Expression of Bacterial Rubisco Genes in Escherichia coli

S. Gutteridge, A. L. Phillips, C. A. Kettleborough, M. A. J. Parry and A. J. Keys

Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

433

Expression of bacterial Rubisco genes in *Escherichia coli*

By S. Gutteridge, A. L. Phillips, C. A. Kettleborough,
M. A. J. Parry and A. J. Keys

Rothamsted Experimental Station, Department of Biochemistry, Harpenden, Herts AL5 2JQ, U.K.

The structural genes for three forms of Rubisco have been isolated from bacteria and introduced into various plasmids. Apart from details of the sequences which have been obtained from these constructs, they are now being exploited for mutagenesis to determine the identity and specific function of the individual amino acid residues that compose the active site. These methods have been applied to a plasmid that contains the structural gene for the simplest form of Rubisco from *Rhodospirillum rubrum*, to obtain mutant enzymes with altered activity. The construct pRR2119 is also expressed to very high levels in *Escherichia coli* and enough recombinant protein of both the wild-type and mutant enzymes can be obtained for detailed physico-chemical studies. Other vectors have now been constructed, containing the genes of prokaryotic Rubisco that assemble into an active form I enzyme. The levels of expression are acceptable and the product is similar to the authentic enzyme. These constructs are now being used for mutagenesis *in vitro* to attempt to alter the relative rates of the oxygenase and carboxylase activities.

Introduction

A comparison of the activities of different species of Rubisco from photosynthetic bacteria to higher plants indicates that the partitioning of the substrate ribulose bisphosphate (RuBP) into carboxylation and oxygenation can vary by an order of magnitude (Jordan & Ogren 1981, 1983). Analysis of the catalytic mechanism has identified the enediolate form of RuBP as the reactive intermediate over which the gaseous substrates, CO_2 and O_2, compete (Gutteridge *et al.* 1984*a*; Saver & Knowles 1982; Sue & Knowles 1982). A detailed understanding of how the geometry of the active site and the resident amino acids dictate the partitioning of the bisphosphate, presumably by discriminating between CO_2 and O_2, is unknown. However, a combination of structural information from X-ray crystallography of different species of the carboxylase and recent attempts to assign regions of the enzyme structure to specific functions by using mutagenesis *in vitro* (Gutteridge *et al.* 1984*b*; Estelle *et al.* 1985; Terzaghi *et al.* 1986) may begin to explain the basis of the substrate discrimination.

The first species of Rubisco to be succesfully cloned and expressed as an active enzyme, with identical kinetic characteristics to the authentic protein, was from *Rhodospirillum rubrum* (Somerville & Somerville 1984; Nargang *et al.* 1984). This enzyme is structurally the simplest Rubisco, composed only of large (L) subunits and active as the dimer. Unfortunately, it does not resemble the enzyme from higher plants which combines both large and small (S) subunits in a complicated hexadecameric (L_8S_8) arrangement. Most recently, plasmid constructs containing the genes for the L and S subunits of Rubisco from cyanobacteria have proved to be acceptable expression vectors, producing relatively large quantities of an active L_8S_8 enzyme that is similar to the native protein (Tabita & Small 1985). These constructs are now being

[129]

31-2

exploited to generate enough higher-plant-type Rubisco for detailed physico-chemical studies, and are also being used for mutagenesis *in vitro* by techniques similar to those already applied to the *R. rubrum* carboxylase. In this way we hope to determine the relative importance of regions of the primary structure on the reaction mechanism and assembly of this complicated enzyme.

Recombinant forms of Rubisco

Form I carboxylase is the enzyme that is composed of the large (55 kDa) and small (14.5 kDa) subunits arranged in a hexadecameric structure, L_8S_8, and is found throughout photosynthetic organisms, from bacteria to higher plants. The form II carboxylase has, so far, been located only in prokaryotic organisms and is composed of only large subunits. These subunits can be arranged either as dimers (L_2), as in the enzyme from *Rhodospirillum rubrum*, or as hexamers (L_6), as found in *Rhodospeudomonas sphaeroides*. The genes for large subunits from all three quaternary forms have been cloned into various plasmids either for sequencing purposes or for structural manipulation. Table 1 lists those constructs that produce active carboxylase.

