PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Expression of Cyanobacterial and Higher-Plant Ribulose 1,5-Bisphosphate Carboxylase Genes in Escherichia coli

D. Bradley, Saskia M. van der Vies and A. A. Gatenby

Phil. Trans. R. Soc. Lond. B 1986 **313**, 447-458
doi: 10.1098/rstb.1986.0051

Phil. Trans. R. Soc. Lond. B **313**, 447–458 (1986)
Printed in Great Britain

Expression of cyanobacterial and higher-plant ribulose 1,5-bisphosphate carboxylase genes in *Escherichia coli*

By D. Bradley[1], Saskia M. van der Vies[2]* and A. A. Gatenby[1]†

[1] *Department of Molecular Genetics, Plant Breeding Institute, Cambridge CB2 2LQ, U.K.*
[2] *Department of Biochemistry, Agricultural University, Wageningen, The Netherlands*

Expression strategies for the synthesis of higher-plant and cyanobacterial RuBP carboxylase genes in *Escherichia coli* have been developed to facilitate the study of the assembly pathway and properties of the enzyme's large (L) and small (S) subunit proteins. The genes for the L and S subunits of the RuBP carboxylase of wheat and of a cyanobacterium, *Synechococcus* 6301 have been cloned into bacteriophage and plasmid vectors such that they are transcribed and translated in *E. coli*.

To date no RuBP carboxylase activity has been detected in extracts prepared from *E. coli* cells synthesizing the wheat L and S subunits, although both gene products were present and soluble. Sucrose gradient analysis of cell extracts from *E. coli* synthesizing both L and S demonstrated that the soluble wheat L polypeptide was present as a large protein aggregate that contained no S subunits.

With the cloned cyanobacterial genes, RuBP carboxylase activity could be recovered in *E. coli* cell extracts when the L and S gene products were synthesized from genes present on the same, or separate, replicons. Solubility and sedimentation studies of the cyanobacterial L subunits synthesized in the absence of S showed that the L subunit was soluble and present in *E. coli* as an L_8 structure. The *E. coli* extracts containing only the L subunit exhibited no detectable RuBP carboxylase activity. Infection of the *E. coli* cells containing L subunits with an M13 phage expressing the cyanobacterial S gene led to the assembly of functional RuBP carboxylase in these cells. This demonstrates the essential role of the S subunit in allowing the formation of an active enzyme.

Introduction

The expression of foreign genes in *E. coli* by means of recombinant DNA technology has proven to be a valuable tool in the study of eukaryotic gene products. Proteins have been produced in *E. coli* because of their economic value (Goeddel *et al.* 1979*a*, *b*) or to accumulate large quantities for structural and biochemical investigations. Protein engineering, by site-directed mutagenesis (Zoller & Smith 1983), is also facilitated through the expression of target genes in *E. coli*. With a functional assay for the protein, mutational variants can be screened rapidly for alterations in protein structure and function (Winter *et al.* 1982). These methods, coupled with the reintroduction of the altered gene or genes into the organism of interest by DNA transformation, provides the basic framework for genetically manipulating organisms in a very specific fashion.

Plant biologists have long discussed the possibility of genetically altering the photosynthetic capabilities of agriculturally important crops with the view towards increasing plant productivity (for review, see Ellis & Gatenby 1984). Central to this discussion has been the enzyme

* Present address: Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, U.K.
† Present address: Biophysics Department, University of Wisconsin, Madison, Wisconsin 53706, U.S.A.

ribulose-1,5-bisphosphate carboxylase (RuBPCase), which occupies a key position in carbon metabolism of all photosynthetic organisms. RuBPCase catalyses the first reaction in the interlocking, but opposing, pathways of photosynthesis (carbon fixation) and photorespiration (carbon consumption) (for review, see Miziorko & Lorimer 1983; Ellis & Gatenby 1984).

There is some interest in RuBPCase because of the possibility of altering the structure of the enzyme to improve the net photosynthetic yield, perhaps by increasing the affinity of the enzyme for carbon dioxide or by increasing the ratio of carboxylation to oxygenation at the active site (Ellis & Gatenby 1984). This modification could be attempted by site-specific mutagenesis, or by the construction of hybrid molecules with subunits from different species, to produce an enzyme with altered kinetic properties. Such studies are difficult, however, because of the structural complexity of the enzyme in most photosynthetic organisms, and because of the physical properties of the RuBPCase subunits from some plant species.

In plants and most photosynthetic bacteria, RuBPCase is composed of sixteen subunits, of which eight are large subunits (L) of molecular mass 50–60 kDa and eight are small subunits (S) of molecular mass 12–18 kDa. The active site of the enzyme is on the L subunit polypeptide (Lorimer & Miziorko 1980; Spreitzer & Mets 1980; Miziorko *et al.* 1982). In plant cells, the L gene is located in the chloroplast and the L polypeptide is synthesized in the organelle (for review, see Ellis 1981). The S polypeptides are synthesized as precursor polypeptides in the cytosol by means of nuclear-encoded mRNA molecules, and are then transported into the chloroplast, matured by proteolysis and assembled with the L polypeptides to form the active L_8S_8 holoenzyme (Broglie *et al.* 1983; Smith & Ellis 1979; Dobberstein *et al.* 1977; Gallagher & Ellis 1982). In addition, a L subunit binding protein has also been implicated in the assembly process (Barraclough & Ellis 1980; Bloom *et al.* 1983). The synthesis and assembly of RuBPCase in higher plants is therefore a very complicated process, and attempts to modify either subunit may require an understanding of the mechanisms of assembly. The roles of the L subunit binding protein and the S protein in holoenzyme assembly have not been determined. A serious impediment to such studies has been the inability to reconstitute higher-plant RuBPCase from its isolated subunits *in vitro* (Voordouw *et al.* 1984). When the L protein is removed from the holoenzyme, it becomes irreversibly denatured and cannot be reconstituted with S to form the hexadecameric holoenzyme. One approach towards understanding higher-plant RuBPCase subunit assembly is to express the L and S genes in *E. coli*, and then to determine the conditions that are required to obtain RuBPCase assembly and activity.

The synthesis and assembly of RuBPCase in photosynthetic cyanobacteria appears to follow a much simpler pathway than in higher plants. This is due to the essentially prokaryotic nature of these autotrophic organisms (Shinozaki *et al.* 1983; Stanier & Cohen-Bazire 1977) at both the level of gene organization and cellular compartmentalization. The genes encoding both the L and S subunits (*rbcL* and *rbcS*) of cyanobacteria have been cloned and sequenced (Curtis & Haselkorn 1983; Reichelt & Delaney 1983; Shinozaki & Sugiura 1983; Nierzwicki-Bauer *et al.* 1984). The genes are closely linked and appear to be cotranscribed (Nierzwicki-Bauer *et al.* 1984; Shinozaki & Sugiura 1985). The apparent simplicity of the cyanobacterial RuBPCase genes has made them particularly attractive as a model system for obtaining expression and assembly of enzymatically active RuBPCase in *E. coli*.

