGTTACCAAGCCA

 261 GAGACAATTAACTACCGAACTCTCAAGCCAGAGATGGATGGCTTATTTTGTGAGCGGATCTTTGG

 326 TCCGAAAGATTGGGAATGCCATTGTGGCAAGTATAAAGAGTCCGTCATAGAGGTATTGTCTGTGA

 391 GCGCTGTGGTGTTAGAAGTCACCGAGTCACGGGTGCGTCGTCACCGTATGGGCTATATCAAACTC

 456 GCCGCACCAGTAGCTCACGTTGGTTACTCAAAGGCATTCCTAGCTATATTCCATTCTGTTGGATA

 521 TGCCCTACGCGATGTCGAGCAGATTGTCTATTTCAACTCTTATGTTGTCCTGAGTCCTGGTAATG

 586 CTGAAACTTTAACTTACAAACAACTACTCAGTGAAGATCAGTGGTTGGAAATAGAAGACCAAATT

 651 TATAGCGAAGATTCTCTGCTGCAAGGCGTAGAGGTAGGTATTGGCGCTGAAGCGCTATTGCGTTT

 716 GCTTGCCGATATCAATTTGGAACAAGAAGCGGAAAGCCTACGGGAAGAAATTGGCAACGCCAAGG

 781 GACAAAAGCGGGCCAAGCTAATTAAGCGACTGCGGGTAATTGACAACTTCATCGCAACTGGTTCC

 846 AAACCAGAGTGGATGGTAATGGCAGTTATTCCTGTGATTCCCCCTGACTTGCGCCCAATGGTACA

 911 GCTAGATGGCGGACGGTTTGCCACCAGCGATTTAAATGATTTATATCGACGGGTAATTAACCGCA

 976 ACAATCGTTTGGCACGCTTGCAAGAAATTTTAGCACCAGAAATTATTGTGCGGAACGAAAAGCGG

1041 ATGCTGCAAGAAGCAGTGGATGCTTTGATTGACAATGGTCGTCGGGGACGCACTGTAGTAGGGGC

1106 AAATAACCGACCACTGAAATCTTTGTCCGACATTATTGAGGGTAAGCAAGGACGTTTCCGACAAA

1171 ACTTGTTAGGTAAACGGGTTGACTACTCTGGACGTTCGGTTATTGTAGTCGGGCCAAAGCTGAAA

1236 ATTCACCAGTGCGGTTTACCTAGAGAAATGGCAATTGAGCTATTTCAACCATTTGTAATTAACCG

1301 CCTGATTCGTAGCGGTATGGTAATACATCAAGCTGCGCCAATGATATCGCGTAATGACCCCAGTG

1366 TTTGGGATGTGCTGGAAGAAGTGATTGAAGGACACCCAGTGATGCTAAACAGGGCGCCAACATTG

1431 CACCGTTTGGGTATTCAGTCTTTTGAACCGATTTTGGTAGAAGGTAGAGCGATTCAACTGCATCC

1496 TCTGGTGTGTCCAGCGTTTAACGCCGACTTTGACGGGGACCAAATGGCGGTACACGTCCCTCTTT

1561 CTTTAGAAAGTCAGGCTGAAGCGCGATTGTTGATGTTGGCTTCTAACAATATTTTGTCACCAGCC

1626 ACGGGTAAACCGATCATCACGCCTAGCCAAGATATGGTGTTGGGAGCCTATTATTTAACAGCAGA

1691 AAATCCCGGTGCGACAAAAGGTGCAGGAAAGTACTTTTCTTCGCTAGAGGATGTAATCATGGCTT

1756 TCCAGCAAGAACAAATTGACTTGCACGCCTATATTTACGTGCGGTTTGACGGTGAAATAGAATCA

1821 GACCAACCGGATACGGAACCCGTGAAAGTGACGGAAAACGAGGATGGTACCCGTACATTACTCTA

1886 TAAGTTCCGTCGAGTCAGACAAGACGCTAAAGGCAATGTACTTTCCCAGTATATATACACAACTC

1951 CAGGTCGCGTTATTTACAATAATGCTATTCAGGAAGCACTAGCAAGCTAAAAGTAAGGAGTGAAA
                                                    Trm

