Plant Molecular Biology 2: 279–290 (1983).

Synthesis of the large subunit of ribulose-1,5-bisphosphate carboxylase in an *in vitro* partially defined *E. coli* system

Jack L. Erion[1], Joseph Tarnowski[2], Susan Peacock, Paul Caldwell, Betty Redfield, Nathan Brot & Herbert Weissbach
Roche Institute of Molecular Biology, Roche Research Center, Nutley, NJ 07110, U.S.A.
[1] *ARCO Plant Cell Research Institute, Dublin, CA 94566, U.S.A.*
[2] *Present Address: Hoffman-La Roche, Inc., Nutley, NJ 07110, U.S.A.*

Keywords: spinach chloroplast DNA, gene expression, initiation of transcription

Abstract

The *in vitro* DNA- or RNA-directed synthesis of the large subunit (LS) of spinach chloroplast ribulose-1,5-biphosphate carboxylase (RuP_2C) has been examined in a highly defined *E. coli* transcription-translation system. Spinach chloroplast DNA, RNA and recombinant plasmids containing the spinach chloroplast LS gene (*rbc*L) have been used as templates in the *in vitro* system and a quantitative assay has been developed to measure LS formation. The *in vitro* formed product contains formylmethionine at the N-terminal position and sediments primarily as a monomer. There is no detectable enzymatic activity associated with the *in vitro* product. To determine where the *E. coli* RNA polymerase used in these systems initiates, we have examined the transcripts produced by this enzyme *in vitro*. Measurements of run-off transcripts indicate that *E. coli* RNA polymerase initiates at the same position on the gene as is seen *in vivo*. In addition, the complete nucleotide sequence of the *rbc*L gene including previously unsequenced 3′ and 5′ flanking regions has been determined. The sequence agrees, except at two nucleotide positions, with previously published sequencing data for this gene (Zurawski, G, Perrot, B, Bottomley, W, Whitfeld, PR, 1981. Nucleic Acids Res. 9:3251–3270).

Abbreviations

LS, large subunit
RuP_2C, ribulose-1,5-bisphosphate carboxylase
*rbc*L, gene for LS of RuP_2C
EF, elongation factor
IF, initiation factor
cp, chloroplast
$NaDodSO_4$, sodium dodecyl sulfate
MDPF 1,2-methoxy-2,3-diphenyl-3-(2H)-furanone
RuP_2, ribulose-1,5-bisphosphate
PVP-40, polyvinylpyrrolidone-40 000 MW
DNAse, Deoxyribonuclease 1
EDTA, ethylene-diaminetetraacetic acid
PMSF, phenylmethylsulfonyl fluoride

Introduction

Recently, our laboratory used a highly defined DNA-directed *in vitro* protein synthesis system to study gene expression in *E. coli* (17–21, 41). A major goal of these studies was to obtain the expression of bacterial genes in a system containing highly purified and defined protein components. This system has not only been useful in identifying new factors required for transcription and translation but also in identifying factors involved in regulating the expression of individual operons. A defined transcription-translation system, containing more than thirty highly purified factors has been described for the DNA-dependent synthesis of β-galactosidase (20) and this same system has been used to study the *in vitro* synthesis of transcription and translation factors coded for on λ*rif*[d]18 and λ*fus*3 phages (41).

Although a partially defined system with *E. coli* factors cannot be used to examine the expression of nuclear genes, protein synthesis in eukaryotic organelles such as mitochondria and chloroplasts bears many similarities to the prokaryotic system. In fact, the chloroplast gene (*rbc*L) for the large subunit (LS) of ribulose bisphosphate carboxylase (RuP_2C) has been expressed *in vivo* in *E. coli* tranformed with a plasmid containing the LS gene (11), *in vitro* using crude *E. coli* extracts (4, 25, 40) and in a heterologous system containing *E. coli* RNA polymerase and a reticulocyte lysate (8). The ability to synthesize the LS in an *E. coli* system may not be surprising, especially when one examines the nucleotide sequence of this gene (42). The basic structural features, which include a 'Shine-Delgarno' sequence (33), a 'Pribnow box' (30), and a '-35' region (26) are all apparently present (42). Whether an *E. coli* system or some heterologous mixture of *E. coli* and chloroplast factors could duplicate regulatory functions *in vivo*, however, is uncertain. In order to examine this possibility we have initiated studies on the *in vitro* synthesis of the LS in a highly defined *E. coli* system using, as templates, spinach chloroplast DNA (cpDNA) and RNA (cpRNA) as well as recently constructed plasmids containing the spinach LS gene (10). Characteristics of the *in vitro* system, and some properties of the *in vitro* products are presented. In addition, the specificity of the interaction of *E. coli* RNA polymerase with this gene is examined.

Materials and methods

Field grown *Spinacia oleracea* L. was purchased locally, 1 to 3 days after harvesting. L-[^{35}S]methionine (600–1 400 Ci/mmol) and sodium [^{14}C]-bicarbonate (58 mCi/mmol) were purchased from Amersham, Arlington Heights, IL, and [^{3}H]formaldehyde (25–100 mCi/mmol) was purchased from New England Nuclear, Boston, MA. 2-Methoxy-2,4-diphenyl-3-(2H)-furanone (MDPF 1) was obtained from Hoffmann-La Roche, Nutley, NJ. Commercial preparations of spinach ribulose-1,5-bisphosphate carboxylase (RuP_2C) were purchased from Calbiochem, La Jolla, CA. [^{14}C]Methylated and unlabeled protein molecular weight markers were obtained from Bethesda Research Laboratories, Bethesda, MD. All other reagents were obtained from commercial sources.

Preparation of chloroplast cpDNA

Spinach leaves were homogenized and cpDNA was prepared by the method described by Kolodner & Tewari (16) except that Proteinase K (Merck, Darmstadt, Germany) was substituted for Pronase. Lysed chloroplasts were centrifuged to equilibrium in CsCl/ethidium bromide gradients (16). The cpDNA bands were removed and the ethidium bromide extracted with isoamyl alcohol. The colorless solution was dialyzed against 10 mM Tris-acetate, pH 8.0, and the DNA precipitated with alcohol and redissolved in 10 mM Tris-acetate, pH 8.0. In an alternative procedure, 1% polyvinyl pyrrolidone-40 000 MW (PVP-40) was added to the homogenizing buffer (16) and the incubations with DNAse and Proteinase K were omitted. Chloroplast isolated in the presence of PVP-40 were washed with a solution containing 300 mM sucrose, 50 mM Tris-HCl, pH 8.0, 20 mM EDTA. The chloroplasts were lysed (16) and the total lysate was centrifuged on a CsCl equilibrium gradient and the DNA isolated as described above. Typical yields from either procedure ranged from 100 to 500 μg of DNA per kilogram of leaves. The restriction patterns of the cpDNA (13) isolated by either procedure were similar but the cpDNA from the PVP-40 procedure did contain some nuclear DNA. However, both DNA preparations were active as templates for the synthesis of the LS in the partially defined *in vitro* protein synthesis system.