In this paper we describe the progress made in designing *E. coli* expression systems that are enabling the assembly and biochemical properties of higher-plant and cyanobacterial RuBPCase genes to be studied in *E. coli*.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Expression of higher-plant RuBPCase genes in *E. coli*

The first higher-plant genes to be cloned into a plasmid expression vector and produce high levels of protein in *E. coli* were the wheat and maize chloroplast genes encoding the L subunit of RuBPCase (Gatenby *et al.* 1981; Gatenby & Castleton 1982). Because of the close similarity between chloroplast and bacterial transcription–translation control sequences (for review, see Whitfeld & Bottomley 1983) the L genes could be expressed in *E. coli* using chloroplast-encoded transcription signals. However, to obtain high levels of maize L protein synthesis in *E. coli*, a bacteriophage lambda promoter was placed 5′ to the chloroplast gene insert. On induction of the lambda P_L promoter, about 2% of the total bacterial protein was found to be the maize L subunit. Additional studies showed that a chloroplast DNA sequence, located between the lambda promoter and the start of L translation, was acting as a transcriptional terminator in *E. coli* (Gatenby & Cuellar 1985). This chloroplast termination signal could be overcome by the presence of several different lambda anti-termination proteins, or by deleting the termination region from the chloroplast DNA insert. The biological significance of this sequence in the chloroplast is not known, but similar problems may have to be overcome when expressing other foreign genes in *E. coli*.

The ability to synthesize high levels of the maize L subunit in *E. coli* has allowed the native properties of the L subunit to be studied (Gatenby 1984). The protein synthesized in *E. coli* was slightly larger (54.3 kDa) than that found in maize chloroplasts (53.3 kDa). This result was consistent with a previous suggestion that higher-plant L subunit is first synthesized as a precursor protein and then post-translationally processed into a mature protein (Langridge 1981). Pulse–chase experiments showed that all of the L subunit made in *E. coli* was stable (Gatenby 1984). No RuBPCase activity could be detected in the *E. coli* extracts, not did the L protein form a stable enzyme–metal–CO_2–[^{14}C]carboxyarabinitol 1,5-biphosphate (CABP) ternary complex. These results were consistent with protein solubility studies, which showed that virtually all the maize L subunits synthesized in *E. coli* were insoluble. Therefore, the inactive and insoluble conformation of the maize L subunit is a natural property of the protein, which will have to be dealt with in attempts to assemble active RuBPCase in *E. coli*.

One solution to the problem of L solubility might be to co-express the L and S subunit genes in *E. coli*. The presence of the S subunit might prevent the L subunit from assuming an insoluble configuration, thus allowing assembly of L and S into active RuBPCase enzyme. This experiment, however, requires that the nuclear-encoded S subunit gene be cloned into an *E. coli* vector such that the the S polypeptide is synthesized without the amino-terminal transit peptide, which is presumably only necessary for transport of the protein into chloroplasts (for review, see Cashmore *et al.* 1985). Such constructions have been made by subcloning a region of the wheat *rbcS* gene, which encoded the mature S polypeptide, from a cDNA clone (Broglie *et al.* 1983) into plasmid and M13 vectors (van der Vies *et al.* 1986; figure 1). The resulting expression clones (p565 and M72) produce the S subunit as a *lacZ*::*rbcS* fusion protein in *E. coli*. The amino-terminals of the S fusion proteins contains 11 (p565) and 10 (M72) amino acids from the *lacZ* protein. The first four amino acids of the mature S polypeptide have also been deleted. Transcription and translation of the S fusion gene are initiated from *lacZ* promoter and ribosome-binding sites, respectively. In order to have a completely homologous L and S RuBPCase assembly system, a bacteriophage M13 clone (M68, figure 1), expressing

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

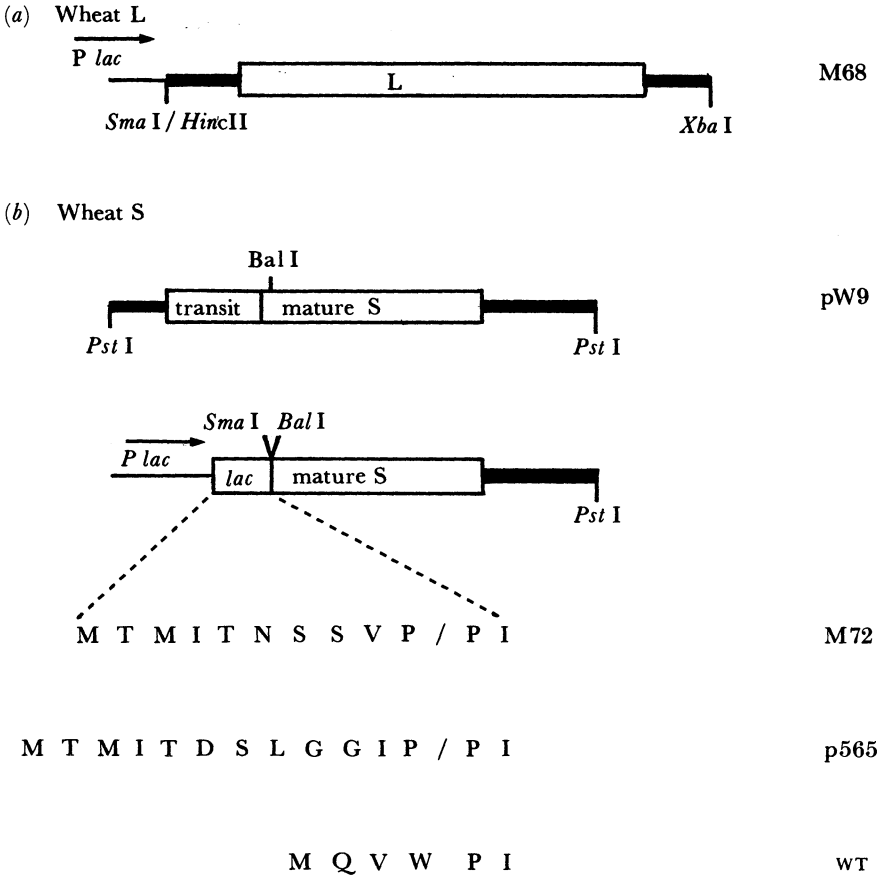

FIGURE 1. (*a*) Structure of wheat RuBPCase L subunit expression clone M68. A 1.9 kb *Hin*cII-*Xba*I DNA fragment encoding the entire wheat L gene (T. A. Dyer, *et al.* personal communication) was subcloned from pTac39 (Howe *et al.* 1982) into M13mp18 DNA that had been digested with *Sma*I and *Xba*I. Transcription of the L gene occurs from the *lac* promoter region (*lac* P, arrow indicates direction of transcription) and is under the control of the *lac* repressor. (*b*) Structure of wheat RuBPCase S clones. pW9 is a cDNA clone that was obtained from N.-H. Chua (Broglie *et al.* 1983). M72 was made by subcloning a 600 bp *Bal*I-*Pst*I fragment from pW9 into M13mp18 (van der Vies *et al.* 1986). p565 is a pBR325 derivative, made by S. J. Rothstein, which also contains a portion of the pW9 wheat S gene. These clones both express the mature wheat S polypeptide as a *LacZ*::S fusion protein. The amino-terminal sequences of the resulting S fusion proteins are shown, as is the wild-type (WT) S subunit sequence. The fusion junction is marked by a slash (/).

the wheat L subunit under the control of the *lacZ* promoter, was made by subcloning from pTac39 (Howe *et al.* 1982).

This combination of wheat L and S expression clones allowed us first to synthesize the S subunit in *E. coli* by means of a plasmid-borne S gene (p565), and then to infect those cells with M68 (wheat L). Phage M68 was also used to infect wild-type *E. coli* cells, to examine the properties of the wheat L subunit in the absence of S. Based on the solubility properties of the maize L subunit in *E. coli* (Gatenby 1984), we expected the wheat L subunit to be insoluble when synthesized in the absence of the S subunit. To examine this, soluble and insoluble fractions of *E. coli* cells infected with M68 in the presence (p565) and absence of wheat S were prepared, electrophoresed on SDS–polyacrylamide gels, and protein was detected by anti-wheat RuBPCase serum (figure 2).