2016 AGTTAGGAGTTAGGAGTTATTAATTACAAACTCCTAACTCCTAACTCCTAACTATTAATTTAGGA
```

```
                Met
2081 CTAATGACTAATGACTAACGAAAAAATGATTTTTCGCAATCGGGTGGTTGATAAAGGTCAACTGA

2146 GAAATTTAATTTCTTGGGCCTTTACCTATTATGGTACGGCGCGAACCGCAGTAATGGCAGACAAA

2211 CTGAAAGATTTGGGATTTCGCTACGCTACCAAAGCAGGGGTTTCCATCAGTGTAGATGACTTGAT

2276 GGTGCCACCAACTAAGCGATTGCTCTTAGAAGCAGCCGAAGAAGAAATTCGGGCTACTGAAACCC

2341 GTTACCAACGAGGAGAAATCACAGAAGTAGAACGCTTCCAAAAGGTAATCGATACCTGGAACGGT

2406 ACTAGTGAAGCTTTGAAAGATGAAGTAGTCGTTCACTTCAAAAAGACTAATCCCCTAAGCTCCGT

2471 GTATATGATGGCATTTTCCGGTGCACGGGGTAACATCTCTCAAGTGCGGCAATTGGTGGGGATGC

2536 GGGGACTGATGGCAGATCCTCAAGGGGAAATTATCGATTTGCCCATCAAAACCAATTTCCGTGAA

2601 GGATTAACTGTAACGGAATACATTATTTCGTCTTACGGTGCCAGAAAAGGATTGGTGGTACAGCC

2666 TTCACGGACGGCTGACTCTGGTTATCTCACCCGCCGATTGGTGGATGTATCCCAGGTGTATTATT

2731 CGGGATTTGACTGCGGCACGCCAGAGCTTAGCATTCGACCAATGACAGAAGGGGCCAAAACCTTG

2796 ATTCCTCTAGCAACTCGCTTGATGGGACGGGTAATTGGCGAAGATGTGCTGCATCCGGTAACAAA

2861 AGAAGTGATTGCAGCACGCAATTCCCCAATTTCTGAGGATTTGGCGAAGAAAATTGAAAAATCTG

2926 GGGTGGGCGAAGTTGTGGTGCGATCGCCACTAACTTGTGAAGCTGCACGTTCTGTCTGTCAGCAC

2991 TGCTACGGCTGGAGTTTGGCACACGCAAGCATGGTAGATTTGGGTGAAGCTGTGGGGATTATTGC

3056 CGCCCAAAGTATCGGCGAACCTGGTACCCAGCTAACCATGCGGACATTCCACACAGGTGGGGTGT

3121 TTACTGGAGAAGTGGCGCAACAAGTGCGTTCCAAAATCGATGGTACTGTCAAGCTTCCTCGCAAA

3186 CTGAAGACCAGAACATATCGTACTCGCCACGGGGAAGATGCCCTCTATGTTGAGGCTAATGGCAT

3251 CATGCTTTTGGAGCCAACAAAAGTAGGTGATGTTACCCCAGAAAACCAAGAGGTTCATCTTACCC

3316 AAGGTTCAACACTATATGTATTTGATGGAAATAAGGTAAACAAGGTCCAGTTGTTAGCAGAGGTT

3381 GCCCTTGGTGGACGTACAACTCGGACTAATACAGAAAAAGCAGTTAAAGACGTCGCTTCTGACTT

3446 AGCAGGGGAAGTGCAATTTGCCGAAGTTGTTCCAGAACAAAAAACTGACCGTCAAGGCAACACTA

3511 CAACCACACGCCGCACGCGTGGCTTGATTTGGATTTTGTCTGGGGAAGTTTACAACTTGCCGCCA

3576 GGGGCCGAATTGGTGGTGAAAAATGGTGATGCGATCGCTTCAAATGGAGTTTTAGCAGAAACCAA

3641 GTTAGCCAGTTTGCACGGCGGTGTGGTGCGCTTGCCAGAAGCTACCCCAGGTAAGAGTACCAGGG

3706 AAATTGAGATTATCACCGCTTCTGTAGTCTTAGACCAGGCAACGGTGACAGTTCAAAGTTCTCAA

3771 GGACGTAATAACTACTTAGTTTCTACTGGCAACAACCAAGTATTTAACCTCCGGGCTACACCAGG

3836 CACAAAAGTGCAAAATGGTCAAGTAGTAGCTGAGTTAATTGATGACCGCTATCGCACAACCACTG

3901 GTGGATTCCTGAAATTTGCTGGTGTAGAAGTCCAGAAAAAAGGCAAAGCCAAGCTGGGTTATGAA

3966 GTCGTGCAGGGCGGTACCTGTTGTGGATCTCCTGAAGAAAGCCACGAAGTTAATAAAGATATCTC

4031 CTTGCTGTTGGTGGAAGACGGCCAGTTTGTGGAAGCTGGCACCGAAGTCGTGAAAGATATCTTCT

4096 GCCAAAACAGTGGTGTGGTAGAAGTGACCCAGAAAAACGACATCCTCCGGGAAGTCGTGGTGAAG

4161 CCAGGGGAACTGCTGATGGTGGACGATCCAGAATCAGTCATCGGGCGAGATAACACCTTCATCCA

4226 ACCAGGTGAGGAATTC
```

Figure 16. **Comparasion of Nucleotide Sequences from Within the Intergenic Regions of rpoB, rpoC1 and rpoC2**: The stretches of sequence with 100% correspondence are boxed and have been aligned for the best fit. The distances (in bp) from the first T in the TTAATT unit to the A of the translation initiation codon of each gene are indicated.


-56
ATG
AAAA C T C AGT A A TT AG CAT TTAG C A C TTAG A A A T CTC TTAATT CAAAC rpoC1
AAAA G T AAGG AGT G A AA AG TTAG G A G TTAG G A G T TA TTAATT A CAAAC rpoC2
-55
ATG

Figure 17. **Comparision of Amino Acid Sequences of the Beta Subunit of E. coli (Ec), N. commune (Nc) and Tobacco Chloroplast (Tc)**: Numbers refer to the residue positions in the E.coli beta subunit sequence.

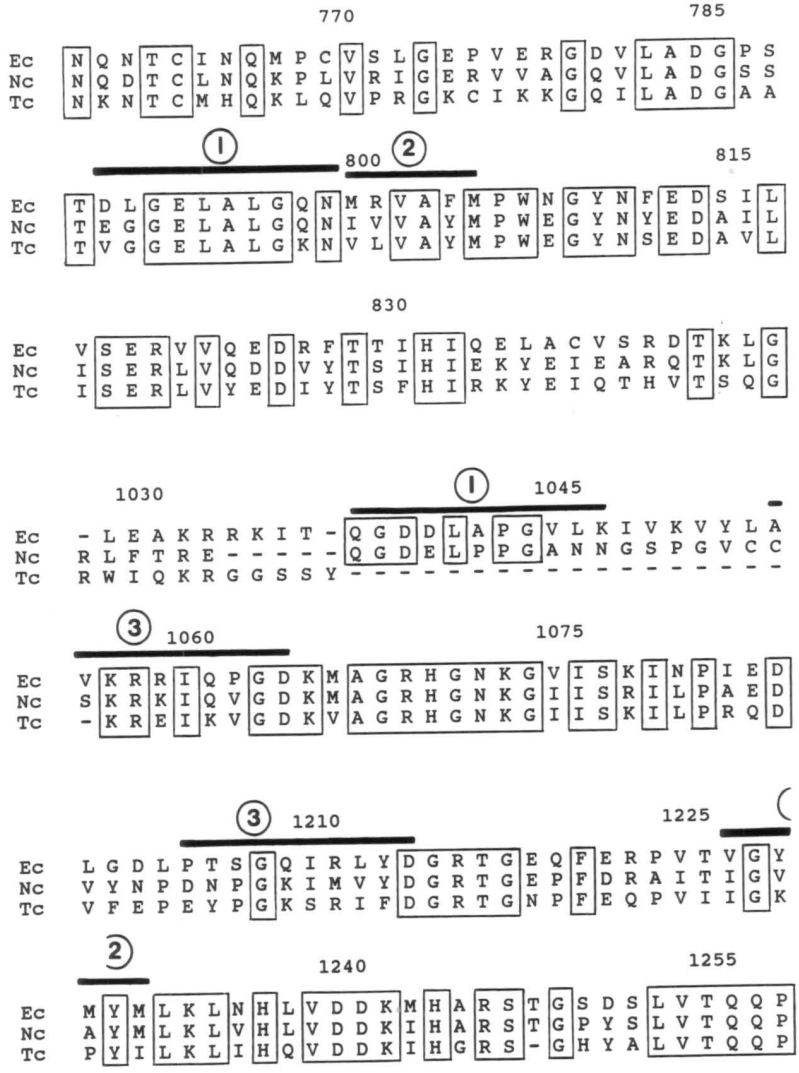

770
785
Ec N Q N T C I N Q M P C V S L G E P V E R G D V L A D G P S
Nc N Q D T C L N Q K P L V R I G E R V V A G Q V L A D G S S
Tc N K N T C M H Q K L Q V P R G K C I K K G Q I L A D G A A
1
800
2
815
Ec T D L G E L A L G Q N M R V A F M P W N G Y N F E D S I L
Nc T E G G E L A L G Q N I V V A Y M P W E G Y N Y E D A I L
Tc T V G G E L A L G K N V L V A Y M P W E G Y N S E D A V L
830
Ec V S E R V V Q E D R F T T I H I Q E L A C V S R D T K L G
Nc I S E R L V Q D D V Y T S I H I E K Y E I E A R Q T K L G
Tc I S E R L V Y E D I Y T S F H I R K Y E I Q T H V T S Q G
1030
1
1045
Ec - L E A K R R K I T - Q G D D L A P G V L K I V K V Y L A
Nc R L F T R E - - - - - Q G D E L P P G A N N G S P G V C C
Tc R W I Q K R G G S S Y - - - - - - - - - - - - - - - - - -
3
1060
1075
Ec V K R R I Q P G D K M A G R H G N K G V I S K I N P I E D
Nc S K R K I Q V G D K M A G R H G N K G I I S R I L P A E D
Tc - K R E I K V G D K V A G R H G N K G I I S K I L P R Q D
3
1210
1225
Ec L G D L P T S G Q I R L Y D G R T G E Q F E R P V T V G Y
Nc V Y N P D N P G K I M V Y D G R T G E P F D R A I T I G V
Tc V F E P E Y P G K S R I F D G R T G N P F E Q P V I I G K
2
1240
1255
Ec M Y M L K L N H L V D D K M H A R S T G S D S L V T Q Q P
Nc A Y M L K L V H L V D D K I H A R S T G P Y S L V T Q Q P
Tc P Y I L K L I H Q V D D K I H G R S - G H Y A L V T Q Q P

Figure 18. **Alignment of Derived Amino Acid Sequence of Gamma (Nc) with the N-terminal Portion of E. coli Beta' (Ec)**: Dashed lines indicate shifts of the sequences for the best fit. Numbering refers to the amino acid residues in the E. coli beta' subunit. Regions of sequence correspondence are enclosed in a single box. A double box around region of sequence indicates the segments in E. coli beta' subunit with extensive sequence similarity to component A of eukaryotes.

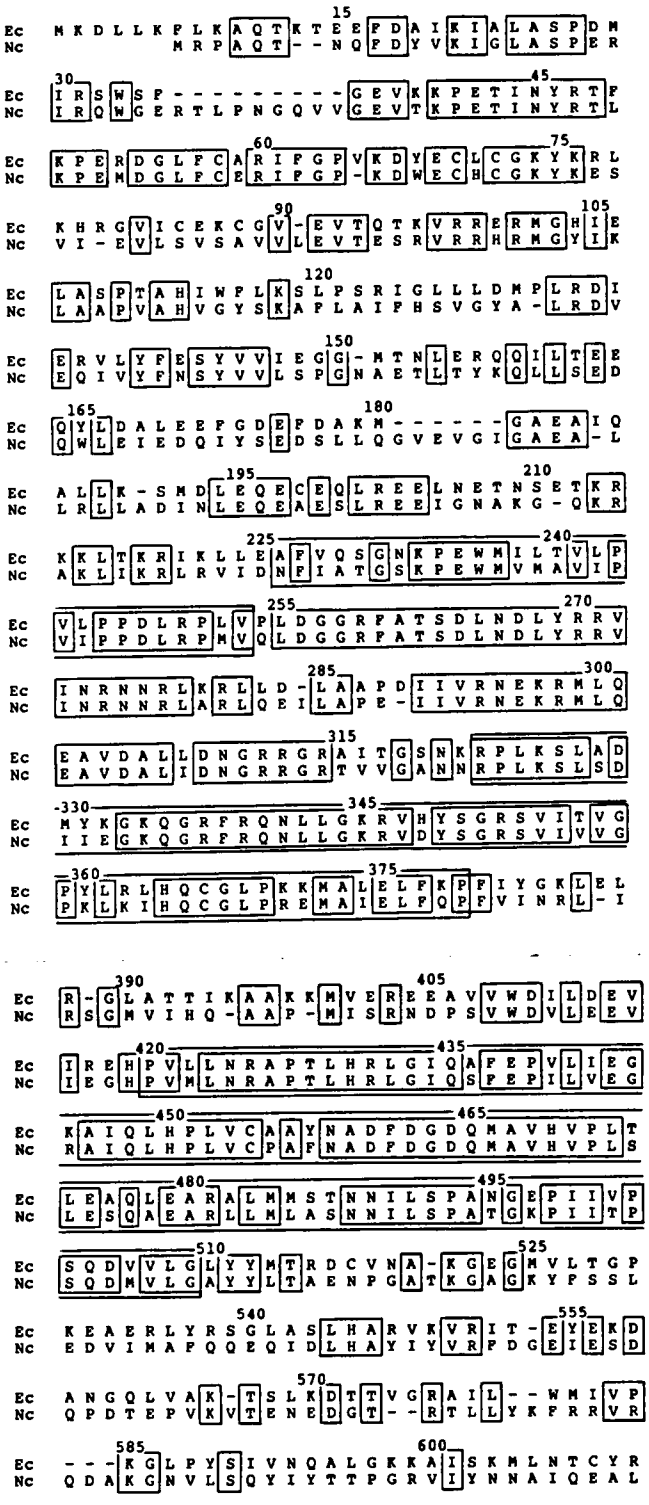

Ec MKDLLKFLKAQTKTEEFDAIKIALASPDM
Nc MRPAQT--NQFDYVKIGLASPER
Ec IRSWSP----------GEVKKPETINYRTF
Nc IRQWGERTLPNGQVVGEVTKPETINYRTL
Ec KPERDGLFCARIFGPVKDYECLCGKYKRL
Nc KPEMDGLFCERIFGP-KDWECHCGKYKES
Ec KHRGVICEKCGV-EVTQTKVRREHMGHIE
Nc VI-EVLSVSAVVLEVTESRVRRHRMGYIK
Ec LASPTAHIWFLKSLPSRIGLLLDMPLRDI
Nc LAAPVAHVGYSKAFLAIFHSVGYA-LRDV
Ec ERVLYFESYVVIEGG-MTNLERQQILTEE
Nc EQIVYFNSYVVLSPGNAETLTYKQLLSED
Ec QYLDALEEFGDEFDAKM------GAEAIQ
Nc QWLEIEDQIYSEDSLLQGVEVGIGAEA-L
Ec ALLK-SMDLEQECEQLREELNETNSETKR
Nc LRLLADINLEQEAESLREEIGNAKG-QKR
Ec KKLTKRIKLLEAFVQSGNKPEWMILTVLP
Nc AKLIKRLRVIDNFIATGSKPEWMVMAVIP
Ec VLPPDLRPLVPLDGGRFATSDLNDLYRRV
Nc VIPPDLRPMVQLDGGRFATSDLNDLYRRV
Ec INRNNRLKRLLD-LAAPDIIVRNEKRMLQ
Nc INRNNRLARLQEILAPE-IIVRNEKRMLQ
Ec EAVDALLDNGRRGRAITGSNKRPLKSLAD
Nc EAVDALIDNGRRGRTVVGANNRPLKSLSD
Ec MYKGKQGRFRQNLLGKRVHYSGRSVITVG
Nc IIEGKQGRFRQNLLGKRVDYSGRSVIVVG
Ec PYLRLHQCGLPKKMALELFKPFIYGKLEL
Nc PKLKIHQCGLPREMAIELFQPFVINRL-I
Ec R-GLATTIKAAKKMVERE EAVVWDILDEV
Nc RSGMVIHQ-AAP-MISRNDPSVWDVLEEV
Ec IREHPVLLNRAPTLHRLGIQAFEPVLIEG
Nc IEGHPVMLNRAPTLHRLGIQSFEPILVEG
Ec KAIQLHPLVCAAYNADFDGDQMAVHVPLT
Nc RAIQLHPLVCPAFNADFDGDQMAVHVPLS
Ec LEAQLEARALMMSTNNILSPANGEPIIVP
Nc LESQAEARLLMLASNNILSPATGKPIITP
Ec SQDVVLGLYYMTRDCVNA-KGEGMVLTGP
Nc SQDMVLGAYYLTAENPGAT KGAGKYFSSL
Ec KEAERLYRSGLASLHARVKVRITE-EYEKD
Nc EDVIMAFQQEQIDLHAYIYVRFDGEIESD
Ec ANGQLVAK-TSLKDTTVGRAIL--WMIVP
Nc QPDTEPVKVTENEDGT--RTLLYKFRRVR
Ec ---KGLPYSIVNQALGKKAISKMLNTCYR
Nc QDAKGNVLSQYIYTTPGRVIYNNAIQEAL
Ec ILGLKPTVIFADQIMYTGFAYAARSG
Nc AS

Figure 19. **Alignment of the Derived Amino Acid Sequence of rpoC2 (beta') with the Carboxy Terminal Portion of the E. coli Beta' Subunit**: The conventions for numbering and labeling are as in Figure 18.

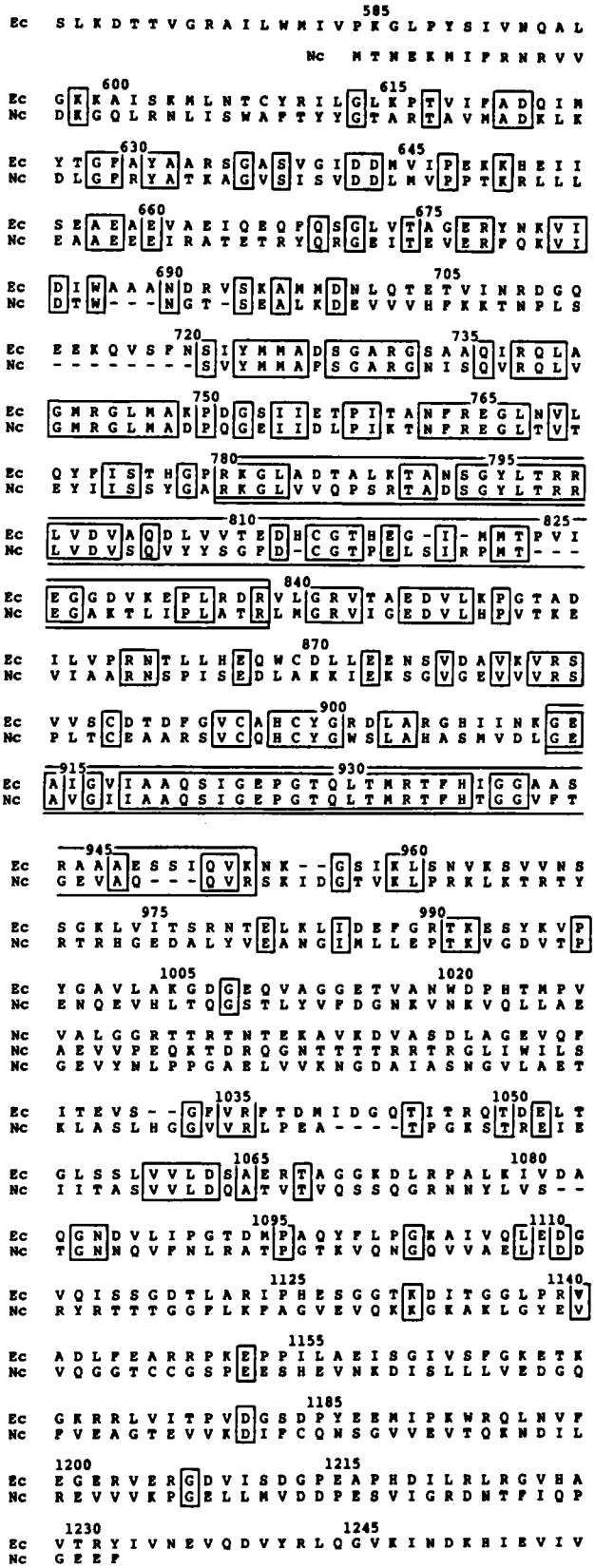


Figure 20. **Predicted Secondary Structure of the Intergenic Region Separating rpoB and rpoC1**: The structure was computed using the algorithm of Zucker and Stiegler (1981).

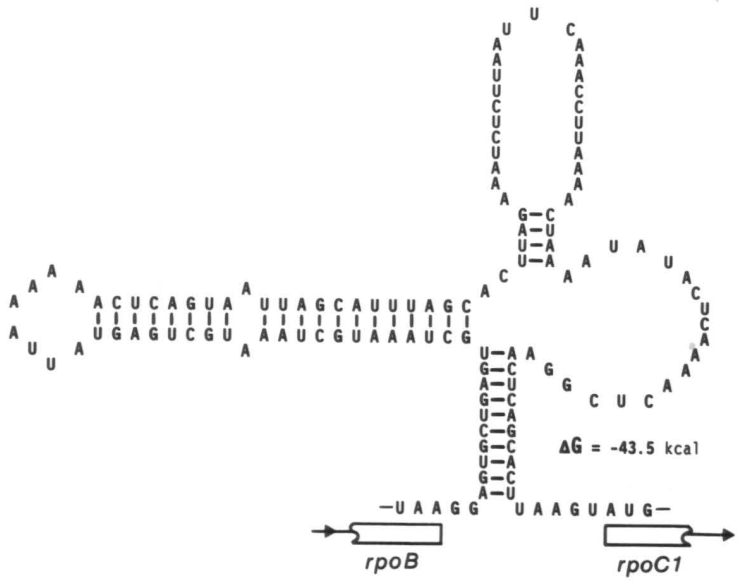

ΔG = -43.5 kcal
rpoB
rpoC1

Figure 21. **Predicted Secondary Structure of the Intergenic Region Separating rpoC1 and rpoC2.**

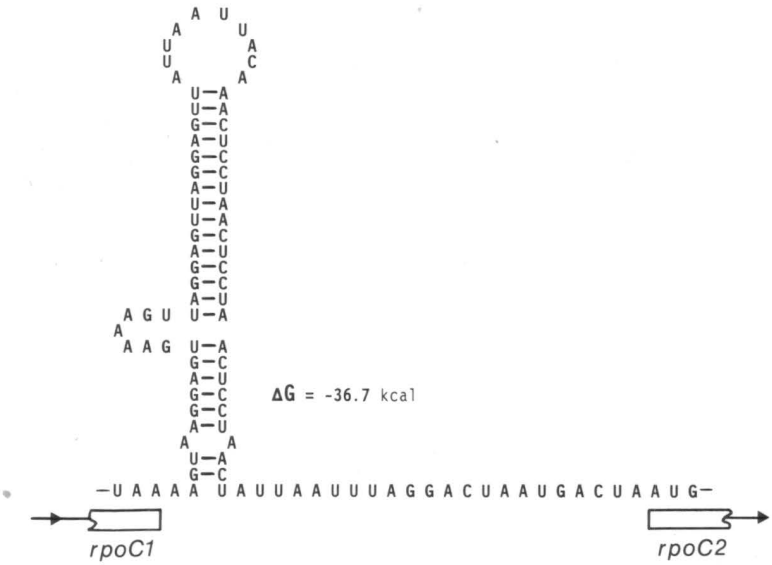