Many sequences of the L and S subunits have now become available as a result of recombinant techniques, and these have produced a measure of the degree of homology of the enzyme throughout photosynthetic organisms. More particularly, this comparison has highlighted the composition of the active site through the low homology between form I and II enzymes. The regions that are conserved have, in nearly all cases, been pinpointed by affinity

TABLE 1. PLASMID CONSTRUCTS CONTAINING THE STRUCTURAL GENES OF RUBISCO THAT PRODUCE ACTIVE ENZYME WHEN EXPRESSED IN *E. COLI*

construct	promoter/ vector	gene source	activity	reference
		form II		
pRR116	pBR325	*Rhodospirillum rubrum*	trace	Somerville & Somerville 1984
pRR2119	pBR322	*Rhodospirillum rubrum*	+	Somerville & Somerville 1984
pRRE188	pBR322	*Rhodospirillum rubrum*	+	Gutteridge *et al.* 1984*b*
pRQ2	pBR322	*Rhodopseudomonas sphaeroides*	trace	Quivey & Tabita 1985
pRQ52	pUC8	*Rhodopseudomonas sphaeroides*	+	Quivey & Tabita 1985
		form I		
pSV55	pLa2311	*Synechococcus* PCC6301	trace	Gatenby *et al.* 1985
pDB50	pUC9	*Synechococcus* PCC6301	+	Gatenby *et al.* 1985
pCS75	pUC9	*Synechococcus* PCC6301	+	Tabita & Small 1985
pSynRES1	pUC9	*Synechococcus* PCC6301	+	This work
pSynRES2	pUC9	*Synechococcus* PCC6301	trace	This work
pANtac	pKK223-3	*Synechococcus* PCC6301	+	This work
pAnX105	pUC19	*Anabaena* 7120	+	Gurevitz *et al.* 1985

labels. These affinity labels have identified individual amino acid residues at the active site, and the sequences have indicated the regions that must compose the site.

For one to be able to attempt genetic manipulation meaningfully, the recombinant protein must be active. The first constructs involving the structural genes for the L subunit of higher plants, although relatively good expression vectors, produced proteins which were not only devoid of activity but did not assemble correctly and were isolated as aggregated insoluble masses (Gatenby *et al.* 1981; Gatenby & Castleton 1982). The prokaryotic form II carboxylase, composed of only one subunit type, was considered more likely to be expressed in an active state. The gene coding for the carboxylase from *Rhodospirillum rubrum* was obtained from a gene library as a 4.0 kilobase fragment of DNA, and cloned into pBR325 (Somerville & Somerville 1984). The resulting construct, pRR116, produced active carboxylase but with very low yields. A 2.4 kb BamH1 fragment was taken from pRR116 and cloned into the Tet^r region of pBR322, along with some of the *lacZ* gene from M13 mp 7. The resulting construct produced significant quantities of active *R. rubrum* enzyme when expressed in *E. coli* hosts. The only significant difference between the recombinant material and authentic enzyme was the extension by 25 amino acid residues of the *N*-terminus of the L subunit; kinetically, the enzymes were indistinguishable. The larger size of the polypeptide was due to the fusion of the *R. rubrum* gene in frame with part of *lacZ* encoding the *N*-terminal region of β-galactosidase. The construction of this plasmid, pRR2119, finally laid to rest any doubts that the carboxylation and oxygenation of RuBP was catalysed by two different enzymes, as the recombinant product exhibited both activities. This plasmid has proved useful for applying site-specific mutagenesis to modify the active site structure of the carboxylase.

SITE-DIRECTED MUTAGENESIS *IN VITRO*

The construction of a recombinant system that generates acceptable quantities of active protein allows mutagenesis of the carboxylase genes by a number of recently well-developed methods (Smith 1986; Botstein & Shortle 1985). Most notable is the replacement of any amino acid residue desired within the primary sequence. In the case of enzymes, these methods have been exploited to determine the role of individual amino acid residues in the catalytic process (Wilkinson *et al.* 1984), to modify the interaction of subunits in complex structures (Robey & Schachman 1985), or to introduce stability into the protein structure (Wells *et al.* 1985).