Plasmid DNA

The construction and isolation of plasmid pSoe3101 has been described elsewhere (10). This plasmid contains an 11.2 kilobase-pair (kbp) BamHI fragment from spinach cpDNA that was inserted into the BamHI site of pBR322. Plasmid pJEa4, which has an insert containing the entire LS gene, was derived from pSoe3101. Following a double digestion of pSoe3101 with BamHI and AvaI restriction enzymes, a 2.2 kbp fragment was isolated by agarose gel electrophoresis and electroelution (38). Similarly the plasmid pBR322 was double digested with BamHI and AvaI, and a 3.3 kbp fragment was isolated which contains the β-lactamase gene and the origin of replication, but lacks 1.05 kbp of the tetracycline gene. The 2.2 kbp BamHI, AvaI digested fragment from pSoe3101 was ligated

in vitro with the 3.3 kbp BamHI, AvaI digested pBR322 fragment. After transformation of *E. coli* strain RR1 and screening (10), plasmid pJEa4 was isolated.

Preparation of chloroplast cpRNA

Chloroplasts were isolated from spinach leaves in the presence of PVP-40 and washed as described above. The isolated chloroplasts were lysed on ice for 15 min by adding 2 volumes of a solution containing 25 mM Tris-HCl, pH 8.0, 25 mM EDTA, 2% $NaDodSO_4$, 2 mg/ml heparin and 200 μg/ml Proteinase K. The solution was stirred vigorously for 2 min and then 1 volume each of phenol (redistilled, water saturated, pH 8.0) and chloroform/isoamyl alcohol (25:1) were added. The mixture was stirred for an additional 2 min and the aqueos phase was separated by centrifugation. The aqueous layer was extracted with phenol/chloroform/isoamyl alcohol two additional times. The nucleic acids were precipitated from the aqueous phase by the addition of 0.1 volume of 1 M sodium acetate, pH 5.0 and 2 volumes of ice cold absolute alcohol and stored at −20 °C overnight. The precipitate was collected by centrifugation and dissolved in water. Three volumes of 4 M sodium acetate, pH 6.0, were added to precipitate the RNA and separate it from DNA (28). The precipitated RNA was washed three times with 3 M sodium acetate, pH 6.0, and then dissolved and dialyzed against water. The cpRNA was stored in a liquid nitrogen freezer.

Preparation of tritiated RuP_2C and fluorescent RuP_2C

[^{3}H]RuP_2C was prepared by reductive methylation using [^{3}H]-formaldehyde, and cyanoborohydride as described (14). The specific activity routinely obtained was approximately 1 500 to 3 000 cpm/μg. The distribution of radioactivity between the large subunit and small subunit was approximately 3.4:1, consistent with the distribution of lysine residues (31, 42). Fluorescent labeled (MDPF 1) RuP_2C was prepared by a modification of the procedure (1, 39) used previously to label several *E. coli* proteins (41). Antisera against RuP_2C was raised in white male New Zealand rabbits. Each new lot of antiserum was titrated to obtain maximum precipitation of [^{3}H]RuP_2C.

In vitro *protein synthesis*

The preparation, source and purity of the protein factors used and the characteristics of the partially defined, coupled *in vitro* system have been described previously (20, 41). The complete system (35 μl) contained 30 mM Tris-OAc (pH 8.0); 10 mM Na-dimethyl-glutarate (pH 6.0); 35 mM NH_4OAc; 65 mM KOAc; 9.3 mM MgOAc; 2 mM dithiothreitol; 0.7 mM UTP, CTP, and GTP; 2.7 mM ATP; 30 mM phosphoenolpyruvate; 0.8 μg of pyruvate kinase; 2.5 mg of polyethylene glycol 6 000; 0.8 mM spermidine; 0.125 mM each of 19 amino acids (minus methionine); 4 nmol of L-[^{35}S]methionine (15 000) or 40 pmol of f[^{35}S]Met-tRNA (80 000 cpm/pmol) plus 4 nmol of methionine; 1.5 nmol of $N^{5,10}$-methenyl-H_4-folate; 50 μg of *E. coli* B tRNA and 0.8 to 1.2 A_{260} units of NH_4Cl-washed *E. coli* ribosomes. The following protein components were then added: 50 to 150 units each of 20 aminoacyl-tRNA synthetases, IF-1 (0.5 μg), IF-2 (0.4 μg), IF-3 (0.5 μg), EF-Tu (9 μg), EF-G /1.2 μg), Ehrlich's ascites cell extract (26 μg), *E. coli* RNA polymerase (0.9 μg), 4.5 to 20 μg of a partially purified fraction from a 1 M DEAE salt eluate (41)* and 2.6 μg of cpDNA, or where indicated, 0.5 to 1.0 μg of plasmid pSoe3101 or pJEa4. When cpRNA (20 μg) was used as template, RNA polymerase was omitted. The reaction mixture was incubated at 37 °C for 60 min unless stated otherwise. At the end of the 60-min incubation, a 5 μl aliquot of the reaction was removed for slab gel electrophoretic analysis, (7.5–15% polyacrylamide gradient) in the presence of 0.1% $NaDodSO_4$, followed by fluorographic detection (2, 22). The remainder of the incubation was brought to 2% $NaDodSO_4$ and heated at 90 °C for

* The 1 M DEAE salt eluate of a DEAE column (41) contains a factor that stimulates protein synthesis with all of the DNA templates tested. The active component has been partially purified as follows. The 1 M DEAE salt eluate was concentrated by ammonium sulfate precipitation (0–80%). About 1 g of protein was placed on a 40 ml DE-53 column and the protein eluted using a gradient between 0.1 and 0.6 M KCl in a buffer that contained 20 mM Tris-Cl, pH 7.4, 1 mM Mg acetate and 1 mM DTT. The active component was eluted at about 0.2 M KCl and the protein was concentrated by ultrafiltration. About 180 mg of this fraction was chromatographed on a 200 ml Ultrogel (ACA-44) column using a buffer containing 20 mM Tris-Cl, pH 7.4, 1 mM Mg acetate, and 1 mM DTT. The active fractions (eluting in the range of 80 000 to 100 000 daltons) were pooled, concentrated by ultrafiltration, and used in the *in vitro* system.

20 min to solubilize the synthesized LS. A 5 μl aliquot of the incubation was heated in 10% Cl_3CCOOH, the precipitate collected on a nitrocellulose filter and the amount of [^{35}S]methionine retained on filter determined (total protein).