ΔG = -36.7 kcal
—U A A A A U A U U A A U U U A G G A C U A A U G A C U A A U G—
rpoC1
rpoC2

Figure 22. **Subclone pX1 (pUC18 + rpoBC1C2 of N. commune).**

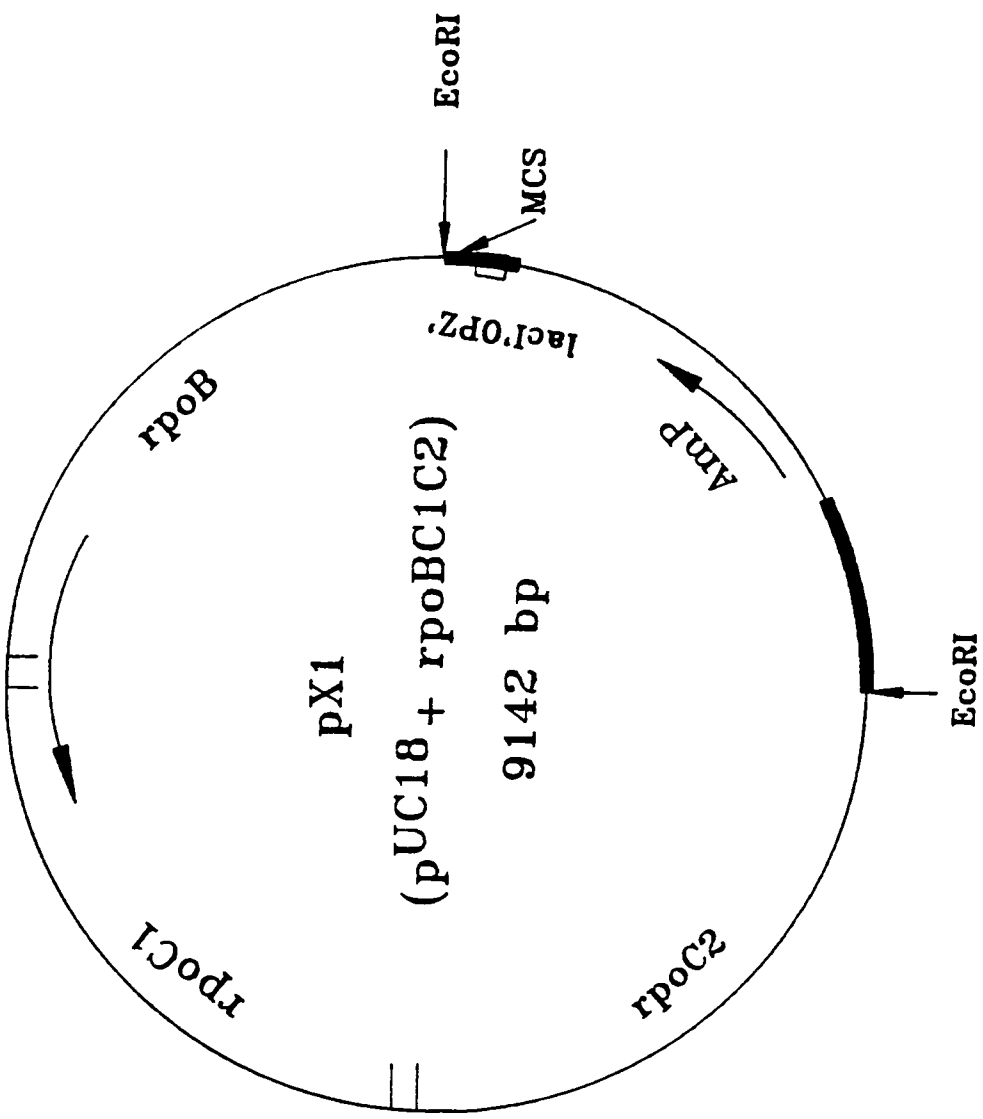

EcoRI
MCS
lacI'OPZ'
AMP
EcoRI
rpoC2
rpoC1
rpoB
pX1
(pUC18+ rpoBC1C2)
9142 bp

Figure 23. **Immuno- and Autoradiographic Analysis of the Gene Products Resulting from Expression of rpoBC1C2**: A: Immunoanalysis of Western blot with Anabaena sp. PCC 7120 anti-core serum. Antiserum was used at a dilution of 1:5000 and protein A - HRP was used at a dilution of 1:10,000. Dark arrows indicate the positions of bands. beta'$_N$ = N terminal portion of N. commune UTEX 584 beta' subunit. The numbers are kilodaltons. a: Lane 1, total proteins isolated from fresh N. commune cells; lane 2, cell-free in vitro transcription-translation minus DNA (endogenous activity) + ^{35}S-methionine; lane 3, as lane 2 with addition of linearized pWQX035 DNA + ^{35}S-methionine; lane 4, as lane 3 but using unlabeled L-methionine; lane 5, total protein of E.coli JM109 (pX1); lane 6, total protein of E. coli JM109 (added in excess over sample in lane 5). B: after immunoanalysis of the Western blot the filter was air-dried and placed in contact with Kodak XAR-5 film for 15 min. For these analysis, levels of ^{35}-incorporation were measured prior to the loading of the gel; equal amounts of radioactive material (cpm) were loaded in lanes 2 and 3.

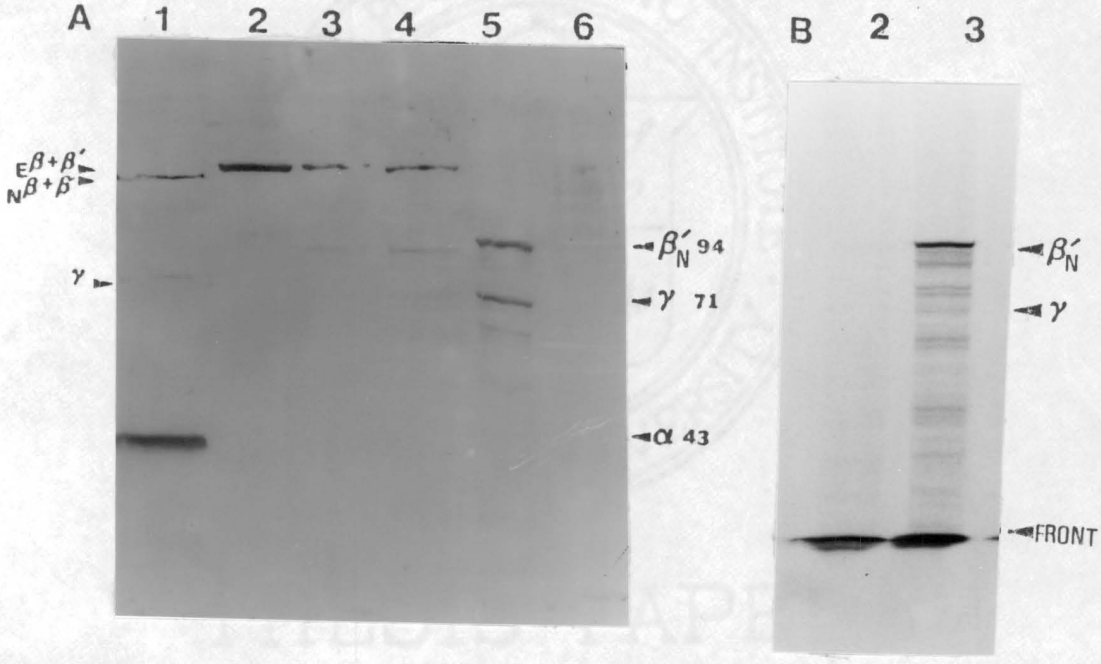

A
1
2
3
4
5
6
$E^{\beta+\beta'}$
$N^{\beta+\beta'}$
γ
β'_N 94
γ 71
α 43
B
2
3
β'_N
γ
FRONT

Figure 24. **In vitro Expression of rpo Genes of N. commune Using a Cell-free Transcription-translation System of E. coli**: A: Coomassie staining of gel. Lane 1, size-markers; lane 2, total proteins of N. commune; lane 3, no sample; lane 4, translated proteins present in supernatant after immuno-precipitation; lane 5, translated proteins present in pellet after immuno-precipitation; lane 6, total translated proteins without immuno-precipitation. B: The gel was dried and exposed using Kodak type 2 film for 15 min.

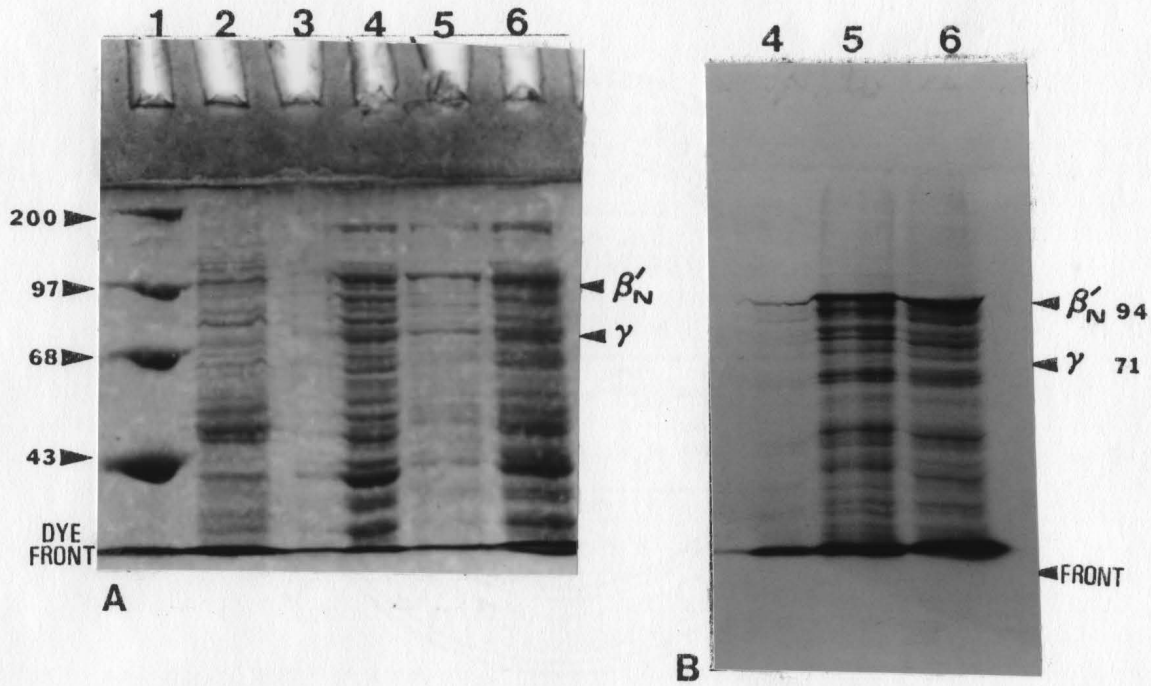
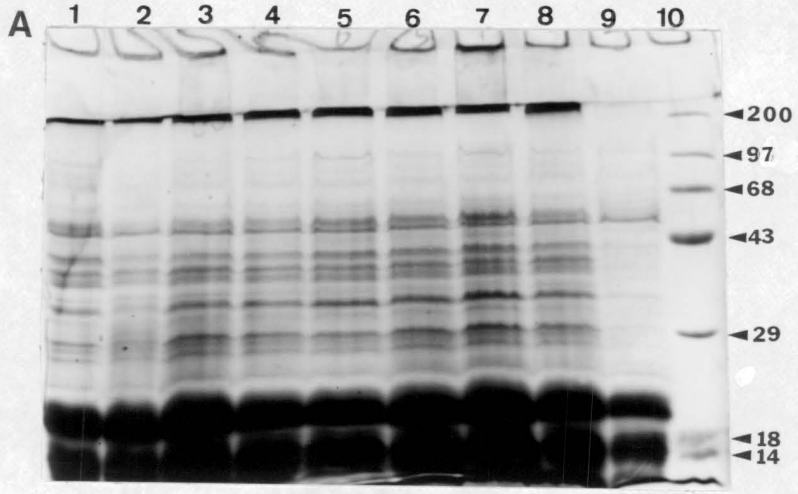

1 2 3 4 5 6
200
97
68
43
DYE FRONT
β'_N
γ
A
4 5 6
β'_N 94
γ 71
FRONT
B

Figure 25. **A: SDS-PAGE Analysis of Proteins Isolated from Dried and Rewetted N. commune Cells**: Proteins were visualized by Coomassie staining. Lane 1, 1-day dried; lane 2, 5-day dried; lane 3, 10-min rewetting; lane 4, 30-min rewetting; lane 5, 1-hour rewetting; lane 6, 5-hour rewetting; lane 7, 24-hour rewetting; lane 8, 4-day rewetting; lane 9, fresh cells; lane 10, protein size-markers. **B: Western Blot of Proteins after SDS-PAGE**: The proteins on the filter were visualized by Amido-black staining.


A
1
2
3
4
5
6
7
8
9
10
200
97
68
43
29
18
14

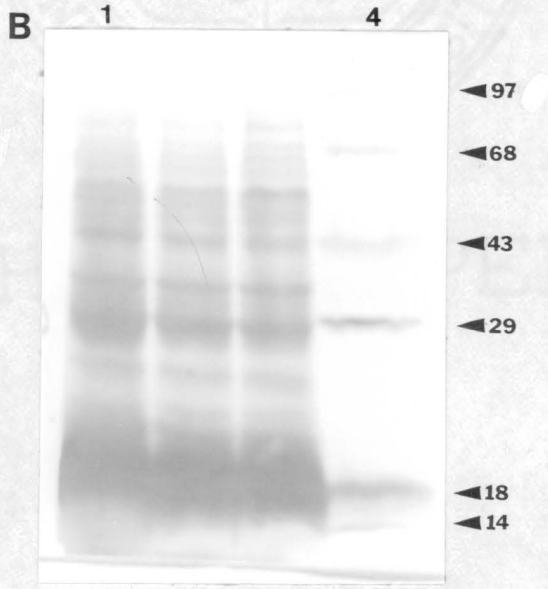

B
1
4
97
68
43
29
18
14

Figure 26. **SDS-PAGE, Western Blot and Immuno-analysis of Proteins Isolated from Dried and Rewetted N. commune Cells:** The lanes here are the same as those in Figure 24. The antiserum was used at a dilution of 1:5000, and the anti-IgG (rabbit) alkaline phosphatase was used at a dilution of 1:2500.

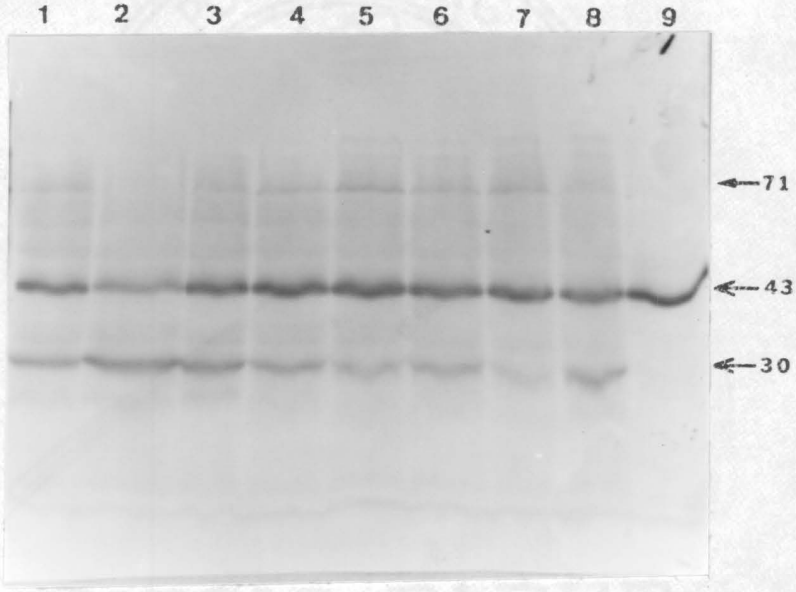


Figure 27. **Electrophoresis of RNA Purified from Fresh N. commune Cells:** RNA was visualized using ethidium bromide. Lane 1, rRNA marker of E. coli; lane 2, total RNA of N. commune with LiCl precipitation; lane 3, same as lane 2 but as much as 5 times of RNA was loaded; lane 4, same as lane 2 without LiCl precipitation; lane 5, same as lane 4 but as much as 5 times sample was loaded.

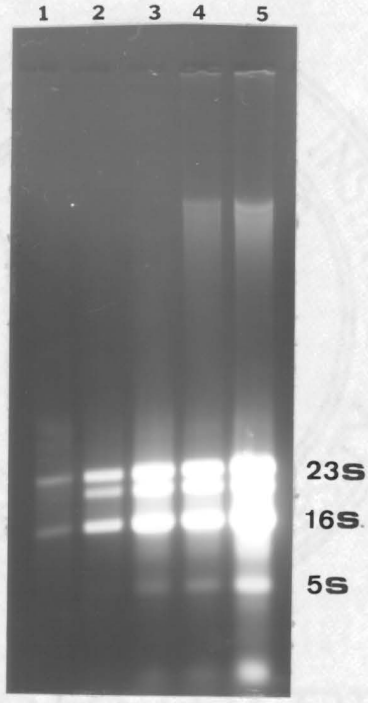


Figure 28. **RNA-RNA Hybridization**: Northern-blotted total RNA isolated from fresh N. commune cells hybridized with rpoC2 ribo-probe (A), with rpoC1 ribo-probe (B), and with rpoB ribo-probe (lanes C and D). D was same as C but electrophoresis was performed for longer time.

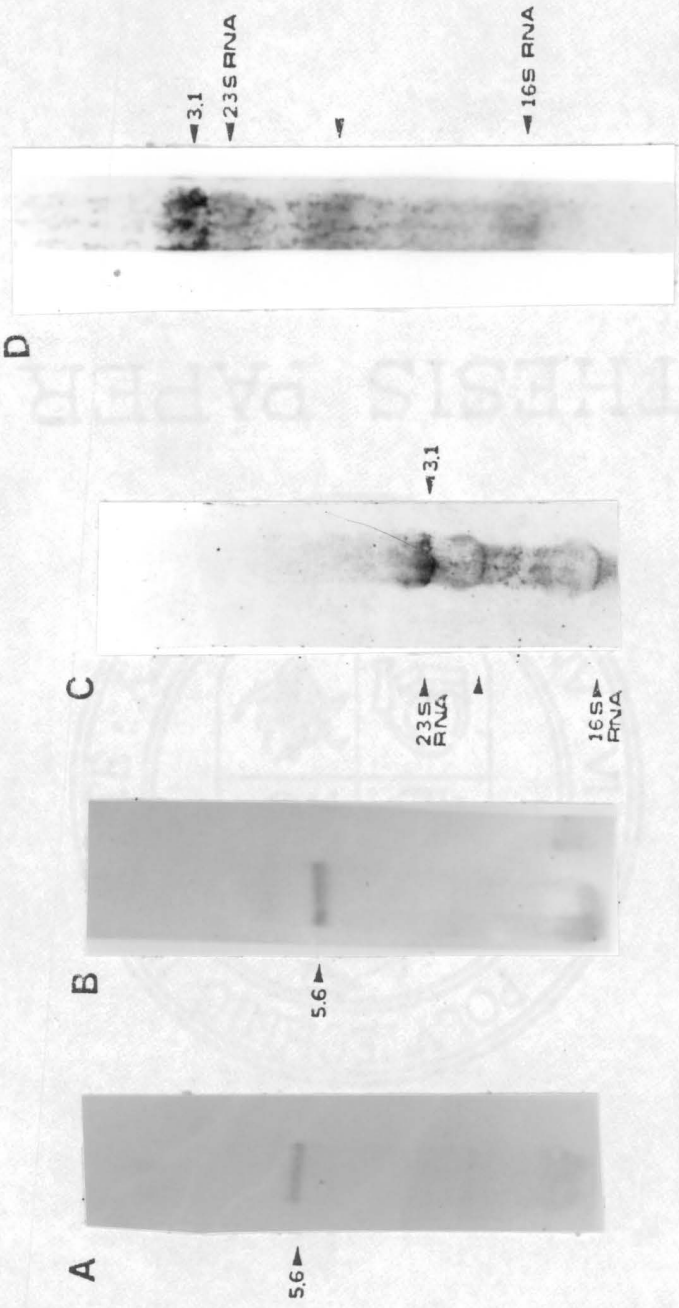
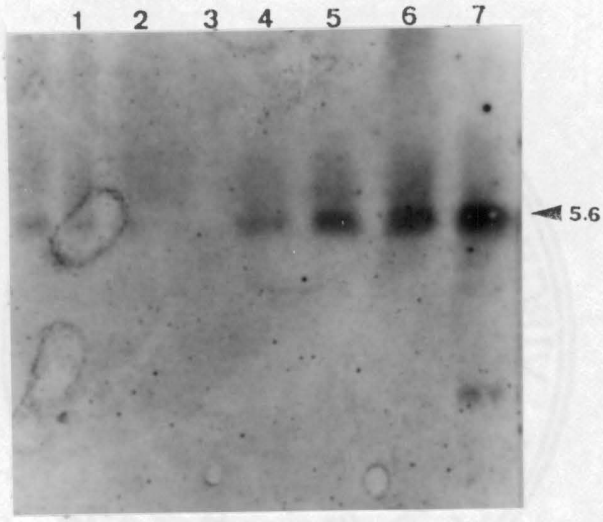

A
5.6
B
5.6
C
23S RNA
16S RNA
3.1
D
3.1
23S RNA
16S RNA

Figure 29. **Northern Analysis of Total RNA Isolated from Dried and Rewetted N. commune Cells:** The probe used here was ^{35}-labelled ribo-probe for rpoC2. A: Lane 1, 1-day dried; lane 2, 5-day dried; lane 3, 5-min rewetting; lane 4, 10-min rewetting; lane 5, 30-min rewetting; lane 6, 60-min rewetting; lane 7, fresh cells. **B: Relative Amounts of mRNA of rpo Genes of N. commune Cells under Different Conditions**: The picture was scanned (reflection scanning) using a Shimadzu CS9000 U. The RNA from fresh cells was used as a control (100%). Scanned areas of other bands are divided by the area of the band of RNA from fresh cells.


1 2 3 4 5 6 7
5.6

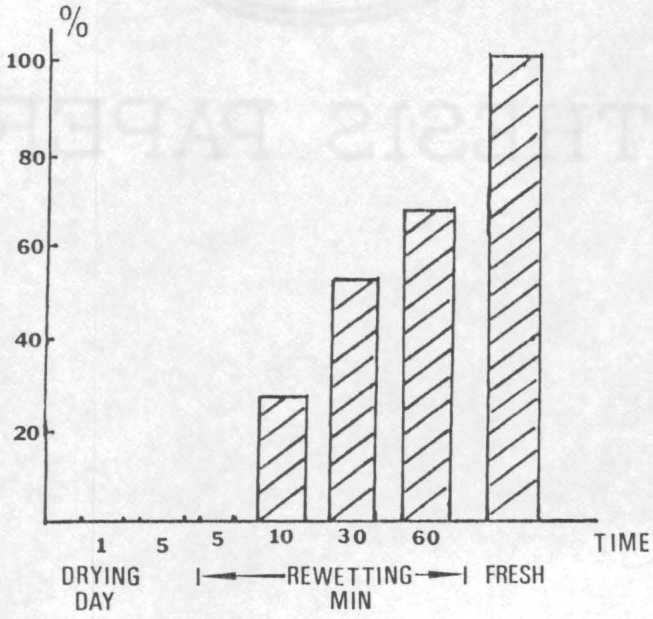
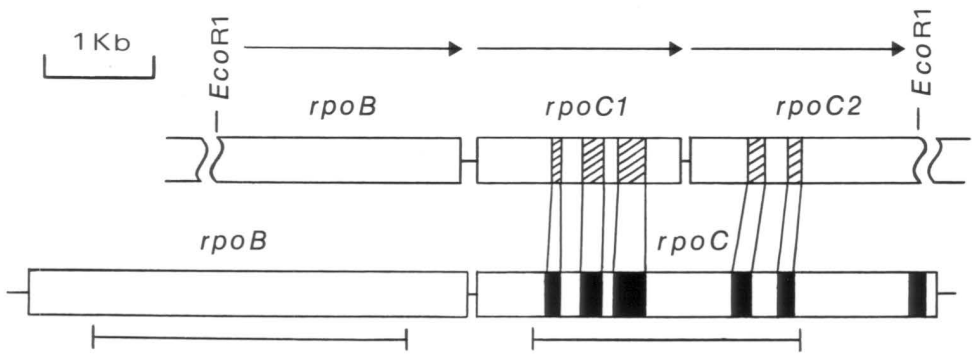

%
100
80
60
40
20
1 5 5 10 30 60
TIME
DRYING DAY
REWETTING MIN
FRESH

Figure 30. **Comparision of rpoBC Genes Between N. commune (upper) and E. coli (lower):** The cloned piece of N. commune rpoBC1C2 genes contains 5 out of the 6 most conserved regions between procaryotic and the eukaryotic RNA polymerase.

Attention Patron:

Page 101 repeated in numbering


1Kb
EcoR1
rpoB
rpoC1
rpoC2
EcoR1
rpoB
rpoC

DISCUSSION

Southern analysis revealed that the N. commune genome contains single copies of rpoB, rpoC1 and rpoC2 genes. The cloned 6.5-kb DNA fragment contained sequences corresponding to the C-terminal portion of rpoB, the complete rpoC1 gene, and the N-terminal portion of the rpoC2 gene. These three genes are linked in the order of rpoB, rpoC1 and rpoC2, and have the same orientation. The intergenic sequences between these three genes have great sequence similarity, are A-T rich (Figs. 12, 14, 15), and have the ability to form stable secondary structure (Fig. 19, 20), which may play an important role either in the transcription and/or the translation process. The intergenic sequences separating the corresponding rpo genes in chloroplasts are A-T rich as well. The significance of these sequences, if any, remains unknown.

The cyanobacterial RNA polymerase has a core structure of $\alpha_2\beta\gamma\beta'$. The genes corresponding to rpoC of E. coli are present as two genes, rpoC1 and rpoC2, in cyanobacteria. The derived amino acid sequences of rpoC1 and rpoC2 show extensive similarity, respectively, with sequences of the N-terminal and the C-terminal portions of the beta' subunit of RNA polymerase of E. coli. Therefore, the function of the beta' subunit of E. coli may reside on the gamma and beta' subunits in cyanobacteria. The gamma (71,000 M_r) and beta' (150,000, M_r) of cyanobacteria together are much larger than the beta' (160,000 M_r) of E. coli. The additional amino acids may provide additional sites for transcriptional regulation. The rpoC1 product (gamma subunit) contains three of six highly

conserved domains shared between the E. coli beta' subunit and the A subunit of eukaryotic RNA polymerase. The cloned part (about half of the full size) of rpoC2 contained another two additional conserved domains (Fig. 29).

Despite the similarity of rpo gene structure and organization between cyanobacteria and chloroplasts (Fig. 30), the larger subunits of RNA polymerase show greater similarity between cyanobacteria and E. coli than the similarity between cyanobacteria and chloroplasts. This may indicate functional differences in the RNA polymerase of endocyobiotic organelles.