An invaluable aid to many of these manipulations has been the availability of the structure of the active centres of the proteins from X-ray data. The choice of which amino acid residue to change, and to what, can be made rationally, especially if a specific function is suspected for that amino acid residue. For carboxylase, without detailed X-ray data, models of the active site have to be based on sequence comparisons and affinity-label results. Figure 1 is just such a model, which combines this information with a prediction of the secondary structure of the important regions. Some of these regions are considered more distant from the site because they are either conserved but not labelled, eg Lys 228, or have been labelled with a 'long-distance' probe (Herndon & Hartman 1984). The sequence around Lys 8 identified by limited proteolysis (Gutteridge *et al.* 1986) may not contribute any one important residue; rather loss of this sequence destroys the enzyme activity. It must also be emphasized that the S subunit also contributes to the integrity of the active site, although no specific amino acid residue or region

has been labelled. Although various amino acid residues have been identified by these probes at the active site, their function is still unclear. Only one amino acid residue has a known function, and that is Lys 201 (191 in *R. rubrum*), which is carbamylated during activation of the carboxylase by CO_2 and Mg^{2+}.
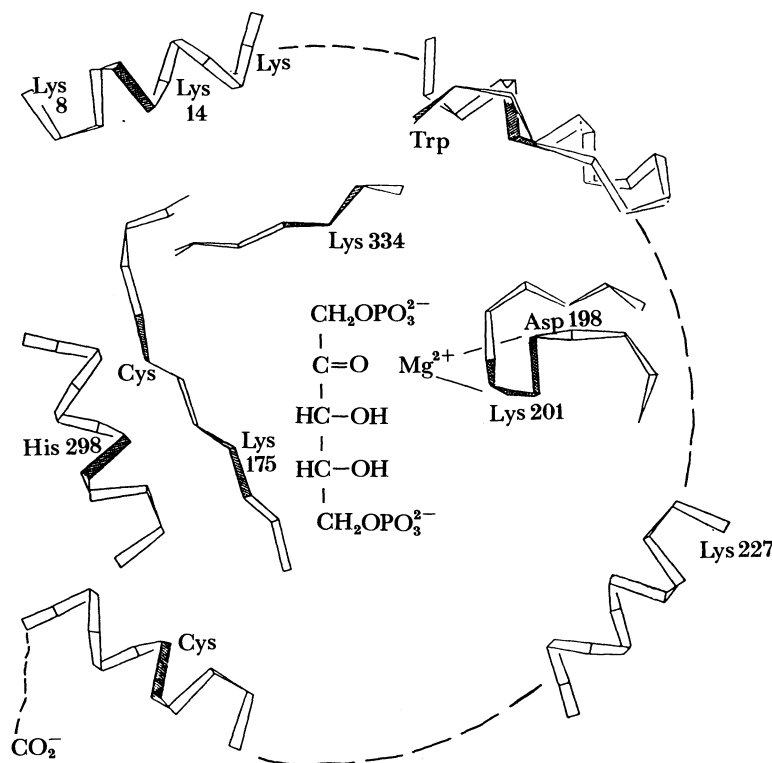

FIGURE 1. A model of the active site of Rubisco showing those regions of the L-subunit primary sequence that are highly conserved, or have been shown to be involved in the integrity of the active site. Those placed on the outer radial are part of the sequence picked out by 'long-range' affinity probes (Ala 53–Thr 65), discovered by limited proteolysis (Met 1–Lys 14) or not labelled but highly conserved (Lys 227–Lys 236). Secondary structure prediction suggests these tend to be α-helical. Those sequences surrounding the labelled Lys 175 and Lys 334 are sheet structures and form the active site; the essential Lys 201 apparently resides in a loop region along with Asp 198. Both amino acid residues form the coordination sphere of the activating metal ion Mg^{2+}, which orients the RuBP substrate and directs the stereochemical course of the reaction of catalysis. (Note that residues are enumerated with respect to the sequence of the spinach enzyme.)