In contrast to the maize L subunit insolubility in *E. coli*, about 60% of the wheat L

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

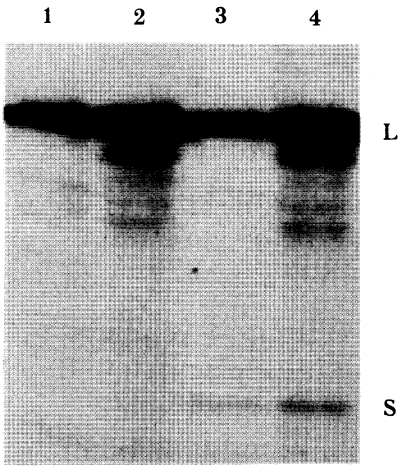

FIGURE 2. Solubility of the wheat L protein in *E. coli* cells in the presence and absence of the wheat S protein. Wheat L protein in insoluble (lane 1), and soluble (lane 2) fractions of *E. coli* strain TG2 infected with M68 phage, and insoluble (lane 3) and soluble (lane 4) fractions of TG2 (p565) infected with M68 phage. 200 ml cultures of TG2 (from T. Gibson, Δ (*lac-pro*), *thi*, *strA*, *hsdR*$^-$, *hsdM*$^-$, *recA*$^-$, *sclC*300::Tn10(TetR)F′ *traD*36, *proA*B, *lac*I^q, *lacZ* ΔM15) and TG2 (p565) (see figure 1) were grown in YT-broth containing 1 mM isopropyl β-D-thio-galactopyranoside (IPTG), to induce high-level transcription from *lac* promoter regions. The TG2 (p565) culture also contained ampicillin (50 μg ml^{-1}) to maintain selection for the p565 plasmid. The cultures were infected at an OD_{600} of 0.8 with 0.2 ml of M68 (wheat L gene, figure 1) phage and grown at 37 °C for another four hours. The cells were then collected by centrifugation, and soluble and insoluble protein extracts were prepared as described by Gatenby *et al.* (1985). Equal amounts of *E. coli* soluble and insoluble proteins were electrophoresed on a 15% SDS–acrylamide gel, transferred to nitrocellulose filter, and immunolabelled with anti-wheat RuBPCase serum and ^{125}I-labelled protein A (Gatenby *et al.* 1985).

subunit was found in the soluble fraction (figure 2; compare lanes 1 and 2). The presence of the S subunit did not seem to have any effect on the amount of L that was found in the soluble fraction. About 80% of the *lacZ*::wheat S protein synthesized in *E. coli* was found in the soluble fraction. RuBPCase assays were carried out on the soluble fractions from both strains of M68 (L) infected cells (TG2 and TG2 (p565); see legend to figure 2 for the genotype of TG2). Despite the apparent solubility of the wheat L in *E. coli* and the presence of the S subunit, no RuBPCase activity could be detected in these fractions.

The soluble proteins from cells expressing both L and S subunits were also fractioned on sucrose gradients to examine subunit assembly. Figure 3 shows a Western blot of fractions of such a sucrose gradient. Most of the S protein was found at the top of the gradient (figure 3, lanes 15–19). A small amount of S protein was also found in fractions comigrating with L. However, the most likely explanation for this result is protein aggregation rather than specific subunit assembly. The wheat L protein was found distributed throughout the gradient, but with a peak abundance of the protein sedimenting at a rate indicating a molecular mass of about 750 kDa (figure 3, lanes 3–7). A sucrose gradient of soluble proteins from cells synthesizing L in the absence of S was also run in parallel with the one grown in figure 3, and an identical distribution of L was observed (data not shown).

Thus, it appears that while the wheat L polypeptide is soluble in *E. coli*, the structure of the protein is that of a large protein aggregate, where the L subunits are probably not accessible for specific assembly with the S subunit. Perhaps the L subunit binding protein (Barraclough & Ellis 1980; Bloom *et al.* 1983) functions in chloroplasts to prevent this L aggregate from forming, and will, in fact, be necessary to obtain L and S assembly in *E. coli*. Another obvious

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

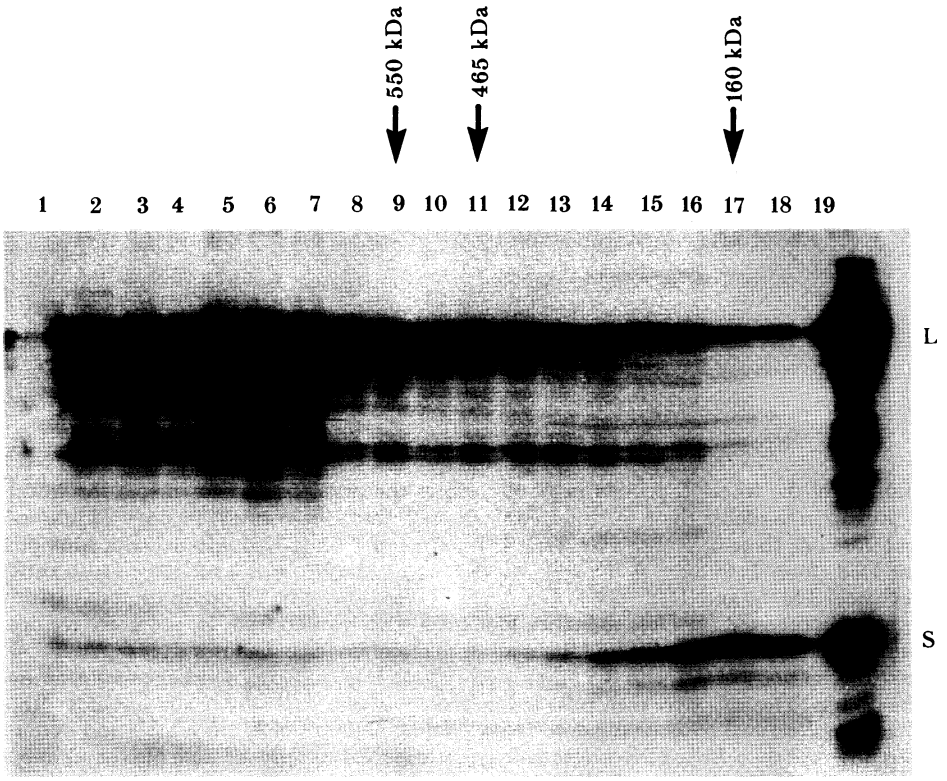

FIGURE 3, Sucrose gradient analysis of wheat L and S protein from a soluble protein extract of *E. coli*. Protein extracts were prepared as for figure 2 and then centrifuged on a sucrose gradient (Gatenby *et al.* 1985). Aliquots from each gradient fraction were then separated by electrophoresis on a 15% SDS–acrylamide gel and Western blotted. Fraction 1 is from the bottom of the gradient; fraction 19 is from the top. The sucrose gradient had an internal molecular-mass marker of β-galactosidase (465 kDa), which was assayed in samples from each gradient fraction. An identical calibration sucrose gradient was always centrifuged at the same time; this contained RuBPCase holoenzyme (550 kDa), β-galactosidase (465 kDa) and aldolase (160 kDa). The sedimentation peaks of the three different protein molecular-mass markers are designated by arrows.

explanation for the lack of RuBPCase activity is that the lacZ::wheat S fusion protein might not be competent for assembly with the wheat L subunit because of its altered amino terminus (see Discussion). New experiments have been initiated to construct a wheat S expression plasmid, which will direct the synthesis of a completely unaltered S protein in *E. coli*, to test this possibility.