Northern blotting revealed separate transcription of rpoBC1C2 genes. The size of the rpoB transcript is expected to be about 3.1 kb, based on the size of the corresponding subunit of RNA polymerase of Anabeana, if the gene is transcribed alone. In N. commune, the transcript corresponding to the rpoB was 3.1 kb which suggested that the rpoB was transcribed alone using its own promoter. The transcript corresponding the rpoC1C2 genes was the same and was about 5.6 kb, which indicated that the rpoC1C2 genes were transcribed with a common promoter in N. commune. The transcription of rpo genes of cyanobacteria is different from the transcription of the corresponding genes of E. coli and of chloroplasts. In chloroplasts of spinach, rpoBC1C2 genes are co-transcribed from the same promoter (Hudson et al., 1988). Using the promoter, one transcript is made from the rpoB gene and terminates at the first intron of the rpoC1 gene; and another transcript is made from the rpoB gene until the end of the rpoC2 gene. The rpoBC genes of E. coli are co-transcribed using promotor L10, together with rpoJ and rpoL genes. The rpo operons of archaebacteria also show different transcription (Zillig et al., 1989). In

Sulfolobus acidocaldarius, having a basic structure of BAC, the genes encoding B and A are co-transcribed, and the gene encoding the C subunit is transcribed by a separate promoter.

There were probably two reasons for the poor hybridization between N. commune rpoB ribo-probe and the Northern-blotted RNA. First, the estimated size of the rpoB transcript (3.1 kb), if it was transcribed alone, is close to the size of 23 S rRNA (3 kb). The massive amount of 23 S rRNA interferred with the separation of this RNA in the agarose gel, causing poor resolution of the band. Second, the massive amount of 23 S rRNA might block the hybridization between the probe and the RNA.

The in vivo expression products of rpoBC1C2 were assayed using heterologous antibodies directed against the core proteins of the RNA polymerase of Anabaena PCC 7120. Two polypeptides, with molecular weights of 94,000 and 71,000, were detected using the antibodies. The two polypeptides were believed to be the products of rpoC2 (incomplete) and rpoC1, respectively, according to their gene sizes. The rpoBC1C2 fragment was cloned out of frame (relative to lacZ) after the lacZ promoter at the multiple cloning site of pUC18. Using this promoter, a fusion transcript of lacZ interrupted by the transcripts of rpoBC1C2 was produced . The transcripts of rpoC1C2 were further translated using their own ribosomel binding sites, whereas, the rpoB (imcomplete) transcript was not translated because of lack of a ribosomel binding site and was not cloned in the frame relative to the lacZ coding sequence.

The cloned rpoC1 had the potential to encode a polypeptide (complete) of 70,200 daltons, which agreed closely with the size of 71,000, estimated by SDS-PAGE. Thus, the polypeptide of M_r of 71,000 probably was the product of rpoC1. The size of 94,000 M_r, estimated by SDS-PAGE, is larger than the predicated size from the sequence data (86,000 M_r) of rpoC2 (truncated). The rpoC2 was fused to the remainder of lacZ and caused the synthesis of a polypeptide of 89,600 M_r. The rest difference might reflect inaccuracy in M_r estimated by SDS-PAGE. This product could only be the product of rpoC2, but not of rpoB, since the cloned rpoB had no ribosomal binding site.

In the case of in vitro translation, a polycistronic transcript was made using T7 RNA polymerase. The second and third cistrons were further translated using their own ribosomal binding sites; the first one could not be translated due to lack of a ribosomal binding site in the cloned piece of DNA for this gene (the same result as found for in vivo expression). The intensity of the band, corresponding to the 94,000 M_r polypeptide was stronger than the one corresponding to the 71,000 M_r polypeptide. This was not due to a differential incorporation of ^{35}S-methionine since both peptides contained about the same numbers of methionine residues (15, 2.4% of total residues in gamma, and 14, 1.9% of total residues in the portion of beta'). The apparent higher level of expression of rpoC2 could be accounted for by its potential ribosomal binding site (AGGA, -10), whereas, rpoC1 has ribosomal binding site GGA (-15). The optimal consensus sequence for the ribosomal binding in E. coli is AGGAGG, located at about -6 bases upstream from the initiation codon. For the E. coli ribosomes, the first cyanobacterial rpo binding site could present a better one than the second one.

The expressed products of the rpoBC1C2 genes in vitro represented only a small portion of the total proteins in the translation system. Autoradiography of this sample, after being separated on a SDS-PAGE, indicated a major band of molecular weight of about 94,000, and a number of smaller bands. In addition, all these ^{35}S-labeled polypeptides could be precipitated with the antibodies, and were absent in the supernatant after immunoprecipitation (Fig. 23B, lane 4). The smaller polypeptides were likely from imcomplete translation of the transcripts. Coomassie staining of the SDS-PAGE of immunoprecipitated samples revealed two major polypeptides (Fig. 23, lane 5) which co-migrated with those two polypeptides produced in in vivo. A band with M_r of about 180,000 was present in all in vitro translation samples with or without the exogenous plasmid DNA (Fig. 22A, lanes 2-4; Fig. 23A, lanes 4-6). This might be the result of cross reaction between the antibodies and the beta and beta' subunits of E. coli RNA polymerase.

The different subunits of RNA polymerase of N. commune, when immobilized on a membrane by Western blotting, showed different cross reactivity with the antibodies used in this research (Fig. 22, lane 1). The alpha subunits had the strongest cross reactivity with the antibodies, whereas, the beta, gamma, and beta' subunits had very weak reaction with the same antibodies. Much effort was made to improve the transfer efficiency of larger polypeptides during Western blotting in order to rule out the possibility that the poor reaction was due to the poor transfer of the larger polypeptides. However, I was unable to improve the sensitivity. The in vitro expressed gamma and beta' subunits (^{35}S-labelled) showed almost non-detectable cross reaction with the antibodies (Fig. 22A, lanes 3, 4) even though the amounts of protein were more