Without the help of a three-dimensional model of the active site of carboxylase, then, site-directed mutagenesis must be applied to those regions of the enzyme structure that are known to compose the catalytic site, preferably involving amino acid residues considered important at that site i.e. those identified within the immediate vicinity.

The first application of these techniques was to change the sequence surrounding the lysine residue that is carbamylated during enzyme activation. The production of the carbamate initiates the formation of the binding site of the essential Mg^{2+} ion. Many of the other amino acids adjacent to this lysine residue are oxygen-containing species that are likely to be part of that co-ordination sphere. One of these amino acids, Asp 188, is highly conserved, and might be just such a candidate. By changing the codon GAC, which codes for Asp 188, to GAA, the aspartate residue could be substituted by a glutamate residue. It was argued that such a modest

change should allow the retention of enzyme activity, as it would not cause marked disruption of the site, and yet should modify the metal-ion site. A further advantage of this change was the introduction of a new restriction site, which was recognized by *Eco*RI and was thus a ready means of identifying mutated plasmids.

The single point mutation was achieved by applying fairly standard techniques to pRR2119, except that the linear form of the duplex was used for the manipulations (Gutteridge *et al.* 1984*b*). A synthetic 20-base oligomer was designed to hybridize at only one region of the coding strand exposed around the site of interest and align the single-base mismatch to convert the C in GAC to an A. The double-stranded plasmid was recovered by polymerase and ligase action and the 20-mer, this time radioactively labelled with ^{32}P, was used to probe transformed colonies for the presence of mutated DNA. Those colonies harbouring the mutant plasmid, pRRE188, were grown to high density, and the isolated plasmid was subjected to restriction digestion to locate the new *Eco*RI site. From the size of the DNA fragments in combined digests with other restriction enzymes, it was clear that approximately 1–3 % of transformants carried the mutation at about the correct position in the DNA. Sequencing of a restriction fragment that encompassed the new site indeed confirmed the successful positioning of the mutation.

Although the incorporation of a mutation into the active site had been shown to be successful at the level of the DNA, it is essential that enough of the recombinant material be isolated to allow the effect of the mutation on the protein function to be assessed in detail. Attempts to obtain the *R. rubrum* enzyme from *E. coli* in continuous culture were unsuccessful, producing inconsistent and inferior amounts of the protein. Analysis of the organism during increasing numbers of generations indicated that the plasmid was unstable and that cultures were becoming dominated by non-plasmid-bearing cells (Pierce & Gutteridge 1985). From assessment of the activity of the β-lactamase that was being generated along with the carboxylase from the plasmid, it was clear that the selection pressure imparted by the ampicillin in the medium was being rapidly depleted. Non-plasmid-bearing organisms that had a superior doubling time were rapidly overgrowing the plasmid-containing cells, resulting in a massive loss of recombinant protein. A model was devised based on the expected rate of loss of the antiobiotic against the inoculum size, i.e. the initial concentration of β-lactamase and the expected time for batch cultures to be dominated by the non-plasmid organisms. By inoculating the culture with a very small amount of cells (*ca.* 1 cell per millilitre of medium), the number of generations obtained with the desired organism, before the antibiotic was depleted below concentrations sufficient to sustain selection, was high enough to obtain reasonable quantities of protein-containing cells. In fact, from a 300 litre fermenter containing glycerol-supplemented L-broth with O_2 gassing, an OD_{650} of about 26 was obtained with *E. coli* (HB101), and each batch of cells yielded approximately 85 g of the recombinant enzyme. In contrast, with *E. coli* (HB101) containing pRRE188 the quantities of enzyme obtained were consistently 10–20 % lower than the wild-type enzyme. Although the mutation altered a GAC codon to GAA, this codon is used elsewhere in the gene for glutamate and therefore was not the cause of the lower levels of expression. Presumably the effect of the single point-mutation is to change the secondary structure of the plasmid or message.