Quantitation of the synthesis of LS by immunoprecipitation and gel analysis

Aliquots (5 to 20 μl) of the $NaDodSO_4$ treated reaction mixture were mixed with 2 to 5 μg of [^{3}H]RuP_2C, 40 to 100 μl of RuP_2C antiserum and 50 μl of immunoprecipitation buffer (200 mM Tris-HCl, pH 7.9, 4% Triton X-100, 2 M NaCl) in a final volume of 200 μl. The immunoprecipitates were incubated at 37 °C for 3 h, and then chilled at 4 °C overnight. The immunoprecipitates were collected by centrifuging in a microcentrifuge for 1 min and the pellets washed three times with one-fourth strength immunoprecipitation buffer. The precipitates were suspended in 50 μl of 2% $NaDodSO_4$ containing 0.1 M 2-mercaptoethanol, and dissolved by heating for 2 min at 90 °C. The dissolved immunoprecipitates were mixed with 3–5 μg of MDPF-RuP_2C, bromophenol blue, and 40% (w/v) sucrose. The samples were electrophoresed on 10% polyacrylamide disc gels (9). The fluorescent labeled RuP_2C allowed direct visualization of the large subunit both during and after disc gel electrophoresis by using UV (352 nm) illumination (41). The gels were either sliced mechanically to obtain a complete radioactive profile or only the fluorescent band was removed manually to determine the radioactivity in the synthesized LS. The gel sections were extracted with 1.0 ml of 0.1% $NaDodSO_4$ for 60 min at 55 °C. Radioactivity was determined after addition of 7 ml of Instabray (National Diagnostics, Parsippany, NJ) in a liquid scintillation counter. The tritiated RuP_2C served as an internal standard for calculating the recovery of the synthesized LS after immunoprecipitation and gel electrophoresis. The amount of LS formed *in vitro* was calculated, after correcting for recovery and aliquots, from the specific activity of the L-[^{35}S]methionine used, and the number of methionine residues in the large subunit (42).

Sucrose gradient analysis of the in vitro *synthesized product*

The contents of ten incubations were pooled and dialyzed *vs.* 50 mM Tris-HCl (pH 8.5), 20 mM $MgCl_2$, 2% glycerol at 4 °C. Insoluble material was removed by spinning for 1 min in a microfuge and the supernatant placed on a 13 ml linear 2–15% sucrose gradient containing 50 mM Tris-HCl (pH 8.5), 20 mM $MgCl_2$, 2% glycerol. The gradients were centrifuged in a SW40 rotor at 35 000 rpm for 30 h and 0.45 ml fraction collected. One hundred μl aliquots were treated with LS antiserum as described above. The immunoprecipitates were washed, solubilized, and the radioactivity determined. Coprecipitation with [^{3}H]RuP_2C standard was used to determine recovery. Molecular weight markers used were: aldolase (160 000), bovine serum albumin (68 000), ovalbumin (45 000), and chymotrypsinogen (26 000).

Assay for enzyme activity of in vitro *synthesized LS*

Using pJEa4 DNA as template, LS was synthesized as described above (without [^{35}S]methionine) and tested for enzyme activity using an assay based on the fixation of $H^{14}CO_3$ into acid stable material dependent on RuP_2. For these experiments the 35 μl *in vitro* incubation (see above) was brought to a final volume of 0.2 ml and contained 40 mM Tris-HCl (pH 7.6), 15 mM $MgCl_2$, 10 mM $NaH^{14}CO_3$ (20 000 cpm/nmole), 4 mM NaCl, 2% glycerol and 0.08 mM EDTA. The solution was preincubated at 30 °C for 20 min to activate the enzyme before initiating the reaction by adding 4 μl of 750 mM RuP_2. After 1 h at 30 °C, 35 μl of 6 N HCl were added to terminate the reaction. The sample was transferred to a scintillation vial and evaporated to dryness under N_2 to liberate free $^{14}CO_2$. The dried samples were dissolved in 0.5 ml water and the radioactivity measured. Control incubations were run to insure that RuP_2C activity (using standard spinach holoenzyme) could be recovered when carried through the entire procedure.

In vitro *run-off assays using* E. coli *RNA polymerase*

RNA transcripts were synthesized *in vitro* using *E. coli* RNA polymerase and various DNA fragments derived from the plasmid pJEa4. Reactions contained 10 mM Tris-HCl, pH 7.9, 10 mM $MgCl_2$, 100 mM KCl, 0.1 mM each of ATP, CTP and GTP, 0.01 mM [α^{32}P]-UTP (20 μCi), 0.5 to 1 unit of *E. coli* RNA polymerase (Worthington Biochemicals)

and 0.1 to 0.5 μg of DNA in a total volume of 50 μl. After transcription (20 min at 37 ° C), samples were diluted with 50 μl of water and extracted two times with chloroform. Samples were precipitated and washed with ethanol. After drying under vacuum, the samples were solubilized in 95% formamide and heated for 30 min at 50 ° C and loaded onto 8% acrylamide sequencing gels (37). Electrophoresis and autoradiography were carried out as previously described (37).

DNA sequencing

DNA was sequenced using the chemical modification procedure of Maxam & Gilbert (27). DNA fragments were labeled at their 3′ ends using [α^{32}P]-dNTP and the Klenow fragment of *E. coli* DNA polymerase 1 (27), or alternatively, for DNA fragments which contained a 3′ overhang, [α^{32}P]-cordycepin triphosphate and terminal transferase, using a kit supplied by New England Nuclear (Boston, MA).

Results

In vitro *products directed by cpDNA and cpRNA*

Analysis of the proteins synthesized in the cpDNA and cpRNA directed *in vitro* incubations was initially done by slab gel electrophoresis. As seen in Fig. 1, with both templates, a number of [^{35}S]methionine-labeled proteins are formed in the partially defined system. A prominent band is present at approximately 55 000 daltons that comigrates with the large subunit of [^{3}H]RuP_2C. Positive identification of the LS protein, as one of the *in vitro* products, was achieved by immunoprecipitation followed by disc gel analysis (see Materials and methods). Analysis of the radioactive products by gel electrophoresis after immunoprecipitation with RuP_2C antiserum is seen in Fig. 2. A [^{35}S]methionine labeled protein, that comigrates with the LS protein ([^{3}H]-standard, Fig. 2) is the primary product on the gel after immunoprecipitation with RuP_2C antiserum in both the cpDNA (Fig. 2A) and cpRNA (Fig. 2B) directed system. As described in Materials and methods, this procedure provided the basis for a quantitative assay for the formation of the LS in the *in vitro* incubations.