than enough to be detected by Coomassie staining (Fig. 23A, lane 5). However, the same polypeptides have barely detectable reaction with the antibodies in the same experiment when cold non-labeled methionine was used instead the ^{35}S-labeled methionine, which obviously was due to more gamma and beta' proteins being present in this sample. This simply reflects that a better translation efficiency can be obtained using a non-labeled methionine than using a ^{35}S-labeled methionine. All evidence suggests a weak cross reaction between the immobilized larger subunits of RNA polymerase of N. commune and the antibodies directed against the core protein of RNA polymerase of Anabaena PCC7120. However, the same gamma and beta' subunits produced in vivo had reasonable cross reaction with the same antibodies (Fig. 20A, lane 5). The reason probably was that the amounts of polypeptides of gamma and beta' present in this sample were much higher than those present in the samples of the in vitro translation and in the total proteins isolated from the N. commune cells. This reflects that the lacZ promoter is a very strong promoter. Nevertheless, the gamma and beta' polypeptides had reasonable affinity for the antibodies when they were present in solution, since both could be precipitated completely with the same antibodies (Fig. 23B, lane 4). Immobilization, therefore, causes the weak cross reaction with the antibodies. The experiment using the antibodies to screen 15 different strains of cyanobacteria, when ^{125}I-label was used, also showed poor reactions with larger subunits of RNA polymerase of some strains (Schneider et al., 1987), while the alpha subunit had a strong reactivity.

Studies of RNA polymerase structure and gene organization provide interesting views of evolution. In eukaryotes it is believed that three different RNA polymerase enzymes arose from the duplication and diversification of a common ancestral gene within eukaryotes. However, recent evidence has shown that RNA polymerase II and III share a common stem with that of archaebacteria, whereas polymerase I bears more resemblance to that of eubacteria (Zillig et al 1989). This suggests coexistance of the three different RNA polymerases of the same age as the primitive eukaryotes.

My finding of splitting of rpoC1C2, corresponding to the single rpoC in E. coli, supports an hypothesis of divergent evolution of RNA polymerases. This event is also true in chloroplast and archaebacterial genomes. The difference of rpoC1C2 genes between chloroplast and cyanobacterial genomes is that there is an intron in rpoC1 gene of chloroplasts. It is not clear if a fission or a fusion is the cause of the difference between rpoC and rpoC1C2 genes. The splitting of the rpoC gene is an event present in prokaryotes only.

It is proposed that the chloroplast evolved from an ancestral form of a cyanobacterial-like organism. My findings give support to this proposal. Comparison of sequences of up- and down-stream rpo genes between cyanobacteria and chloroplasts should provide additional information about this hypothesis. If splitting is the cause, it should happen before the existence of the eukaryotic plant cell. However, the similarities between the rpo genes of cyanobacteria and other eubacteria is larger than those between cyanobacteria and chloroplasts.

Northern blot analysis of RNA samples, isolated from dried, rewetted and fresh N. commune cells, indicated that the synthesis of mRNA for the polymerase ceased upon drying, and preexisting mRNA was degraded as or after the cells were dried. The transcription of RNA polymerase genes gradually recovered when the cells were rewetted for more than 5 min (Fig. 28). The 10-min, 30-min, and 60-min rewetted Nostoc cells, after being dried for 5 days, contained 28%, 53%, and 66%, respectively, of the amount of the mRNA present in fresh cells. Though ribo-probe for rpoC2 was the only probe used in this experiment, it should represent all the mRNA for RNA polymerase as the fact that all subunits of this enzyme are produced stoichoimetrically. The actual amounts present in the rewetted cells could be higher than the numbers estimated here if the condition for rewetting was the same as the one for growing fresh cells (see Method). The fresh cells were grown in a fermenter, while the rewetting process was performed on nylon clothes.

The RNA polymerase enzyme itself, however, behaves differently from its transcripts. The proteins are much more stable under the same conditions. The results of Western blot and immunoanalysis showed that there was almost the same amount of alpha protein present in samples isolated from either dried, rewetted or fresh cells (Fig. 25). This result implies that RNA polymerase of N. commune survives the drying process, since different subunits of RNA polymerase are produced stoichiometrically. But there is a smaller peptide (about 30,000 M_r) present in all protein samples except the one purified from fresh cells. Thus, RNA polymerase indeed undergoes some degradation during the drying and rewetting processes. While I have no direct evidence to assess

from which subunit this fragment is derived, I suspect that this polypeptide is a degradation product of the alpha subunit, based upon the strong immuno-cross reaction with the antibodies.

The consensus sequences of the cyanobacterial promoters remain to be elucidated due to the general lack of S1 mapping information of upstream regions of cyanobacterial genes. The gene encoding indole phosphate hydrolase of N. commune can be expressed in E. coli using its own promotor (Xie et al., 1989b). It has been demonstrated that mRNA of cyanobacterial genes can be translated using their own ribosomal binding sites (Xie et al., 1989c). This suggests that cyanobacterial promoter and ribosomal binding sites share some consensus sequences with those of E. coli.

The use of phage vector instead of a plasmid vector for construction of the genomic library, was to overcome the background problem. A heterologous DNA probe from E. coli was used to screen a genomic library propogated in E. coli cells. It is clear that host cells would give background problems. But using the library made of a phage can minimize the background problem. During the lytic cycle of phage growth, a phage particle can proliferate 200 times and its host chromosomal DNA becomes degraded when new phage particles are released. In addition, gentle conditions can be used to release phage DNA while other non-infected host cells remain intact. All of these minimized the possible background problem, though minor backgrounds could still be seen over negative plaques (Fig. 3).

The use of exonuclease III and exonuclease VII makes it possible to construct overlapping deletion clones from a DNA fragment for subsequent DNA sequencing. Selection of those deletion clones carrying inserts with

desired sizes is tedious, especially when the original DNA fragment to be sequenced is large. Size analysis after purification of plasmid DNA from each deletion clone is also labor intensive. Therefore, a rapid and efficient procedure for screening of deletion clones was developed in this study (Xie and Potts, 1989). Using this procedure, unpurified plasmid DNA with a size difference of approximately 3.5%, could be resolved on an agarose gel (Fig. 9). The size of a plasmid DNA estimated by this procedure was smaller than the actual size which was determined after the plasmid DNA was purified and linearized.

REFERENCES

1. **Ahearn, J. M., Jr., M. S. Bartolomei, M. L. West, L. J. Cisek, and J. L. Corden.** 1987. Cloning and sequence analysis of the mouse genomic locus encoding the largest subunit of RNA polymerase II. J. Biol Chem. 262:10695-10705.
2. **Allison, L. A., M. Moyle, M. Shales, and C. J. Ingels.** 1985. Extensive homology among the large subunits of eukaryotic and prokaryotic RNA polymerase. Cell 42:599-610