EXPRESSION OF CYANOBACTERIAL RUBPCASE GENES IN *E. COLI*

Recently, several groups have reported the synthesis of active RuBPCase in *E. coli* by expressing the L and S genes from several different species of cyanobacterium (Gatenby *et al.* 1985; Gurevitz *et al.* 1985; Tabita & Small 1985; Christeller *et al.* 1985). In each case, the RuBPCase L and S genes were cloned into *E. coli* plasmids such that the cotranscription of these genes was dependent on a plasmid-borne promoter. In addition, the active enzymes were shown to contain the S polypeptide and to be assembled into multimeric L_8S_8 when synthesized from *lac* promoter expression plasmids (Gurevitz *et al.* 1985; Tabita & Small 1985). However, when *Synechococcus* L and S genes were expressed in *E. coli* by means of two different *E. coli*

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

transcription vectors, the ratios of soluble L and S polypeptides were found to be non-stoichiometric (Gatenby *et al.* 1985). Estimates of subunit composition, based on RuBPCase specific activity (Andrews & Ballment 1983) and protein gel electrophoresis, were consistent with an average RuBPCase subunit structure of L_8S_{2-5}, with the S content varying with the expression vector used (Gatenby *et al.* 1985). The analysis of enzyme structure and RuBPCase activity was done in a way that made it impossible to determine if the carboxylase activity was all associated with L_8S_8 molecules or with L_8S_{2-5} structures. In contrast, when the *Anabaena* and *Anacystis*† RuBPCase was synthesized and assembled in *E. coli*, the enzyme was found to have about the same specific activity as that purified from the cyanobacteria (Gurevitz *et al.* 1985; Tabita & Small 1985; Christeller *et al.* 1985).

Gurevitz *et al.* (1985) showed that the ratio of soluble L subunit to S subunit was consistent with an L_8S_8 structure, but that in addition a large amount of unassembled L was also synthesized, all of which was found as insoluble protein. Based on these observations, they proposed a model for the assembly of RuBPCase L_8S_8 holoenzyme from L and S subunit heterodimers, the S subunit being postulated to play an essential role in maintaining L solubility and in initiating holoenzyme assembly.

Andrews & Ballment (1983) used purified *Synechococcus* RuBPCase in a reconstitution system *in vitro* to study RuBPCase assembly and activity. With a mild acid treatment, they were able to strip about 90% of the S subunit from purified *Synechococcus* holoenzyme . The L subunits were then recovered as an intact L_8 core that could be used in reconstitution experiments with purified S subunits. They found that the specific activity of the reconstituted holoenzyme had a linear dependence on the amount of S subunit present and that an almost fully reconstituted L_8S_8 enzyme could be recovered. Based on these experiments, they proposed that only L–S complexes were catalytically active, and that RuBPCase activity was associated with L–S pairs in a L_8S_n structure, regardless of the value of n. These experiments demonstrated the essential role of S in maintaining the holoenzyme in a catalytically active structure. However, questions about the role of the S subunit, if any, in initiating L assembly, or in maintaining L as a soluble protein, still remain unresolved. The reconstitution experiments showed that, once an L_8S_8 structure has been formed, S can be reversibly removed from the L_8 core structure. The fact that L stays in an L_8 structure in the absence of S, however, still leaves open the possibility that S might play a role in the initial formation of the L_8 core, as has been postulated by Gurevitz *et al.* (1985). Once formed, the L_8 structure would no longer be dependent on S for maintaining stability or solubility.

To examine these issues further and to determine the most efficient expression system for obtaining high levels of *Synechococcus* RuBPCase activity in *E coli*, we have subcloned the L and S genes from pDB50 (figure 4, Gatenby *et al.* 1985; van der Vies *et al.* 1986) into different bacteriophage M13 and pUC plasmid vectors. In each case, the expression of the isolated RuBPCase subunit gene was under the control of the *E. coli lacZ* promoter. The ability to synthesize the L subunit in *E. coli*, in the absence of S, allowed us to investigate the self-assembly potential of the L polypeptide, and to determine the solubility properties of the L subunit. In addition, by infecting a strain bearing an L subunit plasmid with the appropriate S subunit M13 clone, the S subunit gene could be introduced into the *E. coli* cells and expressed synchronously. Finally, we also used the M13 S gene phage (M67, figure 4) to infect an

† *Anacystis* 6301 and *Synechococcus* 6301 are the same cyanobacterial species (Christeller *et al.* 1985).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

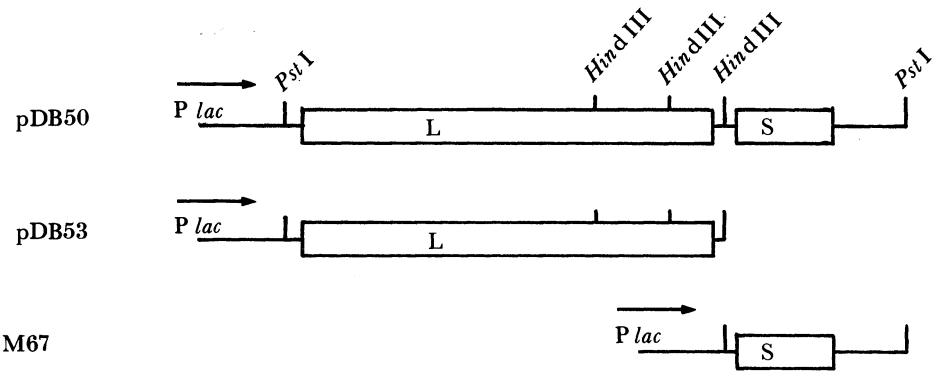

FIGURE 4. Structure of *Synechococcus* RuBPCase L and S expression clones. The details of the DNA constructions have been described in Gatenby *et al.* (1985) and van der Vies *et al.* (1986). All three of the clones express the *Synechococcus* RuBPCase genes with a vector *lac* promoter region (P *lac*; arrow shows the direction of transcription).

E. coli strain containing pDB50, which synthesizes both L and S but only assembles into a L_8S_{5-6} RuBPCase, in an attempt to drive the assembly of all the soluble L subunits into L_8S_8 structures.

After induction of the *lacZ* promoters and infection of the plasmid cells with the appropriate M13 phage, assembly and activity of the RuBPCase subunits synthesized in *E. coli* were assayed in cell extracts. The specific activity of RuBPCase activity in *E. coli* cells expressing L and S from pDB50 (table 1) was 21.5 nmol min^{-1} mg^{-1} of soluble protein at 30 °C. A ^{14}C-labelled transition-state analogue of 2-carboxyarabinitol 1,5-bisphosphate (CABP), which binds tightly to the L subunit active site (Miziorko 1979; Miziorko & Sealy 1980), was used to quantitate the amount of L subunit present in the soluble fraction of pDB50 *E. coli* extracts. Calculations from these measurements showed that the RuBPCase activity of the pDB50 cell extract was consistent with the formation of L_8S_{5-6} enzyme. The same pDB50 cells were infected with M67 (S gene, figure 4) and the RuBPCase activity was measured (table 1). The increase in specific activity relative to the L concentration (determined by CABP binding) suggested that the average subunit structure of RuBPCase in M67-infected pDB50 cells had increased slightly compared with that of an L_8S_{6-7} enzyme. In addition, the amount of RuBPCase activity per

TABLE 1. *SYNECHOCOCCUS* RuBPCASE ACTIVITY IN *E. COLI*

plasmid	phage	RuBPCase[1]	specific activity[2]	estimated structure
pSV55 (L+S)	—	8.0	1.07	L_8S_{2-3}
pDB50 (L+S)	—	21.5	1.93	L_8S_{5-6}
pDB50 (L+S)	M67 (S cyano)	64.8	2.24	L_8S_{6-7}
pDB53 (L)	—	no detectable activity		—
pDB53 (L)	M67 (S cyano)	33.0	ND[3]	L_8S_{4-5}
Synechococcus	RuBPCase *in vivo*		3.1[4]	L_8S_8

These determinations of *Synechococcus* RuBPCase activity in *E. coli* extracts were taken from Gatenby *et al.* (1985) and van der Vies (1986).