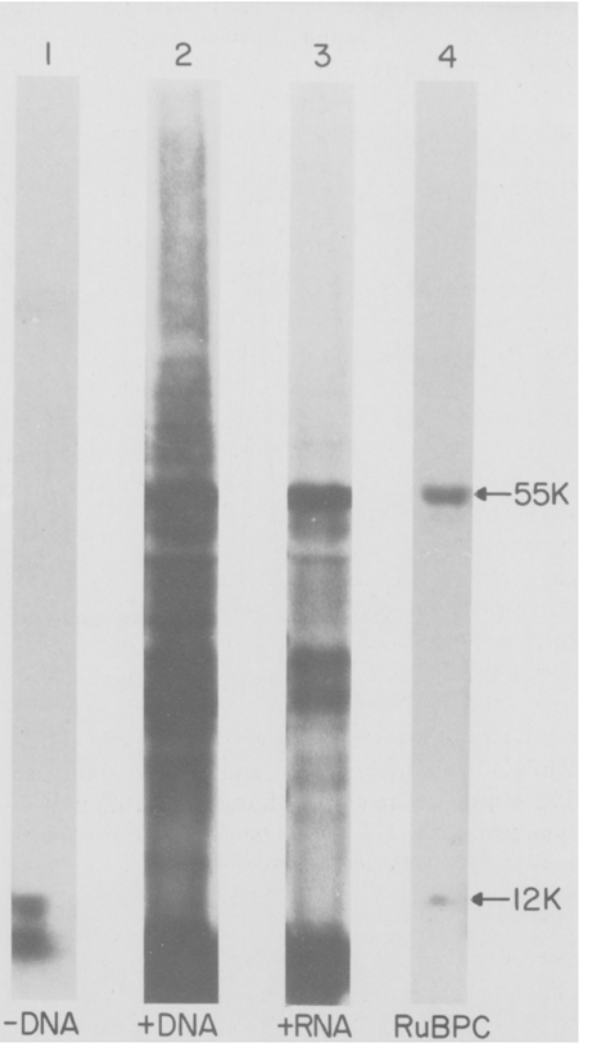

Fig. 1. Gel separation of [^{35}S]methionine labeled protein products synthesized in the partially defined *in vitro* protein synthesis system. The products were electrophoresed through a NaDodSO$_4$/polyacrylamide gradient gel (7.5 to 15%) and fluorographed as described in Materials and methods. Track 1: no cpDNA or cpRNA; track 2: 2.5 μg cpDNA; track 3: 20 μg cpRNA; track 4: [^{3}H]RuP_2C standard.

Characteristics of the in vitro *system*

The effect of template concentration on the *in vitro* formation of the LS protein in the partially defined system was determined. Although 2.5 μg of cpDNA saturate the reaction, about 40 μg of total cpRNA are required for saturation (data not shown). With both templates, LS synthesis in-

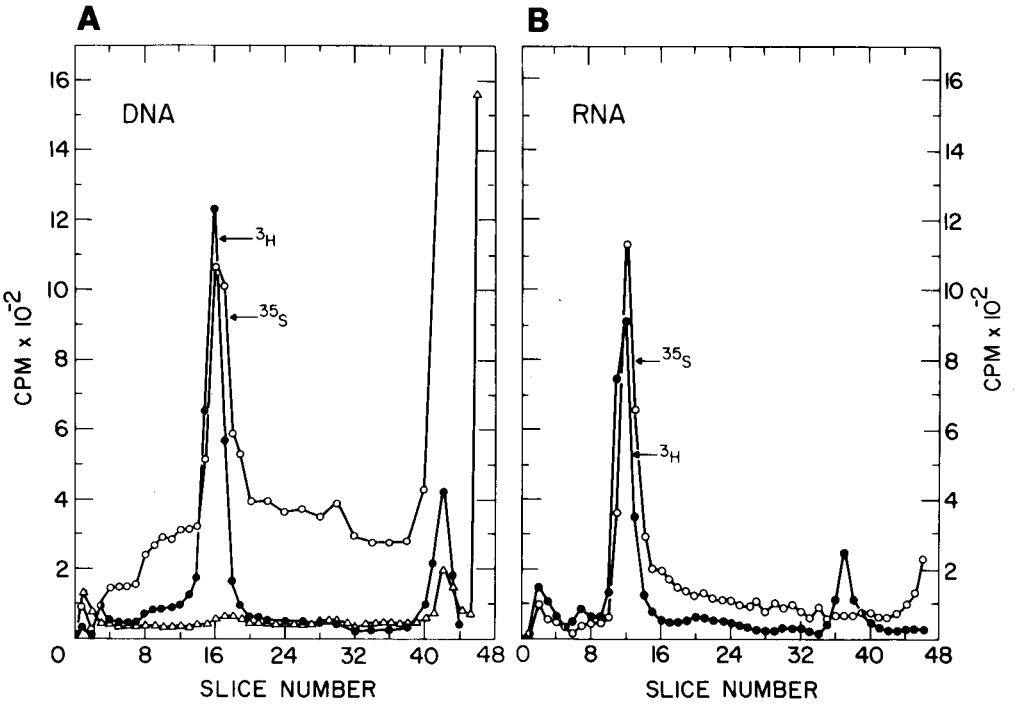

Fig. 2. Identification of LS-RuP$_2$C by gel analysis after immunoprecipitation.
A: A 20-μl aliquot of the reaction mixture with cpDNA as template was used for immunoprecipitation (see Materials and methods). The solubilized immunoprecipitate was analyzed on a 10% polyacrylamide gel (2), and the gels were sliced into 2 mm sections and assayed for radioactivity. (●—●): [^{3}H]RuP$_2$C; (○—○): [^{35}S]methionine products from incubations with cpDNA; (△—△): [^{35}S]methionine products from incubation without added DNA.
B: A 5-μl aliquot of the reaction mixture with cpRNA as template was used for immunoprecipitation (see Materials and methods). The solubilized immunoprecipitate was analyzed as described above. (●—●): [^{3}H]RuP$_2$C; (○—○): [^{35}S]methionine products from incubations with cpRNA.

creases with time for up to 60–90 min (Fig. 3) and a sharp Mg^{+2} optimum was obtained between 8.5 and 9.5 mM. Generally, between 0.2 and 0.6 pmol of LS was formed in typical incubations with cpRNA. However, the activity with different cpDNA preparations was quite variable with values ranging between 0.03 and 0.2 pmol of LS synthesized in typical incubations. It was calculated that even in the cpRNA directed system, the synthesis of the LS represented only a few percent of the total [^{35}S]methionine incorporated into hot Cl_3CCOOH insoluble products.

Dependencies for LS synthesis

One of the prime advantages of the partially defined *in vitro* system is that it permits studies on the role of individual factors in protein synthesis. Pre-

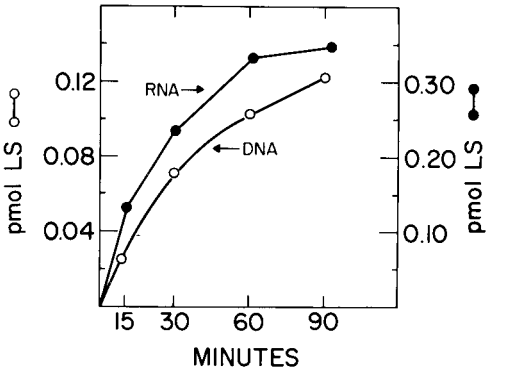

Fig. 3. Time course for LS synthesis. Details are described in Materials and methods. 2.5 μg of cpDNA and 20 μg of cpRNA were used.