3. **Angeloni, S. and M. Potts.** 1986. Polysome turnover in immobilized cells of N. commune (cyanobacteria) exposed to water stress. J. Bacteriol. 168:1036-1039.
4. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J.G. Seidman, J. A. Smith, and K. Struhl.** 1987. Current protocols in molecular biology. John Wiley & Son.
5. **Buzby, J. S., R. D. Porter, and S. E. Stevens, Jr.** 1985. Expression of the Escherichia coli lacZ gene on a plasmid vector in a cyanobacterium. Science (NY) 230:805-807
6. **Burgess, R. R.** 1969a. Separation and characterization of the subunits of ribonucleic acid polymerase. J. Biol. Chem. 244:6168-6176.
7. **Burgess, R. R.** 1969b. A new method for the large scale purification of E. coli deoxyribonucleic acid-dependent ribonucleic acid polymerase. J. Biol. Chem. 244:6160-6167.
8. **Cameron, R. E.** 1962. Species of Nostoc Vauch. occurring in the Sonoron desert in Arizona. Trans Am Microsc Soc 81:379-384.
9. **Cangelosi, G. A., C. M. Joanne, J. J. Rosen, and J. C. Meeks.** 1986. Cloning and expression of a Nostoc sp. leucine biosynthetic gene in Escherichia coli. Arch. Microbiol 145:315-321
10. **Carr, N. G. and B. A. Whitton.** 1982. The biology of cyanobacteria. Blackwell Scientific Publications, Oxford, London, Edinburgh, Boston, Melbourne, pp.688.
11. **Chambon, P.** 1975. Eukaryotic nuclear RNA polymerase. Annu. Rev.

Biochem. 44:613-638.

12. **Cho, K. W., Y. Khalili, K. Zandomeno, and R. Weunmann.** 1985. The gene coding for the large subunit of human RNA polymerase II. J. Biol. Chem. 260:15204-15210
13. **Curtis, S. E. and R. Haselkorn.** 1984. Isolation, sequencing and expression of two members of the 32 Kd thylakoid membrane protein gene family from cyanobacterium Anabaena PCC7120. Plant Mol. Biol. 3:249-258.
14. **Defrancesco, N. and M. Potts.** 1988. Cloning of nifHD from N. commune UTEX 584 and of a flanking region homologous to part of Azotobacter vinelandii nifU gen. J. Bacteriol. 170:3297-3300
15. **Dennis, P. P., V. Nene, and R. Glass.** 1985. Autogenous post-translational regulation of RNA polymerase Beta and Beta' subunits synthesis in E. coli. J. Bacteriol. 161:803-806.
16. **Falkenburg, D., B. Dworniczak, D. M. Faust and E. K. Bautz.** 1987. RNA polymerase II of Drosophila : Relation of its 140,000 M_r subunit to the Beta subunit of Escherichia coli RNA polymerase. J. Mol. Biol. 195:929-937.
17. **Faust, D. M., R. Renkawitz-Polh, D. Falkenburg, A. Gasch, S. Bialojan, R. A. Yougn, and E. K. Bautz.**(1986). Cloning and identification of the genes coding for the 140-kd subunit of Drosophila RNA polymerase II. EMBO J. 5:741-746.
18. **Fourney, R. M., J. Miyakoshi, R. S. Day, III, and M. C. Paterson.** 1988. Northern blotting: efficient RNA staining and transfer. Focus 10:1:5-7.
19. **Golden, J. W., S. J. Robinson, and R. Haselkorn.** 1985. Rearrangememt of nitrogen fixation genes during heterocyst differentiation in the cyanobacterium Anabeana. Nature (London) 314:419-423.
20. **Golden, S. S., J. Brusslan, and R. Haselkorn.** 1982. Expression of a family of psbA genes encoding a photosystem II polypeptide in the cyanobacterium Anacystis nidulans R2. EMBO 5:2789-2798
21. **Grossman, A. D., J. W. Erickson, and G. A. Gross.** 1984. The htpR gene product of E. coli is a sigma factor for the heat-shock promoters. Cell 38:383-390.
22. **Gorden, J., D. L. Cadena, J. M. Ahearn and M. E. Dahmus** 1985. A unique structure at the carboxyl terminus of the largest, subunit of eukaryotic RNA

polymerase II. Proc. Nat. Acad. Sci., U.S.A. 82:7932-7938.

23. **Greenburg, B. M., J. O. Narita, C. Deluca-Flaherty, W. Gruissem, K. A. Rushlow, and R. B. Hallic.** 1984. Evidence for two RNA polymerase activities in Euglena gracillis chloroplasts. J. Biol. Chem. 259:14880-14887.
24. **Haldenwang, W. G., N. Lang, and R. Losick.** 1981. A sporulation induced sigma like regulatory protein from Bacillus subtilis. Cell 23:615-624.
25. **Haldenwang, W. G. and R. Losick.** 1980. Novel RNA polymerase sigma factor from Bacillus subtilis. Proc. Natl. Acad. Sci. USA 77:7000-7004.
26. **Hanahan, D.** 1985. Techniques for transformation of E. coli. In: DNA cloning, vol. I, A practical approach. (ed. D. M. Glover) JRL press.
27. **Herrero A. and P. Wolk.** 1986. Genetic mapping of the chromosome of the cyanobacterium Anabaena variabilis. Proximity of the structural genes for nitrogenase and ribulose-bisphosphate carboxylase. J. Biol. Chem. 261:7748-7754
28. **Herzfeld, F. and W. Zillig.** 1971. Subunit composition of DNA-dependent RNA polymerase of Anacystis nidulans. Eur. J. Biochem. 24:242-248.
29. **Holn, B.** 1979. In vitro packaging of lambda and cosmid DNA. Methods in Enzymology 68:299-309.
30. **Home, M., M. Tanaka, J. Chunwongse, K. Shinozaki, and M. Sugiura.** 1986. A tobacco chloroplast DNA sequence possibly coding for a polypeptide similar to E. coli RNA polymerase beta subunit. FEBS 200:87-90.
31. **Houmard, J., D. Mazel. C. Moguet, D. A. Bryant and N. T. De Marsac.** 1986. Organization and nucleotide sequence of genes encoding core components of the phycobilisomes from Synechococcus 63301. Mol. Gen. Genet. 205:404-410
32. **Hudson, G. S., T. A. Holton, P. R. Whitfeld, and W. Bottomley.** 1988. Spinach chloroplast rpoBC genes encode three subunits of the chloroplast RNA polymerase. Mol. Biol. 200:639-654.
33. **Ingles, C. J., J. Biggs, J. K-C Wang, J. R. Weeks, and A. L. Greenleaf.** 1983. Identification of a structure gene for a RNA polymerase II polypeptide in Drosophila melanogaster and mammalian species. Proc. Natl. Acad. Sci. USA 80:3396-3400.
34. **Ingles, C. J., H. J. Himmelfarb, M. Shales, A. L. Greenleaf, and J. D.**

Friesen. 1984. Identification, molecular cloning, and mutagenesis of Saccharomyces cerevisiae RNA polymerase genes. Proc. Natl. Acad. Sci. USA 81:2157-2161.

35. **Jager K. and M. POtts.** 1988. In vitro translation of mRNA from Nostoc commune (cyanobacteria). Arch Microbiol 149:225-231.
36. **Leffers, H., F. Gropp, F. Lottspeich, W. Zillig, and R. Garrett.** 1989. Sequence, organization, transcription and evolution of RNA polymerase subunit genes from the archaebacterial extreme Halobacterum halobium and Halococcus morrhuae. J. Mol. Biol. 206:1-17.
37. **Lewis, M. K., and R. R. Burgess.** 1982. In The Enzymes, Nucleic acids, part B (Boyer, P. B., ed) Vol. XV, pp.117-128, Academic Press, New York.
38. **Little. M. C. and R. Hollick.** 1988. Chloroplast rpoA, rpoB, and rpoC genes specify at least three components of a chloroplast DNA-dependent RNA polymerase active in tRNA and mRNA transcription. J. Biol. Chem. 263:14302-14307.
39. **Maniatis, T., E. F. Fritsch and J. Sambrook.** 1982. Molecular cloning, A laboratory manual. (Cold Spring Harbor laboratory)
40. **Miller, S. S., F. M. Ausubel, and L. Bogorad.** 1979. Cyanobacterial ribonucleic acid polymerase recognize lambda promoter. J. Bacteriol. 140:246-250
41. **Mazur, B. J., D. Rice, and R. Haselkorn.** 1980. Identification of blue-green algal nitrogen fixation genes by using heterologous DNA hybridization probes. Proc. Natl. Acad. Sci. USA 77:186-190
42. **Murphy, C. C., D. A. Bryant, R. D. Porter, and N. T. De Marsac.** 1987. Molecular cloning and characterization of the recA gene from the cyanobacterium Synechococcus sp. strain PCC 7002. J. Bacteriol. 169:2739-2747