[1] Units: nmol CO_2 fixed (min at 30 °C)$^{-1}$ (mg *E. coli* protein)$^{-1}$.

[2] Units: μmol CO_2 fixed (min at 30 °C)$^{-1}$ (mg L subunit of RuBPCase protein)$^{-1}$.

[3] ND, not determined.

[4] From Andrews & Ballment (1983).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

milligram of *E. coli* protein increased threefold. This result suggests that the level of S synthesis from pDB50 was limiting the amount of active RuBPCase assembled in *E. coli* and that this could be increased by infecting the pDB50 with M13 phage expressing the S gene product.

To examine the properties of the *Synechococcus* L subunit in *E. coli* when the protein was expressed in the absence of the S polypeptide, soluble protein extracts were prepared from *E. coli* cells containing pDB53 (L gene alone, figure 4). In a parallel experiment, cells containing pDB53 were infected with M67 (*Synechococcus* S gene) and soluble extracts were also prepared. Both protein samples were then sedimented on sucrose gradients and aliquots from each fraction were assayed for RuBPCase activity and analysed by SDS–gel electrophoresis and Western blotting with anti-wheat RuBPCase serum.

The results of these experiments showed that about 20% of the soluble L polypeptide, synthesized in the absence of S, sedimented at a position in the gradient consistent with an L_1 or L_2 structure (van der Vies *et al.* 1986). The remaining L protein had a mobility in the sucrose gradient that was consistent with the formation of an L_8 structure. When RuBPCase activity was measured in each fraction of the uninfected pDB53 gradient, no activity could be detected.

The sucrose gradients prepared from pDB53 (*Synechococcus* L gene) cells that had also been infected with M67 (*Synechococcus* S gene) showed a different sedimentation profile for the L protein (van der Vies *et al.* 1986). Western blot analysis of the gradient fractions showed that L monomer or dimer structures were not found in these extracts. Instead, the major peak of the L protein in the gradient sedimented at about 550 kDa. RuBPCase assays on aliquots from each sucrose gradient fraction showed that RuBPCase activity could be measured from these extracts and that the peak of RuBPCase activity coincided with the peak of the L subunit protein.

These results confirm the observations of Andrews & Ballment (1983) that there was an absolute requirement for the S subunit in order to assemble L structures that exhibit RuBPCase activity. The specific activity of the RuBPCase in the M67-infected pDB53 *E. coli* cells was calculated to be 33.0 nmol min^{-1} mg^{-1} of *E. coli* soluble protein (table 1). This is a greater amount of RuBPCase activity than was found in *E. coli* cells expressing *Synechococcus* L and S from pDB50 (table 1), which contains the L and S genes as arranged *in vivo* (figure 4). Previous investigators have suggested that the tandem arrangement of the cyanobacterial L and S genes might be important in localizing the translation products of the two genes so that assembly of the L and S polypeptides could occur efficiently (Gurevitz *et al.* 1985). This is clearly not the case with the *Synechococcus* subunits in *E. coli*, because high amounts of RuBPCase activity could be recovered from *E. coli* cells where the two subunits were transcribed individually. The solubility of the *Synechococcus* L protein synthesized in *E. coli*, in the presence and absence of the S subunit, was examined by doing the following experiment. Equal amounts of *E. coli* soluble protein from samples of cells containing either pDB50 (*Synechococcus* L and S) or pDB53 (*Synechococcus* L gene) were fractioned by electrophoresis on a sodium dodecyl sulphate (SDS)–polyacrylamide gel and Western blotted with wheat RuBPCase antisera (figure 5). The different RuBPCase plasmid cultures were either infected with M67 (*Synechococcus* S gene) or left uninfected.

The amount of *Synechococcus* L polypeptide present in the soluble fraction from cultures containing pDB53, was increased only slightly by the presence of the S subunit (figure 5, compare lanes 2 and 3). In addition, when the amount of S subunit in pDB50 (L and S)-bearing cells was increased by M67 phage infection (figure 5, lanes 4 and 5) there was also no detectable

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

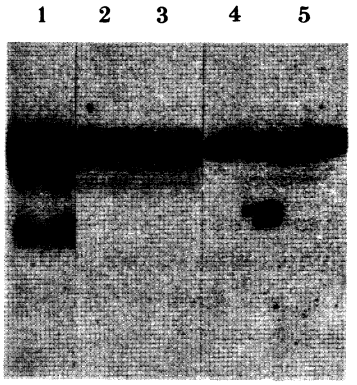

FIGURE 5. Solubility of *Synechococcus* RuBPCase L subunit protein in *E. coli* in the presence and absence of the S polypeptide. Soluble protein extracts were prepared from cells and Western blotted as before (legends to figure 2). Lane 1, total wheat leaf protein; lane 2, TG2 (pDB53) uninfected; lane 2, TG2 (pDB53) infected with M67; lane 4, TG2 (DB50) uninfected; lane 5, TG2 (pDB50) infected with M67.

change in the amount of soluble L protein. In the same experiment, however, RuBPCase activity in the pDB50 cells was increased by threefold on infection, suggesting that more S had been synthesized and assembled into active RuBPCase because of the M67 infection. Taken together, these experiments show that the S subunit is not required for the solubility of *Synechococcus* L in *E. coli*, nor is it required for the self-assembly of L into multimeric structures. The S subunit is, however, absolutely required for RuBPCase activity in *E. coli*.

DISCUSSION

The genes encoding the L and S subunits of wheat RuBPCase have been cloned into bacterial vectors, and both the gene products have been expressed in *E. coli* together. No RuBPCase activity could be detected in *E. coli* cell extracts that contained both the wheat RuBPCase gene products. In contrast to maize L subunit protein synthesised in *E. coli*, which was almost completely insoluble (Gatenby 1984), about 60% of the wheat L protein was found to be soluble. The synthesis of the wheat S polypeptide in *E. coli* cells before the introduction of a phage-borne wheat L gene (M68) did not alter the fraction of L protein present as soluble protein (figure 2). Sedimentation studies (figure 3) demonstrated that the L subunits form a variety of structures that are found throughout the sucrose gradient. However, the major L sedimentation product was found at about 750 kDa, suggesting that the wheat L protein forms a high-molecular-mass aggregate that is incapable of interacting with S to form an active RuBPCase holoenzyme.