Table 1. Dependencies for the synthesis of the large subunit of RUBPC.

Template	Omissions	%
cpRNA	None	100
	EF-Tu	0
	IF-3	10
	Ribosomes	0
cpDNA	None	100
	Polymerase	0

Details of the *in vitro* system including the amounts of each of the components used are listed in Materials and methods. 100% values for the cpRNA- and cpDNA-directed systems were 0.3 pmol and 0.04 pmol, respectively.

vious studies on the synthesis of β-galactosidase (β-gal) in a similar system showed excellent dependencies on many of the purified factors (17–21). Although dependencies on all of the individual *E. coli* components have not been examined in the present study, essentially complete dependencies on RNA polymerase, ribosomes, EF-Tu, and IF-3 were obtained (Table 1).

Studies with plasmids containing the LS gene

The relatively low *in vitro* activity obtained using spinach cpDNA as template plus the difficulty in obtaining reproducible DNA preparations from spinach chloroplasts presented serious problems for continuing studies on the regulation of expression of the *rbc*L gene. Recently, Erion *et al.* (10) have described the construction of recombinant plasmids containing the spinach *rbc*L gene. Plasmids such as pSoe3101 and pJEa4 (see Materials and methods) are excellent templates in the *in vitro* system. As seen in Fig. 4, with saturating levels of pJEa4, it was possible to routinely synthesize 0.4 pmol of LS in 60 min, and values greater than 1.0 pmol were very often reached. Similar results were obtained with pSoe3101 although with this template somewhat lower levels of LS were synthesized *in vitro*. Because of the ease of obtaining the plasmid DNA and the higher *in vitro* activity, the studies presented below were done with plasmid pJEa4 as template.

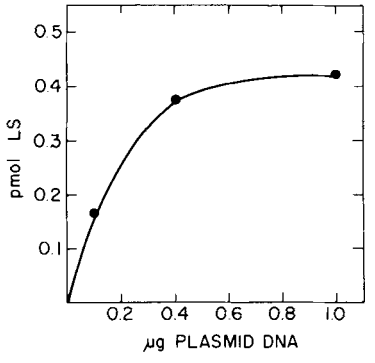

Fig. 4. Synthesis of LS directed by different concentrations of plasmid pJEa4. Details of the *in vitro* system are described in Materials and methods.

Characteristics of the in vitro *formed product; N-terminal identification*

There are limited data on the N-terminal amino acid in mature LS since this protein contains an N-blocked amino terminus as isolated from many species (29). It has been reported that barley LS has the dipeptide alanyl-glycine at the N-terminal positions (29) which would indicate that the first 14 amino acids of the nascent LS protein are removed. It has also been reported that *in vitro* synthesized spinach LS is 1 000 to 2 000 daltons larger than chloroplast synthesized LS (23), suggesting that processing of the spinach LS may also take place. In the highly defined *in vitro* system used in these studies, it is probable that little or no processing of the synthesized product takes place. To ascertain this, standard incubations were carried out using either [^{35}S]Met-tRNA (reactions included unlabeled methionine) or [^{35}S]methionine and the radioactivity in the product determined. The results are seen in Table 2. The formylated initiator amino acid was not only incorporated into LS, but almost equivalent amounts of product were formed using either fMet-tRNA or free methionine showing that the initiator fMet has not been removed from the *in vitro* synthesized product. The possibility that the f[^{35}S]Met-tRNA (40 pmol) was deformylated to

Table 2. In vitro synthesis of large subunit using [^{35}S]fMet-tRNA or [^{35}S]Met.

[^{35}S] Component	LS Synthesized
	pmol
Methionine	0.50
fMet-tRNA	0.42

Details of the incubation are described in the text. The calculations to determine the pmoles of products were based on ten moles of methionine per mole of protein when [^{35}S]-methionine was used, and one mole of formyl-methionine per mole of protein when [^{35}S]fMet-tRNA was used.

[^{35}S]methionine which was then incorporated into product was excluded by the presence of a large excess (4 nmol) of unlabeled methionine in the incubations.

Is the in vitro *product a monomer of aggregate?*

RuP_2C is a high molecular weight aggregate containing eight LS and eight small subunit chains. Whether the *in vitro* synthesized LS is also present as an aggregate was examined by determining the size of the product using sucrose gradient analysis. The results of a typical experiment are shown in Fig. 5. When the *in vitro* synthesized product was centrifuged through a 2–15% sucrose gradient (see Materials and methods), the radioactive material that reacted with antiserum to the LS emerged as a low molecular weight species (approximately 50 000) with no evidence of higher molecular weight forms. Under these conditions RuP_2C holoenzyme did not dissociate, and sedimented as a high molecular weight species of about 500 000 daltons (data not shown).

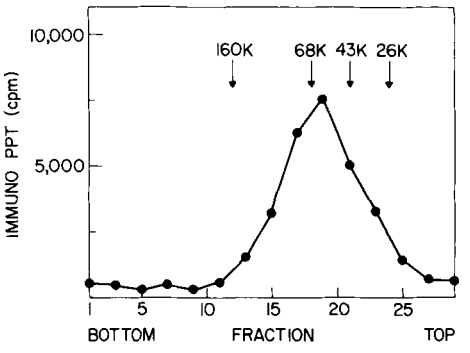

Fig. 5. Sucrose density gradient centrifugation of *in vitro* [^{35}S] labeled LS. The LS was synthesized *in vitro* as described in Materials and methods and centrifuged through a linear 2–15% sucrose gradient for 30 h at 35 000 rpm in a SW40 rotor at 4 °C. Fractions (0.45 ml) were collected and aliquots precipitated with RuP_2C antiserum. Molecular weight markers (aldolase, bovine serum albumin, ovalbumin and chymotrypsinogen) were centrifuged in parallel and the absorbance at 260 nm measured (indicated by arrows).

Enzymatic activity of the in vitro *synthesized product*

All attempts to demonstrate that the *in vitro* synthesized product could fix CO_2 dependent on RuP_2 have been unsuccessful (Table 3). Control experiments in which RuP_2C was added to the *in vitro* incubations demonstrated that the activity of the enzyme could be recovered through the procedure. As little as 0.001 pmol of the enzyme (based

Table 3. Assay of *in vitro* synthesized LS for RuP_2C activity.

Components added	LS formed pmol	Co_2 fixed	
		cpm	nmol
Minus DNA	0	161	–
Complete	2.0	149	–
Complete + RuP_2C (0.5 μg, 1 pmol)	2.0	280 000	14.0

The incubation conditions are described in the text. Where indicated, spinach RuP_2C (1 pmol) was added to the *in vitro* system to show that the carboxylase holoenzyme could be recovered through the procedure (see Materials and methods). The CO_2 fixed in this case was completely dependent on RuP_2.