43. **Nomura, T., A. Ishihama, M. Kajitani, T. Takahashi, N. Nakeda, and K. Yoshinaga.** 1984. Promoter selectivity of E coli RNA polymerase. II. Altered promoter selection by mutant holoenzymes. Mol. Gen. Genet. 193:8-16.
44. **Ohme, M., M. Tanaka, J. Chunwongse, K. Shinozaki, and M. Sugiura.** 1986. A tobacco chloroplast DNA sequence possibly coding for a

polypeptide similar to E coli RNA polymerase beta subunit. FEBS Lett. 200:87-90.

45. **Ohyama, K., T. Kohchi, T. Sanoo, and Y. Yamada.** 1988. Newly identified groups of genes in chloroplasts. TIBS 13:19-22.
46. **Olie, J. J. and M. Potts.** 1986. Purification and biochemical analysis of the cytoplasmic membrane from the desiccation-tolerant cyanobacterium N. commune UTEX 584. Appl. Environ Microbiol., 52:706-710.
47. **Ovchinnikov, Yu. A., G. S. Monastyrskaya, V. V. Gubanov, S. O. Guryev, I. S. Salomatina, T. M. Shuvaea, V. M. Lipkin, and E. D. Sverdlov.** 1982. The primary structure of E. coli RNA polymerase. Nucleotide sequence of the rpoC gene and amino acid sequence of the beta' subunit. Nucl. Acids Res. 10:4035-4044.
48. **Potts, M. and M. Bowman.** 1985. Sensitivity of N. commune UTEX 584 (cyanobacteria) to water stress. Arch. Microbiol. 141:51-65.
49. **Potts, M.** 1986. The protein index of N. commune UTEX 584 (cyanobacteria): changes induced in immolilized cells by water stress. Arch. Microbiol. 146:87-95
50. **Puhler, G., H. Leffers, F. Gropp, P. Palm, H. P. Klenk, F. Lottspeich, R. A. Garrett, and W. Zillig.** 1989. Archaebacterial DNA-dependent RNA polymerase testify to the evolution of eukaryotic nuclear genome. Proc. Natl. Acad. Sci. USA in press
51. **Rice, D., B. Mazur, and R. Haselkorn.** 1982. Isolation and physical mapping of nitrogen fixation gene from the cyanobacterium Anabaena 7120. J. Biol. chem. 257:13157-13163.
52. **Riva, M., S. Memet, J. Y. Micouin, J. Huet, I. Treich. J. Dassa, R. Yougn, J. M. Buhler, A. Sentenac, and P. Fromageot.** 1986. Isolation of structure genes for yeast RNA polymerase by immunological screening. Proc. Natl. Acad. Sci., USA 83:1554-1558.
53. **Roeder, R. G.** 1976. In RNA polymerase (Losick, R & Chamberlin, M., eds), pp. 285-329. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

54. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA 71:1342-1346.

55. **Schallenberg, J., M. Mose, M. Truss. W. Reiser, M. Thomm, K. O. Stetter, and A. Klein.** 1988. Cloning and physical mapping of RNA polymerase genes from Methanobacterium thermoautotrophicum and comparison of homologies and gene orders with those of RNA polymerase genes from other methanogenic archaebacteria. J. Bacteriol. 170:22247-2253.

56. **Scherer, S., A. Ernst, T-W. Chen, and P. Boger.** 1984. Rewetting of drought-resistant blue-green algae: time course of water uptake and reappearance of respiration, photosynthesis, and nitrogene fixation. Oecologia (Berlin), 62:418-423.

57. **Scherer, s. and M. Potts.** 1989. Novel water stress proteins from a desiccation-tolerant cyanobacterium. J. Biol. Chem. 264:12546-12553.

58. **Schneider, G. J. and R. Haselkorn.** 1988. RNA polymerase subunit homology among cyanobacteria, other eubacteria and archaebacteria. J. Bacteriol. 170:4136-4140.

59. **Schneider, G. J., N. E. Tumer, C. Richaud, G. Borbely, and R. Haselkorn.** 1987. Purification and characterization of RNA polymerase from the cyanobacterium AnabaenaPCC7120. J. Biol. Chem. 262:14633-14639.

60. **Searles, L. L., R. S. Jokerst, P. M. Bingham, R. A. Voelker, and A. L. Greenleaf.** 1982. Molecular cloning of sequences from a Drosophila RNA polymerase II locus by P element transposon tagging. Cell 31:585-592.

61. **Sentenac, A.** 1985. Eukaryotic RNA polymerases. CRC Crit. Rev. Biochem. 18:31-91.

62. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Preparation of SDS protein extracts (in: Experiments With Gene Fusions) Cold Spring Harbor laborotory Press, Cold Spring Harbor, NY.

63. **Williams, J. G. and P. J. Mason.** 1985. Hybridization in the analysis of RNA. in: Nucleic acid hybridization, a practical approach. eds. B. D. Hames & S. J. Higgins. IRL Press Inc., Oxford, Washington DC.

64. **Woese, C. R.** 1987. Bacterial evolution. Microbiol. Rev. 51, 221-271.

65. **Xie, W. Q. and M. Potts.** 1989a. Quick screening of plasmid deletion clones carrying inserts of desired sizes for DNA sequencing. Gene Anal. Techn.

6:17-20.

66. **Xie, W. Q., B. A. Whitton, J. W. Simon, K. Jager, D. Reed, and M. Potts.** 1989b. Nostoc commune UTEX 584 genes expressing indole phosphate hydrolase activity in E coli. J. Bacteriol. 171:708-713.
67. **Xie, W. Q., K. Jager, and M. Potts.** 1989c. Cyanobacterial RNA polymerase genes rpoC1 and rpoC2 correspond to rpoC of Escherichia coli. J. Bacteriol. 171:1967-1973
68. **Yang, R., J. Lis and R. Wu.** 1979. Elution of DNA from agarose gels after electrophoresis. in: Methods In Enzymology 68:176-183
69. **Yannish-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103-119.
70. **Young, R. A. and R. W. Davis.** 1983. Yeast RNA polymerase II genes: isolation with antibody probes. Science 222:778-782.
71. **Zillig, W. and K. O. Stetter.** 1985. Distinction between the transcription system of archaebactria and eubacteria. in: The Bacteria Vol. VIII (I. C. Gunsalus, J. R. Sokatch, L. N. Ornston, C. R. Woese and Wolfe, R. S. eds.) pp.525-538.
72. **Zillig, W., H. P. Klenk, P. Palm, H. Leffers, G. Puhler, F. Gropp, and R. A. Garrett.** 1989. Did eukaryotes originate by a fusion event? Endocyt. C. Res. in presss
73. **Zillig, W. and F. Gropp.** 1987. The distribution of antigenitic determinants in the B components of DNA-dependent RNA polymerase of archaebacteria, eubacteria and eukaryotes suggests the primeval character of the extremely thermolphilic archaebacteria. in: RNA polymerase and the regulation of transcription. A Steenbock symposium. eds by W. S. Reznikoff et al.. Elsevier.
74. **Zillig, W., H-P. Klenk, P. Palm, G. Puhler, and F. Gropp.** 1989. The phylogenic relations of DNA-dependent RNA polymerase of archaebacteria, eukaryotes, and eubacteria. Can. J. Microbiol. 35:73-80.

APPENDIX

Appendix 1

Buffer 1: 0.1 M Tris-HCl, pH 7.5, 0.15 M NaCl

Buffer 3: 0.1 M Tris-HCl, pH 9.5, 0.1 M NaCl, and 50 mM $MgCl_2$

Cracking buffer: 62.5 mM Tris-HCl, pH 6.8, 2% (w/v) SDS, 10% (w/v) glycerol, 5% (w/v) 2-mercaptoethanol, and 0.001% bromophenol blue

50x Denhardt's solution: 1% (w/v) BSA, 1% (w/v) Ficoll,and 1% w/v Polyvinylpolypyrrolidone

DNA hybridization buffer: 5xSSC, 1xDenhardt's solution , 1 mM EDTA, 10 mM Tris-HCl (pH 7.5), 0.1% SDS, and 30% de-ionized formamide

Exonuclease III buffer: 50 mM Tris-HCl, pH 8.0, 5 mM $MgCl_2$, 1 mM EDTA

Exonuclease VII buffer: 1x = 50 mM KPO_4, pH7.0, 8 mM EDTA, and 1 mM DTT

LB medium: 10 g Bacto-tryptone, 5 g Bacto-yeast extract, 10 g NaCl, pH 7.5 in 1 liter

Lysis buffer: 2.5 mM EDTA, 2% (w/v) SDS, 5% (w/v) sucrose, and 0.04% (w/v) bromophenol blue

Packaging buffer: 6 mM Tris-HCl, pH 8.0, 50 mM spermidine, 50 mM putrescine, 20 mM $MgCl_2$, 30 mM ATP and 30 mM 2-mercaptoethanol

Protoplasting buffer: 20 mM Tris-HCl, pH 7.5, 5 mM EDTA, 50 mM NaCl, 20% w/v sucrose, 100 ug/ml RNase A, and 50 ug/ml lysozyme