The failure to obtain an active wheat RuBPCase in *E. coli*, despite the synthesis of both subunits, could be due to several factors. First, the L gene product may be synthesized in wheat chloroplasts as a precursor protein that is then processed to a protein product that is competent for assembly. There is some evidence that this occurs in spinach, pea and maize (Langridge 1981; Roy *et al.* 1982; Gatenby 1984). Similarly, there is the question of the effects of the lac::wheat S fusion on the functionality of the S subunit. We think that the effects of the fusion peptide on the wheat S are minor because this protein can still assemble with the *Synechococcus* L subunit to form an active hybrid RuBPCase enzyme (van de Vies *et al.* 1986). Even so, both of these factors can be eliminated by constructing the appropriate L and S expression clones. The most important consideration for future efforts

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

to obtain assembly and activity of higher plant RuBPCase in *E. coli* is the role of the L subunit binding protein in RuBPCase assembly (Barraclough & Ellis 1980; Bloom *et al.* 1983). It may very well prove necessary to express the L subunit binding protein gene in *E. coli*, in addition to the L and S subunit genes, to develop a workable system for the study of higher-plant RuBPCase activity in bacteria.

The results obtained from the expression of cyanobacterial RuBPCase genes in *E. coli* are, however, much more straightforward than those obtained from higher-plant genes. By using cloned *Synechococcus* RuBPCase genes, we have been able to demonstrate that the two subunits will assemble efficiently in *E. coli* to form an active RuBPCase enzyme that has about the same specific activity as that found in the cyanobacteria (table 1). The fact that both subunits can be efficiently synthesized and assembled in *E. coli* from genes carried on either plasmid or M13 phage vectors, now makes it possible to introduce site-specific alterations in the coding portion of either gene. The phenotype effects of the mutations, if any, on RuBPCase assembly and activity could then be rapidly determined in *E. coli* cell extracts. In this way we should be able to begin learning which polypeptide domains within the L and S subunits are important for holoenzyme assembly and catalytic function. Only then will we have much hope of designing a RuBPCase enzyme with the enzymic properties that will result in the more efficient fixation of atmospheric CO_2 and in higher crop yields.

We are grateful to N.-H. Chua for sending us the wheat cDNA clone pW9. We would also like to thank T. A. Dyer and R. B. Flavell for helpful comments on the manuscript, and M. Longstaff for technical assistance. Financial support was provided by the AFRC. D.B. was supported by a N.S.F. postdoctoral fellowship in Plant Biology (PCM-8412385).

References

Andrews, T. J. & Ballment, B. 1983 The function of the small subunits of ribulose bisphosphate carboxylase-oxygenase. *J. biol. Chem.* **258**, 7514–7518.

Barraclough, R. & Ellis, R. J. 1980 Protein synthesis in chloroplasts. IX. Assembly of newly synthesized large subunits into ribulose bisphosphate carboxylase in isolated intact pea chloroplasts. *Biochim. biophys. Acta* **608**, 19–31.

Bloom, M. V., Miles, P. & Roy, H. 1983 Light-dependent assembly of ribulose-1,5-bisphosphate carboxylase. *Proc. natn. Acad. Sci. U.S.A* **80**, 1013–1017.

Broglie, R., Coruzzi, G., Lamppa, G., Keith, B. & Chua, N.-H. 1983 The small subunit of wheat ribulose-1,5-bisphosphate carboxylase. *Biotechnology* **1**, 55–61.

Cashmore, A., Szabo, L., Timki, M., Kausch, A., Van den Broeck, G., Schreier, P., Bohnert, H., Herrera-Estrella, L., Van Montagu, M. & Schell, J. 1985 Import of polypeptides into chloroplasts. *Biotechnology* **3**, 803–808.

Christeller, J. T., Tezaghi, B. E., Hill, D. F. & Laing, W. A. 1985 Activity expressed from cloned *Anacystis nidulans* large and small subunit ribulose bisphosphate carboxylase genes. *Pl. Molec. Biol.* **5**, 257–263.

Curtis, S. E. and Haselkorn, R. 1983 Isolation and sequence of the gene for the large subunit of ribulose-1,5-bisphosphate carboxylase from the cyanobacterium *Anabaena* 7120. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1825–1829.

Dobberstein, B., Blobel, G. & Chua, N.-H. 1977 *In vitro* synthesis and processing of a putative precursor for the small subunit of ribulose-1,5-bisphosphate carboxylase of *Chlamydomonas reinhardtii*. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1082–1085.

Ellis, R. J. 1981 Chloroplast proteins: synthesis, transport and assembly. *A. Rev. Pl. Physiol.* **32**, 111–137.

Ellis, R. J. & Gatenby, A. A. 1984 Ribulose bisphosphate carboxylase: properties and synthesis. In *The genetic manipulation of plants and its application to agriculture* (ed P. J. Lea & G. R. Steward), pp. 41–60. Oxford University Press.

Gallagher, T. F. & Ellis, R. J. 1982 Light-stimulated transcription of genes for two chloroplast polypeptides in isolated pea leaf nuclei. *EMBO J.* **1**, 1493–1498.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Gatenby, A. A. 1984 The properties of the large subunit of maize ribulose bisphosphate carboxylase/oxygenase synthesized in *Escherichia coli*. *Eur. J. Biochem.* **144**, 361–366.

Gatenby, A. A. & Castleton, J. A. 1982 Amplification of maize ribulose bisphosphate carboxylase large subunit synthesis in *E. coli* by transcriptional fusion with the lambda *N* operon. *Molec. gen. Genet.* **185**, 424–429.

Gatenby, A. A. & Cuellar, R. E. 1985 Antitermination is required for readthrough transciption of the maize *rbcL* gene by a bacteriophage promoter in *Escherichia coli*. *Eur. J. Biochem.* **153**, 355–359.

Gatenby, A. A., Castleton, J. A. & Saul, W. W. 1981 Expression in *E. coli* of maize and wheat chloroplast genes for the large subunit of ribulose bisphosphate carboxylase. *Nature, Lond.* **291**, 117–121.

Gatenby, A. A., van de Vies, S. M. & Bradley, D. 1985 Assembly in *E. coli* of a functional multi-subunit ribulose bisphosphate carboxylase from a blue-green alga. *Nature, Lond.* **314**, 617–620.

Goeddel, D. V., Heyneker, H. L., Hozumi, T., Arentzeu, R., Itakura, K., Yansura, D. G., Ros, M. J., Miozzari, G., Crea, R. & Seeburg, P. H. 1979*a* Direct expression in *Escherichia coli* of a DNA sequence coding for human growth hormone. *Nature, Lond.* **281**, 544–548.

Goeddel, D. V., Kleid, D. G., Bolivar, G., Heyneker, H. L., Yansura, D. G., Crea, R., Hirose, T., Kraszewski, A., Itakura, K. and Riggs, A. D. 1979*b* Expression in *Escherichia coli* of chemically synthesized genes for human insulin. *Proc. natn. Acad. Sci. U.S.A.* **76**, 106–110.

Gurevitz, M., Somerville, C. R. & McIntosh, L. 1985 Pathway of assembly of ribulose-bisphosphate carboxylase/oxygenase from *Anabaena* 7120 expressed in *Escherichia coli*. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6546–6550.

Howe, C. J., Bowman, C. M., Dyer, T. A. & Gray, J. C. 1982 Localization of wheat chloroplast genes for the beta and epsilon subunits of ATP synthase. *Molec. gen. Genet.* **186**, 525–530.

Langridge, P. 1981 Synthesis of the large subunit of spinach ribulose bisphosphate carboxylase may involve a precursor polypeptide. *FEBS Lett.* **123**, 85–89.

Lorimer, G. H. & Miziorko, H. M. 1980 Carbamate formation on the ϵ-amino group of a lysyl residues as the basis for the activation of ribulose bisphosphate carboxylase by CO_2 and Mg^{2+}. *Biochemistry, Wash.* **19**, 5321–5328.

Miziorko, H. M. 1979 Ribulose-1,5-bisphosphate carboxylase. Evidence in support of the existence of district CO_2 activator and CO^2 substrate sites. *J. biol. Chem.* **254**, 270–272.