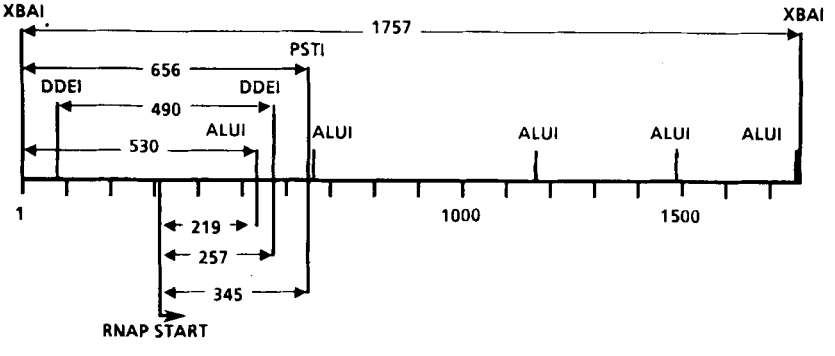

Fig. 6. Map of DNA fragments used in transcription run-off assays. Partial restriction map of a 1757 base-pair Xba 1 fragment containing most of the *rbc*L gene. The position of the promoter for transcription is from the data of Zurawski *et al.* (42). Lower measurements from the promoter site to downstream restriction sites are the predicted size of run-off transcripts in the *in vitro* assay.

on pure spinach RuP_2C activity) could have been detected by the assay used.

Interaction of E. coli *RNA polymerase with the* rbc*L gene*

Our data as well as previous reports (4, 8, 10, 25) indicate that *E. coli* RNA polymerase efficiently transcribes the chloroplast *rbc*L gene. However, it is important to establish whether *E. coli* RNA polymerase initiates transcription at the same site as the chloroplast RNA polymerase. The *in vivo* RNA start site on the *rbc*L gene has been mapped at a position which is 178 base-pairs (bp) upstream from the first AUG codon in the coding sequence (42). Using this data we compared the size of the RNA transcripts produced from various restriction fragments containing the presumed *in vivo* transcription promoter (see Figs. 6 and 7). A 1757 bp Xba 1 fragment isolated from plasmid pJEa4 was restricted with three different enzymes. These enzymes (Alu 1, Pst 1, and Dde 1) cut the original fragments in such a way that one can readily discern whether transcripts initiate from specific or non-specific start sites, such as from the ends of DNA fragments (see Fig. 6). For example, cleavage of the Xba 1 fragment with Alu 1, yields five fragments with sizes of 530 bp, 498 bp, 332 bp, 265 bp and 129 bp. Using the Alu 1 digested Xba 1 fragment, *E. coli* RNA polymerase produced a major run-off transcript which was approximately 220 bases in length (Fig. 7, Lane 3). Since there are no Alu 1 fragments of this size, transcription must be initiating from a specific site. The minor transcript of about 235 bases as well as the higher molecular weight products probably arise from non-specific start sites. Similarly, transcription of the Xba 1 fragment after cleavage with Pst 1, which yields 656 bp and 1098 bp fragments, resulted in a run-off transcript which measured 345 bases in length (Fig. 7, Lane 1). The differences in the size of the run-off transcripts produced by the Alu 1 and the Pst 1 restricted Xba 1 fragment corresponds to the distance between one of the Alu 1 sites and the Pst 1 site on the Xba 1 fragment. Taken together, these results indicated that the run-off transcripts initiate at a site which is 310 bp from the left end of the Xba 1 fragment (see Fig. 6). A comparable analysis of the run-off transcript produced by the Dde 1 digested, Xba 1 fragment (Fig. 7, Lane 2) again indicated that transcription was initiating from a site 310 bp from the left end of the Xba 1 fragment. This position is identical to that found for the start site for the *in vivo* LS RNA transcripts (42), corresponding to a position which is 178 bp upstream from the first AUG triplet in the coding region of the gene.

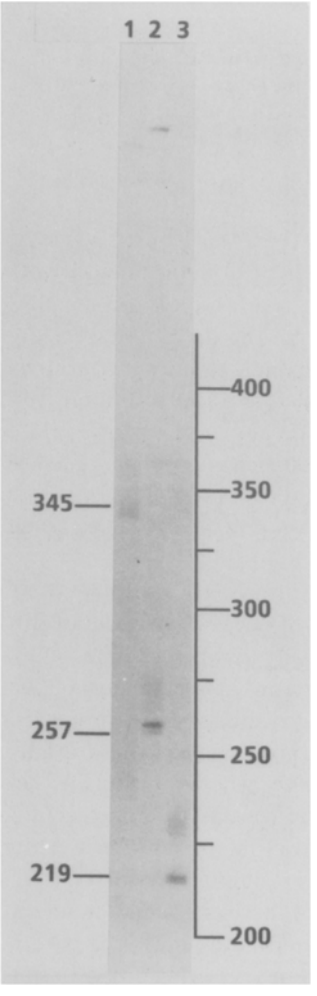

Fig. 7. Autoradiogram of *in vitro* products from transcription run-off assays. The 1757 base-pair Xba 1 restriction fragment, isolated from the plasmid pJEa4, was digested with: 1) Pst 1, 2) Dde 1, and 3)Alu 1. The digested DNA fragments (1 μg) were added to an *in vitro* transcription system using *E. coli* RNA polymerase (see Materials and methods). Radiolabeled RNA transcripts were analyzed on a 7M urea, 6% acrylamide sequencing gel (see Materials and methods). The gel was calibrated (left axis, in bases) using [^{32}P]-end labeled DNA fragments of known sizes (lanes not shown). Right axis indicates size in bases ($\pm$ 3 bases) of major *in vitro* products. Only the relevant portion of the gel is shown.

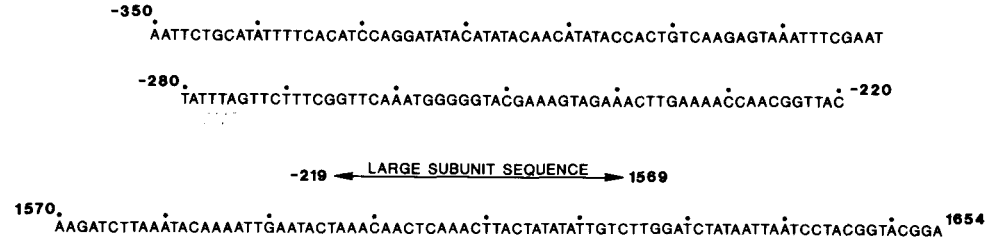

Fig. 8. DNA sequence of the 5′ and 3′ flanking regions of the *rbc*L gene. The DNA sequence from nucleotide -350 to -220 and from 1570 to 1654 are shown. The sequence from -220 to 1570 is identical to that previously described (42) except at position 35 and 657 (see text).