RNA extraction buffer: 0.25 M sucrose, 0.2 M NaCl, 0.1 M Tris-HCl, pH 9.0, and 10 mM Mg acetate, 5 mM DTT*.

RNA hybridization buffer:5xSSC, pH 7, 50 mM $NaPO_4$, pH 6.5, 0.5% (w/v) SDS, 4.5x Denhard's solution, and 55% (v/v) deionized formamide

Sonication buffer: 20 mM Tris-HCl, pH 8, 1 mM EDTA, 3 mM $MgCl_2$, and 5 mM 2-mercaptoethanol

SSC (1x): 0.15 M NaCl, 15 mM sodium citrate, pH 7.0)

SOB: 2% (w/v) Bacto tryptone, 0.5% (w/v) Bacto Yeast extract, 10 mM NaCl, 0.5 mM KCl, 10 mM $MgCl_2$ and 10 mM $MgSO_4$

* Just before use the buffer (without DTT) was boiled for several min, cooled, then DTT was added.

Appendix 2

strains	characteristics	source
Nostoc commune UTEX584	(Het^+ Nif^+)	M. Potts
BHB 2688	F^- recA λ^r(λEam4b2, red3, imm434, cIts, sam7	BRL
BHB 2690	F^- recA, λ^r(λDam15b2, red3, imm434, cIts, Sam7	BRL
DH5	F^-, endA1, hsdR17, supE44, thi-1, lbd(-), recA1, gyrA96, relA1, phi80dlacZ(d)M15	BRL
C600	thi-1, thr-1, leuB6, lacY1, tonA21, supE44, λ^-	Promega
C600hflA	C600 hflaA150::Tn10	Promega
JM109*	$recA^-$,recA1, endA1, gyrA96, thi, hsdR17, supE44,relA1, λ^-,Δ(lac-proAB), {F', traD36, $proAB^+$, $lacI^Q$, lacZΔM15}	T. Larson
Plasmids:		
pGEM-4	2870 bp, amp^r, T7 and SP6 promoters flanking by a pUC18 derived multiple cloning region	Promega
pUC18	2686 bp, amp^r, $lacZ^+$	BRL
pPD489	pBR322 + 2.6 kb EcoRI fragment of rpoC of E. coli	P. Dennis
pPD490	pBR322 + 2.8 kb EcoRI fragment of rpoB of E. coli	P. Dennis
pWQX001, pWQX005	6456 bp EcoRI-EcoRI rpoBC1C2 of N. communein pGEM-4	This study
pX1	6456 bp EcoRI-EcoRI rpoBC1C2 of N. commune in pGEM-4	This study

Appendix 2 cont.

Phage:		
λ^{gt10}	b527, srl $\lambda 3^{o}$, imm^{434}, srl $\lambda 4^{o}$, srl $\lambda 5^{o}$,	Promega

* The strain used here lacked F' factor.

Appendix 3

RESTRICTION-SITE ANALYSIS
Done on DNA sequence NCRPOB
Total number of bases: 6455.
Analysis done on the complete sequence.

List of cuts by enzyme

Aat I : 5647
Acc I : 959
Acy I : 1386 3633 5644
Afl II : 2342
Afl III : 804 3761 5739
Aha III : 773 3157
Alu I : 1114 1764 2682 2712 3011 3125 3443 3488 3546 4210 4629
4678 4970 5181 5253 5300 5387 5895 6078 6167 6279
Apa LI : 4705
Ase I : 4248 4283
Asp 718I : 4080 5292 6192
Asu I : 91 187 388 1015 1807 2538 3004 3435 3746 4375 4996
5573 5791
Ava II : 91 388 2538 3746 5573
Bbe I : 3636
Bbv I : 4533 5097
Bbv I* : 1896 2869 3244 3533 5044 5168
Bbv II : 335 2210 2860 5409 6265
Bgl I : 2563
Bin I : 44 549 2538 6212
Bin I* : 1782 4758 6393
Bsp MI : 2363
Bss HII : 1209

Appendix 3 (continued)

Bst NI : 95 350 1313 1682 2012 2791 4165 4610 4932 5290 5788 5902 5915 5955 6046 6376 6442

Bst XI : 51 1746 2727 5325 6318

Cau II : 633 1334 3910 6036 6357

Cfr 10I : 1422

Cfr I : 6263

Cla I : 4603 4783 5371

Cvi JI : 17 188 209 245 307 319 829 836 905 997 1017 1114 1318 1537 1750 1764 1808 1900 2149 2392 2397 2471 2501 2515 2656 2682 2712 2965 3006 3011 3125 3437 3443 3488 3546 3789 3813 3837 3865 3885 3966 4210 4377 4524 4542 4629 4678 4877 4890 4970 4998 5181 5212 5253 5300 5387 5454 5477 5746 5793 5859 5895 6039 6078 6162 6167 6220 6265 6279 6374

Dde I : 935 1826 2224 2245 2260 2281 2335 2784 2826 4356 4501 4514 4674 4971 5107 5657 5948 5999 6079

Dpn I : 38 296 543 1789 1984 2532 2836 3852 4765 5162 5823 6206 6400

Dra II : 187

Dra III : 3617

Dsa I : 1389 3838 5426

Eco 31I* : 2345

Eco 47III : 533 2606 2917

Eco 57I : 1780 2932 3811 5421 6232

Eco 78I : 3634

Eco NI : 899

Eco RI : 1

Eco RII : 93 348 1311 1680 2010 2789 4163 4608 4930 5288 5786 5900 5913 5953 6044 6374 6440

Eco RV : 1122 2939 3559 6239 6302

Esp I : 4971

Fnu 4HI : 1910 2671 2883 3258 3547 4522 4731 4958 5058 5086 5182
Appendix 3 (continued)
5269 5735 5784
Fnu DII : 110 116 228 651 695 1211 1213 1262 1289 2045 2192 2743
3563 3797 4171 4400 5741
Fok I : 648 964 1087 1510 1792 2523 3083 3266 3285 3597 3966 4089
4757 4935
Fok I* : 128 3691 5047 6336 6420
Gsu I : 371 5234 5359
Gsu I* : 287 2003 4147
Hae II : 535 2608 2911 2919 3636
Hae III : 188 1017 1808 3006 3437 4377 4998 5793 6265
Hga I : 167 529 688 1394 3310 4131
Hga I* : 574 684 2628
Hgi AI : 4709
Hgi CI : 3632 4080 4490 4851 5292 6192 6282
Hha I : 118 228 534 615 1067 1211 1213 1851 2045 2607 2910 2918
3112 3551 3635 3797 4400 5351 5884
Hin P1I : 116 226 532 613 1065 1209 1211 1849 2043 2605 2908 2916
3110 3549 3633 3795 4398 5349 5882
Hind II : 234 3404 4353
Hind III : 4627 5385
Hinf I : 284 498 628 1085 1155 1597 1824 2023 2179 2194 2577 2628
2786 2874 3095 3518 3695 4029 4111 4893 5009 6118 6405
Hpa I : 234
Hpa II : 405 632 1333 1423 3910 4041 4702 5064 6035 6356
Hph I : 673 811 865 1129 1212 1260 1833 2469 4032 5259 5502 5816
5826 5975 6380
Hph I* : 1448 2615 2637 3444 3823 4898 5924
Kpn I : 4084 5296 6196
Mae I : 146 865 2006 2113 2709 3126 3469 3862 3948 4203 4621 5016
Mae II : 452 750 804 1047 1503 2380 2408 2686 3371 3416 3763 4006
5186 5606 5644 5987

Mae III : 148 666 1137 1821 2462 2621 2629 2643 2691 3829 4061 4554
Appendix 3 (continued)
4713 4823 5066 5495 5963 6332
Mbo I : 36 294 541 1787 1982 2530 2834 3850 4763 5160 5821 6204
6398
Mbo II : 337 340 928 1033 1185 1441 1606 2215 2844 2865 2883 2985
3606 4539 4542 5061 5132 5414 5445 6225 6270
Mbo II* : 3934 6296
Mlu I : 5739
Mme I : 779 3370 3401
Mme I* : 1052 1475 1486 2134 2708 2824
Mnl I : 59 259 311 794 849 975 1534 2171 2174 3718 3778 4777
5023 5401 5451 6042 6363
Mnl I* : 2581 2887 3349 3943 4066 4556 5101 5443 5506 5581 6439
Mse I : 22 103 233 725 772 928 1475 1562 1940 1975 1988 2030
2252 2293 2305 2343 2482 2808 3016 3156 3181 3508 3727
3894 4249 4284 4365 4818 5638 6028 6084 6229
Nar I : 3633
Nla III : 36 226 315 498 517 552 894 1621 1695 1758 3965 5237
5309 5468
Nla IV : 12 189 837 1016 1317 3056 3634 3747 3884 4051 4082 4492
4853 4997 5294 5476 5792 6194 6284
Nru I : 1289 2192
Nsp BII : 1866
Nsp HI : 552 894 1758
Pfl MI : 3580 3872 6272
Ple I : 292 1605 2585 2636 2794 4119
Ple I* : 622 1818 4887
Pss I : 190
Pvu I : 1985 5163 5824
Rsa I : 241 470 600 654 730 745 947 1274 1472 1488 1556 1588
1710 2078 3121 3760 3936 4082 4088 4138 4394 4619 4873
5294 5378 5420 5609 5912 6194

Sca I : 3936

Appendix 3 (continued)

Scr FI : 95 350 633 1313 1334 1682 2012 2791 3910 4165 4610 4932 5290 5788 5902 5915 5955 6036 6046 6357 6376 6442

Sdu I : 87 380 4709

Sec I : 87 895 1389 1480 1547 2989 3838 4930 5426 5528 5597 5787 5900 5914 6375

Sfa NI : 3713 5069 5470

Sfa NI* : 909 3244 3263 3610 4735 5426 5807

Sna BI : 453 2381

Spe I : 4620

Sso II : 93 348 631 1311 1332 1680 2010 2789 3908 4163 4608 4930 5288 5786 5900 5913 5953 6034 6044 6355 6374 6440

Ssp I : 1514 3824

Sty I : 87 895 1480 1547 2989 5528 5597

Taq I : 251 544 2166 2749 3170 4109 4603 4783 4979 5371

Taq IIA : 3628

Taq IIA* : 1828 5235

Taq IIB : 1742

Tth111II : 575 1019 2111 2832 5246

Tth111II* : 2919 3568

Xho II : 1787 2530 4763 6204

Xmn I : 2410 4534 5686

Total number of cuts is : 793.

Appendix 4

One - and three-letter codes for amino acids

Alanine	Ala	A	Leucine	Leu	L
Arginine	Arg	R	Lysine	Lys	K
Asparagine	Asn	N	Methionine	Met	M
Aspartic acid	Asp	D	Phenylalanine	Phe	F
Cysteine	Cys	C	Proline	Pro	P
Glutamic acid	Glu	E	Serine	Ser	S
Glutamine	Gln	Q	Threonine	Thr	T
Glycine	Gly	G	Tryptophan	Trp	W
Histidine	His	H	Tyrosine	Tyr	Y
Isoleucine	Ile	I	Valine	Val	V

The vita has been removed from the scanned document