Miziorko, H. M. & Lorimer, G. H. 1983 Ribulose-1,5-bisphosphate carboxylase-oxygenase. *A. Rev. Biochem.* **52**, 507–535.

Miziorko, H. M. & Sealy, R. C. 1980 Characterization of the ribulosebisphosphate carboxylase–carbon dioxide–divalent cation–carboxypentitol bisphosphate complex. *Biochemistry, Wash.* **19**, 1167–1172.

Miziorko, H. M., Behuke, C. E. & Houkom, E. C. 1982 Protein liganding to the activator cation of ribulose bisphosphate carboxylase. *Biochemistry, Wash.* **21**, 6669–6674.

Nierzwicki–Bauer, S. A., Curtis, S. E. & Haselkorn, R. 1984 Cotranscription of genes encoding the small and large subunits of ribulose-1,5-bisphosphate carboxylase in the cyanobacterium *Anabaena* 7120. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5961–5965.

Reichelt, B. Y. & Delaney, S. F. 1983 The nucleotide sequence for the large subunit of ribulose-1,5-bisphosphate carboxylase from a unicellular cyanobacterium, *Synechococcus* PCC6301. *DNA* **2**, 121–129.

Roy, H., Bloom, M., Milos, P. & Monroe, M. 1982 Studies on the assembly of large subunits of ribulose bisphosphate carboxylase in isolated pea chloroplasts. *J. Cell Biol.* **94**, 20–27.

Shinozaki, K. & Sugiura, M. 1983 The gene for the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase is located close to the gene for the large subunit in the cyanobacterium. *Nucl. Acids Res.* **11**, 6957–6954.

Shinozaki, K., Yamada, C., Takahata, N. & Sugiura, M. 1983 Molecular cloning and sequence analysis of the cyanobacterial gene for the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4050-4054.

Smith, S. M. and Ellis, R. J. 1979 Processing of the small subunit precursor of ribulose bisphosphate carboxylase and its assembly into whole enzyme are stromal events. *Nature, Lond.* **278**, 662–664.

Spreitzer, R. J. & Mets, L. J. 1980 Non-mendelian mutation affecting ribulose-1,5-bisphosphate carboxylase structure and activity. *Nature, Lond.* **285**, 114–115.

Stanier, R. Y. & Cohen-Bazire, G. 1977 Phototrophic prokaryotes: the Cyanobacteria. *A. Rev. Microbiol.* **31**, 225–274.

Tabita, F. R. & Small, C. L. 1985 Expression and assembly of active cyanobacterial ribulose-1,5-bisphosphate carboxylase/oxygenase in *Escherichia coli* containing stoichiometric amounts of large and small subunits. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6100–6103.

van der Vies, S. M., Bradley, D. & Gatenby, A. A. 1986 Genetic complementation and assembly of active hybrid ribulose-1,5-bisphosphate carboxylase from wheat and *Synechococcus* genes expressed in *Escherichia coli*. *EMBO J.* (Submitted.)

Voordouw, G., van der Vies, S. M. & Bouwmeister, P. P. 1984 Disassociation of ribulose-1,5-bisphosphate carboxylase/oxygenase from spinach by urea. *Eur. J. Biochem.* **141**, 313–318.

Whitfeld, P. R. & Bottomley, W. 1983 Organisation and structure of chloroplast genes. *A. Rev. Pl. Physiol.* **34**, 279–310.

Winter, G., Fersht, A. R., Wilkinson, A. J., Zoller, M. J. & Smith, M. 1982 Redesigning enzyme structure by site-directed mutagenesis: tyrosyl tRNA synthetase and ATP binding. *Nature, Lond.* **299**, 756–758.

Zoller, M. J. & Smith, M. 1983 Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors. *Meth. Enzym.* **100**, 468–500.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

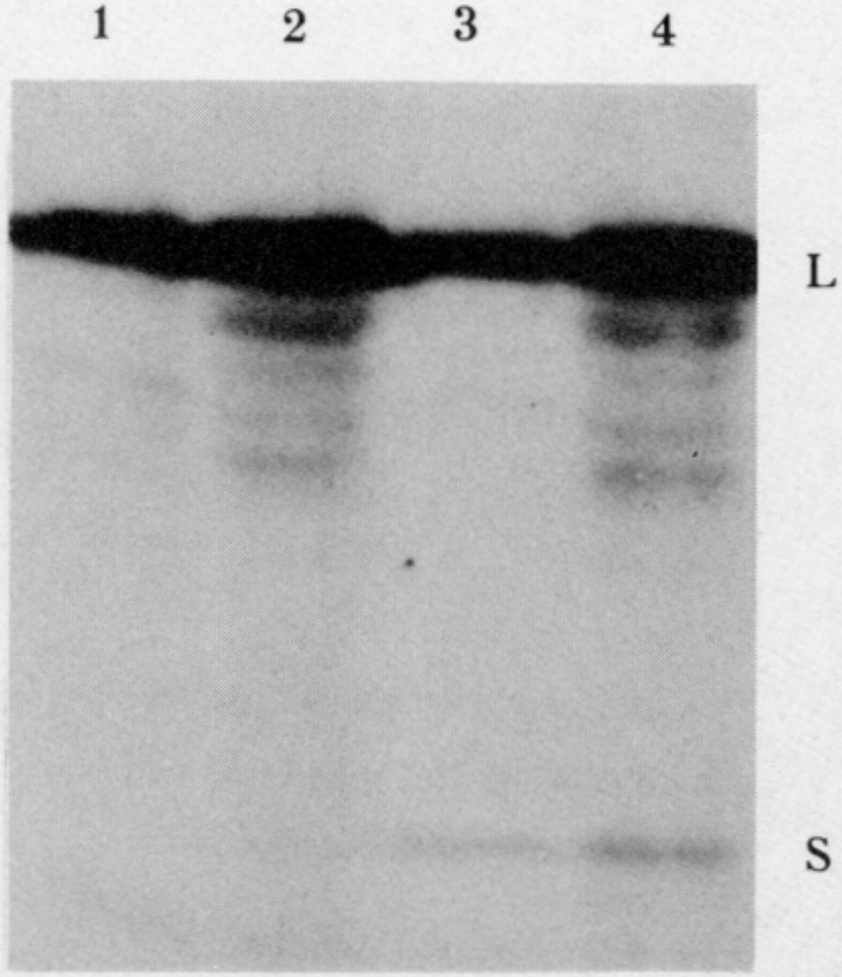

FIGURE 2. Solubility of the wheat L protein in *E. coli* cells in the presence and absence of the wheat S protein. Wheat L protein in insoluble (lane 1), and soluble (lane 2) fractions of *E. coli* strain TG2 infected with M68 phage, and insoluble (lane 3) and soluble (lane 4) fractions of TG2 (p565) infected with M68 phage. 200 ml cultures of TG2 (from T. Gibson, Δ (*lac-pro*), *thi*, *strA*, *hsdR*$^-$, *hsdM*$^-$, *recA*$^-$, *sclC*300::Tn10(TetR)F′ *traD*36, *proA*B, *lac*I^q, *lacZ* ΔM15) and TG2 (p565) (see figure 1) were grown in YT-broth containing 1 mM isopropyl β-D-thio-galactopyranoside (IPTG), to induce high-level transcription from *lac* promoter regions. The TG2 (p565) culture also contained ampicillin (50 μg ml^{-1}) to maintain selection for the p565 plasmid. The cultures were infected at an OD_{600} of 0.8 with 0.2 ml of M68 (wheat L gene, figure 1) phage and grown at 37 °C for another four hours. The cells were then collected by centrifugation, and soluble and insoluble protein extracts were prepared as described by Gatenby *et al.* (1985). Equal amounts of *E. coli* soluble and insoluble proteins were electrophoresed on a 15% SDS–acrylamide gel, transferred to nitrocellulose filter, and immunolabelled with anti-wheat RuBPCase serum and ^{125}I-labelled protein A (Gatenby *et al.* 1985).