Sequence of the spinach rbc*L gene*

Although the complete sequence of the spinach *rbc*L gene has been determined by others (42, 43), data are not available on the nucleotide sequence upstream from the determined *in vivo* RNA start site (42). This segment might contain regions which could serve as fortuitous promoters for *E. coli* RNA polymerase or have a normal regulatory role. The complete sequence of the *rbc*L gene was determined including additional 5′ sequence data from positions -350 through -226 and 3′ sequence data from position 1575 through 1654.

Figure 8 shows only these new sequences since comparison of our sequence data for the *rbc*L gene with previously published data (42, 43) indicates that there are only two nucleotide positions where disagreement occurs. At position 657, the substitution of a guanosine nucleotide for adenosine at the third position in the triplet (CTA) coding for leucine has no effect on the amino acid coded for. At position 35 we find a guanosine residue and not an adenosine as previously reported (42). This results in a change of the amino acid sequence from glycine (GGA) to glutamate (GAA). It is to be noted that GGA also appears at this position in the maize LS sequence (8). Whether this difference in the spinach *rbc*L gene DNA sequence reflects an error in reading or represents heterogeneity in the gene, which is either naturally occurring or a mutation artifactually induced during cloning, is not known.

In addition, examination of the data indicates that there is only one region in the entire sequence which contains a prokaryotic-like promoter (at -213 to -177) and that region corresponds exactly to the position predicted from the start site determined from the *in vitro* transcription of this gene using *E. coli* RNA polymerase (see above). From these observations, it seems likely that the DNA sequence recognized by chloroplast RNA polymerase *in vivo* is very closely related to the sequences recognized by bacterial RNA polymerases (35).

Discussion

In previous studies on the *in vitro* expression of several *E. coli* genes, a partially defined system that included more than 30 highly purified factors was used (20, 41). The advantage of such a system is that it can be used to identify new factors involved in transcription and/or translation as well a for studies on the regulation of expression of specific genes. In the present study, this system has been used to examine the synthesis of the LS of spinach RuP_2C, a protein coded by cpDNA. The prokaryotic nature of the chloroplast has made it possible to use crude *E. coli* extracts (4, 10, 12, 40) or, as described here, a partially defined system containing purified *E. coli* factors to obtain expression of the *rbc*L gene. Both cpRNA and cpDNA, and plasmids carrying the spinach *rbc*L gene have been used in the present study and an assay has been developed that allows quantitation of the formed product *in vitro*. The cpRNA-directed system was more efficient than the cpDNA system, but the most active templates used in this study were plasmids such as pSoe3101 or pJEa4 that contain the gene for the LS of spinach RuP_2C (10).

One advantage of the defined *E. coli* system is the potential ability to replace a specific *E. coli* factor with a chloroplast factor. Although our attempts at such studies are not reported here, an effort was made to replace the *E. coli* RNA polymerase with a

chloroplast polymerase. Spinach cpRNA polymerase was purified about 300-fold by a modification of the procedure used for maize by Bogorad and coworkers (3, 15, 36) but the spinach chloroplast enzyme was found to have low activity in transcribing plasmid DNA or total cpDNA and could not replace the *E. coli* polymerase in the coupled system. All efforts to obtain transcription of the plasmid template, using the cpRNA polymerase, by the addition of spinach chloroplast factors, were also unsuccessful.

Previous studies using a defined *in vitro* system had shown that formylation of the initiator Met-tRNA was essential for the expression of the lactose operon (21). However, the presence of a formylated LS gene product has not been demonstrated previously. In the present studies, it was shown that the LS synthesized in the defined *in vitro* system contains fMet at the N-terminal position. Since the *in vitro* product migrated on a polyacrylamide gel under denaturing conditions with a molecular weight of about 55 000, it was evident that the system could synthesize the complete protein, and suggested that there was little degradation of the product under the incubation conditions used. Our finding, that *E. coli* RNA polymerase specifically initiates on the *rbc*L gene at the same position as chloroplast RNA polymerase *in vivo*, agrees with data obtained by Shiozaki & Sugiura (34), who examined the initiation of transcription by *E. coli* RNA polymerase of the *rbc*L gene from tobacco cpDNA using S1 nuclease mapping. The nucleotide sequence that makes up the transcriptional promoter for the tobacco and spinach *rbc*L gene are virtually identical (see (34)). In addition, Brait *et al.* (5) have shown that the products obtained from the *in vitro* transcription of cloned spinach chloroplast DNA by either chloroplast or *E. coli* RNA polymerase were identical. These results reinforce the idea that the use of an *E. coli* system to study the regulation of chloroplast genes, *in vitro*, has validity.

The *in vitro* synthesized LS had no carboxylase activity. Sucrose gradient analysis under non-denaturing conditions also indicated that the *in vitro* product was a monomer, showing that there was no significant aggregation of the LS in these experiments. Thus far, preliminary attempts to form an octomer, as in the native enzyme, or obtain enzymatically active LS by the addition of the small subunit or spinach chloroplast extracts have been unsuccessful (data not shown).

In a recent study, we have examined the expression of the *rbc*L gene on pJEa4 using a simplified *in vitro* system in which the formation of the first peptide, fMet-Ser, was measured (6). In the present study we have developed a quantitative procedure for the formation of the LS using a defined *in vitro* system. Having available these *in vitro* systems to study the expression of the *rbc*L gene should make it possible to investigate the regulation of the expression of this gene. In addition, the *in vitro* system described here may facilitate the study of the properties of the LS in the absence of the small subunit.

References

1. Barger BO, White FC, Pace JL, Kemper DL, Ragland WL: Estimation of molecular weight by polyacrylamide gel electrophoresis using heat stable fluorophors. Anal Biochem 70:327–335, 1976.
2. Bonner WM, Laskey RA: A film detection method for tritum-labelled proteins and nucleic acids in polyacrylamide gels. Eur J Biochem 46:83–88, 1974.
3. Bottomley W, Smith HJ, Bogorad L: RNA polymerase of maize: partial purification and properties of the chloroplast enzyme. Proc Natl Acad Sci USA 68:2412–2416, 1971.
4. Bottomley W, Whitfeld PR: Cell-free transcription and translation of total spinach chloroplast DNA. Eur J Biochem 93:31–39, 1979.
5. Briat JF, Drun M, Loiseaux S, Mache R: Structure and transcription of the spinach chloroplast rDNA leader region. Nucleic Acids Res 10:6865–6878, 1982.
6. Cenatiempo Y, Robakis N, Meza-Basso L, Brot N, Weissbach H, Reid BR: Use of different $tRNA^{Ser}$ isoacceptor species in vitro to discriminate between the expression of plasmid genes. Proc Natl Acad Sci USA 79:1466–1468, 1982.
7. Chua NH, Blomberg F: Immunochemical studies of thylakoid membrane polypeptides from spinach and Chlamydomonas reinhardtii. J Biol Chem 254:215–223, 1978.
8. Coen DM, Bedrock HR, Bogorad L, Rich A: Maize chloroplast DNA fragment encoding the large subunit of ribulose bisphosphate carboxylase. Proc Natl Acad Sci USA 74: 5487–5491, 1979.
9. Davis BJ: Disc gel electrophoresis II. Method and application to human serum proteins. Ann NY Acad Sci 121: 404–427, 1964.
10. Erion JL, Tarnowski J, Weissbach H, Brot N: Cloning, mapping and in vitro transcription-translation of the gene for the large subunit of ribulose-1,5-bisphosphate carboxylase from spinach chloroplasts. Proc Natl Acad Sci USA 78:3459–3463, 1981.
11. Gatenby AA, Castleton JA, Sage MW: Expression in E. coli of maize and wheat chloroplast genes for large subunit of ribulose bisphosphate carboxylase. Nature 291:117–121, 1981.

12. Hartley MR, Wheeler A, Ellis RJ: Protein synthesis in chloroplasts. V. Translation of messenger RNA for the large subunit of fraction I protein in a heterologous cell-free system. J Mol Biol 91:67–77, 1975.
13. Hermann RG, Feierabend J: The presence of DNA in ribosome-deficient plastids of heat-bleached rye leaves. Eur J Biochem 104:603–609, 1980.
14. Jentoft N, Dearborn DG: Labeling of proteins by reductive methylation using sodium cyanoborohydride. J Biol Chem 254:4359–4365, 1979.
15. Jolly SO, Bogorad L: Preferential transcription of cloned maize chloroplast DNA sequences by maize chloroplast RNA polymerase. Proc Natl Acad Sci USA 77:822–826, 1980.
16. Kolodner R, Tewari KK: The molecular size and conformation of the chloroplast DNA from higher plants. Biochim Biophys Acta 402:372–390, 1975.
17. Kung H-F, Spears C, Schultz T, Weissbach H: Studies on the in vitro synthesis of β-galactosidase: necessary components in the ribosomal wash. Arch Biochem Biophys 162:578–584, 1974.
18. Kung H-F, Spears C, Weissbach H: Purification and properties of a soluble factor required for the deoxyribonucleic acid-directed in vitro synthesis of β-galactosidase. J Biol Chem 250:1556–1562, 1975.
19. Kung H-F, Spears C, Weissbach H: DNA-directed in vitro synthesis of β-galactosidase: dependencies on elongation factor Tu and tRNA. Arch Biochem Biophys 174:100–104, 1976.
20. Kung H-F, Redfield B, Treadwell BV, Eskin B, Spears C, Weissbach H: DNA directed in vitro synthesis of β-galactosidase: studies with purified factors. J Biol Chem 252: 6889–6894, 1977.
21. Kung H-F, Eskin B, Redfield B, Weissbach H: DNA-directed in vitro synthesis of β-galactosidase: requirement for formylation of methionyl-tRNA$_f$. Arch Biochem Biophys 195:396–400, 1979.
22. Laemmli UK: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature London 227:680–685, 1970.
23. Langridge P: Synthesis of the large subunit of spinach ribulose bisphosphate carboxylase may involve a precursor polypeptide. FEBS Lett 123:85–89, 1981.
24. Link G, Coen DM, Bogorad L: Differential expression of the gene for the large subunit of ribulose bisphosphate carboxylase in maize leaf cell types. Cell 15:725–731, 1978.
25. Malnoe P, Rochaix RD, Chua NH, Spahr PF: Characterization of the gene and messenger RNA of the large subunit of ribulose 1,5-diphosphate carboxylase in Chlamydomonas reinhardtii. J Mol Biol 133:417–434, 1979.
26. Maniatis T, Ptashne M, Backman K, Kleid D, Flashman S, Jeffrey A, Maurer R: Recognition sequences of repressor and polymerasein the operators of bacteriophage lambda. Cell 5:109–114, 1975.
27. Maxam AM, Gilbert W: Sequencing and end-labeled DNA with base-specific chemical cleavages. Methods Enzymol 65:499–580, 1980.
28. Palmiter RD: Ovalbumin messenger ribonucleic acid translation. J Biol Chem 248:2095–2106, 1973.
29. Poulsen C, Martin B, Svendsen I: Partial amino-acid sequence of the large subunit of ribulose-bisphosphate carboxylase EC-4.1.1.39 from barley. Carlsberg Res Commun 44:191–199, 1979.
30. Pribnow D: Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter. Proc Natl Acad Sci USA 74:784–788, 1975.
31. Rutner AC, Lane MD: Nonidentical subunits of ribulose disphosphate carboxylase. Biochem Biophys Res Commun 28:531–537, 1967.
32. Saluja AK, McFadden BA: Modification of active site histine in ribulose bisphosphate carboxylase/oxygenase. Biochemistry 21:89–95, 1982.
33. Shine J, Dalgarno L: The 3′-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc Natl Acad Sci USA 71:1342–1346, 1974.
34. Shinozaki K, Sugiura M: Sequence of the intercistronic region between the ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit and the coupling factor β subunit gene. Nucleic Acids Res 10:4923–4934, 1982.
35. Siebenlist U, Simpsom RB, Gilbert W: E. coli RNA polymerase interacts homologously with two different promoters. Cell 20:269–281, 1980.
36. Smith HJ, Bogorad L: The polypeptide subunit structure of the DNA-dependent RNA polymerase of Zea mays chloroplasts. Proc Natl Acad Sci USA 71:4839–4842, 1974.
37. Smith AJH: DNA sequence analysis by primed synthesis. Methods Enzymol 65:560–580, 1980.
38. Southern E: Gel electrophoresis of restriction fragments. Methods Enzymol 68:152–176, 1980.
39. Weigele M, De Bernardo S, Leimgruber W, Cleeland R, Grunberg E: Fluoroscent labeling of proteins. A new methodology. Biochem Biophys Res Commun 54:899–906, 1973.
40. Whitfeld PR, Bottomley W: Mapping of the gene for the large subunit of ribulose bisphosphate carboxylase on spinach chloroplast DNA. Biochemistry International 1:172–178, 1980.
41. Zarucki-Schulz T, Jerez C, Goldberg G, Kung H-F, Huang KH, Brot N, Weissbach H: DNA-directed in vitro synthesis of proteins involved in bacterial transcription and translation. Proc Natl Acad Sci USA 76:6115–6119, 1979.
42. Zurawski G, Perrot B, Bottomley W, Whitfeld PR: The structure of the gene for the large subunit of ribulose 1,5-bisphosphate carboxylase from spinach chloroplast DNA. Nucleic Acids Res 9:3251–3270, 1981.
43. Zurawski G, Bottomley W, Whitfeld PR: Structure of the genes for the beta and epsilon subunits of spinach chloroplast ATPase indicates a dicistronic mRNA. Proc Natl Acad Sci USA 79:6260–6264, 1982.

Received 6 June 1983; in revised form 6 September 1983; accepted 8 September 1983.