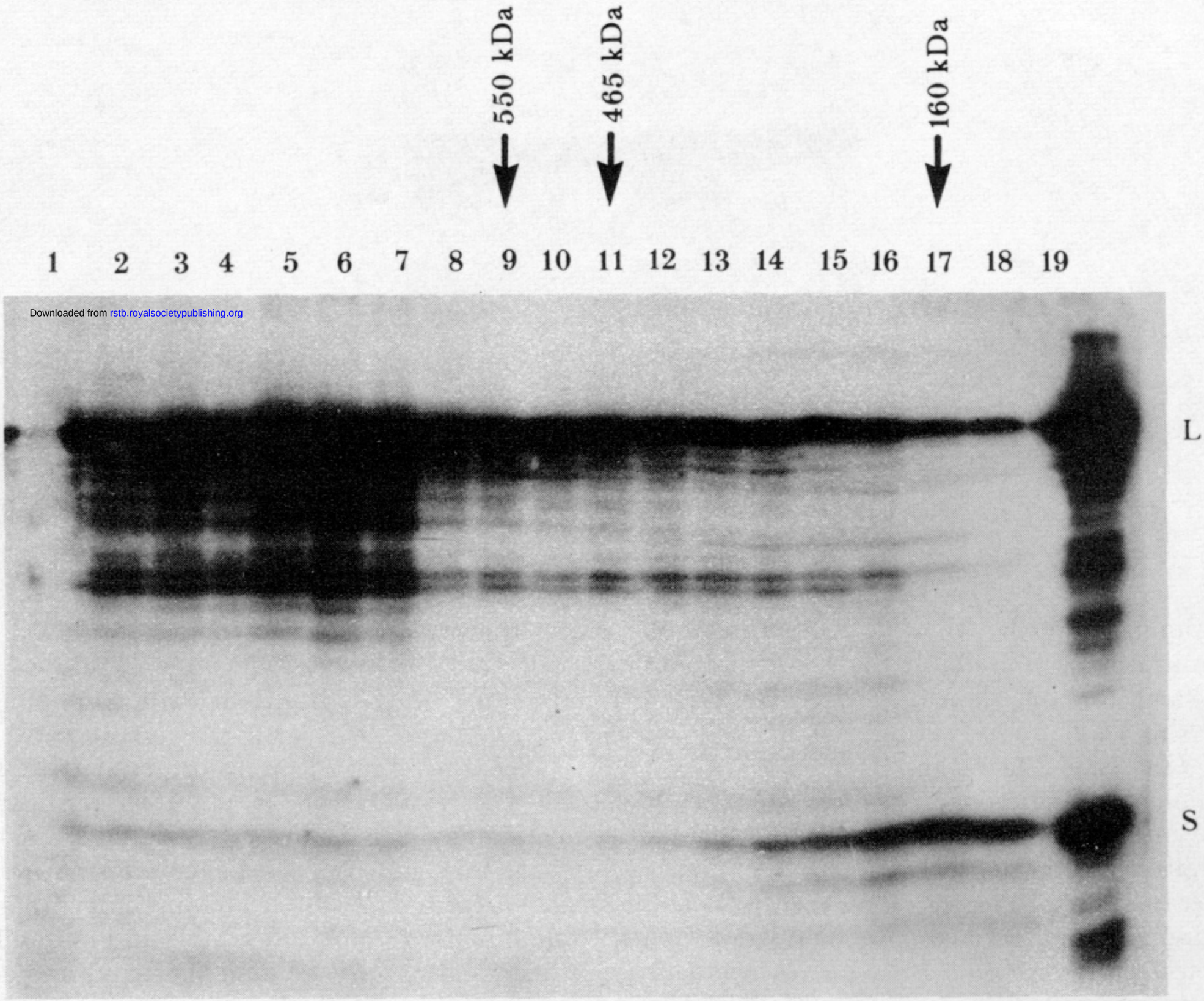

FIGURE 3, Sucrose gradient analysis of wheat L and S protein from a soluble protein extract of *E. coli*. Protein extracts were prepared as for figure 2 and then centrifuged on a sucrose gradient (Gatenby *et al.* 1985). Aliquots from each gradient fraction were then separated by electrophoresis on a 15% SDS–acrylamide gel and Western blotted. Fraction 1 is from the bottom of the gradient; fraction 19 is from the top. The sucrose gradient had an internal molecular-mass marker of β-galactosidase (465 kDa), which was assayed in samples from each gradient fraction. An identical calibration sucrose gradient was always centrifuged at the same time; this contained RuBPCase holoenzyme (550 kDa), β-galactosidase (465 kDa) and aldolase (160 kDa). The sedimentation peaks of the three different protein molecular-mass markers are designated by arrows.

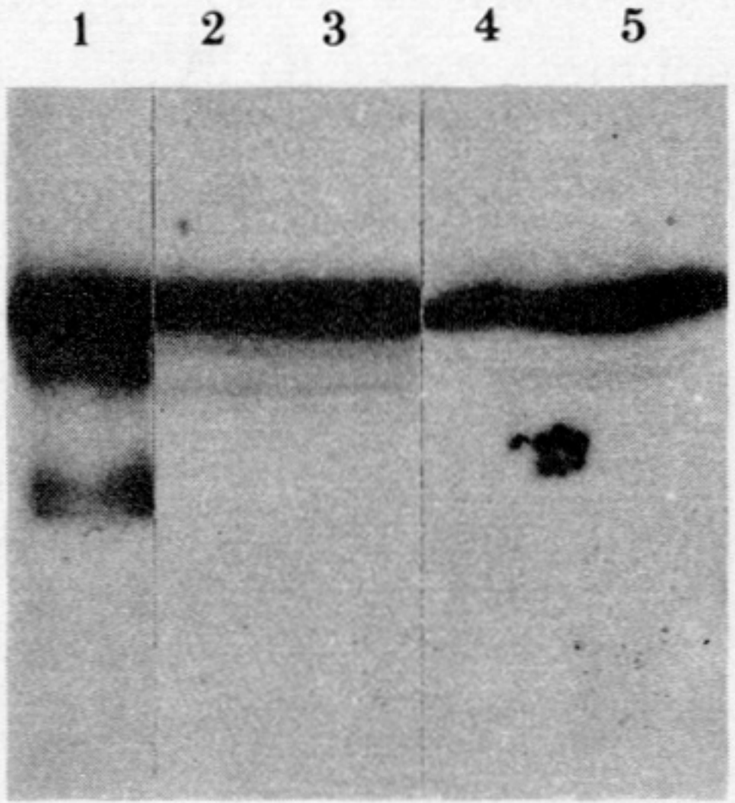

FIGURE 5. Solubility of *Synechococcus* RuBPCase L subunit protein in *E. coli* in the presence and absence of the S polypeptide. Soluble protein extracts were prepared from cells and Western blotted as before (legends to figure 2). Lane 1, total wheat leaf protein; lane 2, TG2 (pDB53) uninfected; lane 2, TG2 (pDB53) infected with M67; lane 4, TG2 (DB50) uninfected; lane 5, TG2 (pDB50) infected with M67.