The wild-type and mutant enzymes were compared for changes in kinetic parameters and for the effect of this mutation on the metal-binding site. The specificity factor was unchanged, but the k_{cat} for both oxygenase and carboxylase had declined by 30 %. There was a marginal but significant increase in the K_m for CO_2. The most dramatic difference between the two

enzyme forms was the nature of the metal site due to the mutation. The Mg^{2+} was substituted by Mn^{2+} during the activation of the enzymes, trapped at the active site by carboxyarabinitol bisphosphate (CABP); the spectra of the frozen samples was recorded. It was clear from the change in position of some of the features, and the increase in hyperfine resolution, that the replacement of aspartate by glutamate perturbed the metal–ion site, and yet the overall conservation of the spectrum line shape indicated that the mode of CABP binding had been unaffected. Thus the introduction of an extra methylene group within the co-ordination sphere of the metal has perturbed the site enough to decrease the turnover of both catalytic activities.

pRR2119 has been subjected to other site-directed mutations. Somerville and colleagues have replaced the lysine residue at 191 (201 in spinach) that is carbamylated during activation, by a glutamate residue (Estelle *et al.* 1985). The view was that the γ-carboxyl group of a glutamate residue might be the equivalent of a stable carbamylated ε-amino group of lysine. Unfortunately, the enzyme was catalytically inactive, although the bisphosphate-binding site was still intact. A second example has been the conversion of Met 330 to Leu in the *R. rubrum* protein (Terzaghi *et al.* 1986); the leucine residue, together with Lys 329, is conserved at the active site of form I carboxylase L subunits. This mutation did not make the *R. rubrum* enzyme more like a higher plant enzyme but, as outlined in table 2, decreased the affinity for CO_2 significantly.

TABLE 2. SITE-DIRECTED MUTATIONS IN pRR2119

amino acid residue	location	effect	reference
Asp → Glu	188	τ unchanged $k_{cat} \times 0.7$ $K_m(CO_2) \times 2$ metal site disturbed	Gutteridge *et al.* 1984*b*
Lys → Glu	191	inactive	Estelle *et al.* 1985
Met → Leu	330	$K_m(CO_2) \times 15$ $k_{cat} \times 0.2$ $K_m(O_2) \times 36$	Terzaghi *et al.* 1986

EXPRESSION AND MUTAGENESIS OF THE FORM I (L_8S_8) ENZYME IN BACTERIA

The first plasmids containing the structural genes of the form I enzyme were constructed mainly for sequencing purposes and thus contained only L or S structural genes rather than both. One of these, pZMB1B, was used to supply the gene for the L subunit of maize which was introduced into a vector containing the thermoinducible promoter (P_L) from λ phage. The resulting construct, pPBI3 expressed the L subunit to relatively high levels, but the protein could be isolated only as an insoluble aggregate (Gatenby & Castleton 1982). No soluble L subunit was detected, nor was there any RuBP-dependent CO_2-fixing ability. The same problems were encountered with the wheat L subunit gene in a similar construct (Gatenby *et al.* 1981). Clearly, the question arises whether the S subunit is essential to ensure the solubility of the L subunit.

An important step in understanding the protein-synthetic processes that lead to the assembly of such a complicated quaternary structure has been the construction of plasmids that produce active form I enzyme. The structural genes for the L subunit were first isolated from cyanobacteria, and were again used for sequencing purposes before it was realized that the DNA fragments that had been isolated also contained the S subunit gene (Shinozaki *et al.* 1983; Shinozaki & Sugiura 1983). In the case of the carboxylase from *Synechococcus* PCC6301, the

organization of the gene is shown in figure 2. The L subunit gene precedes the S subunit by some 93 bases, and both are transcribed as one large polycistronic message (Shinozaki & Sugiura 1985). There are obvious regions of secondary structure within the intergenic region and the 3' region of the S subunit. A ribosome-binding site is also evident 5' to the two genes.

Two plasmids have been constructed by using the *Pst*1 fragment from *Synechococcus* that has both genes. The first, pSV55, put the control of expression under a P_L promoter, but with only limited success. The construct produced active enzyme, but it was of low specific activity and at low levels (Gatenby *et al.* 1985). We have constructed pSynRES1 with the genes under the control of the *lac* promoter in pUC9 (figure 2). In the *E. coli* host JM83, the amount of active carboxylase that can be detected suggests that almost 15 % of the soluble cell contents are due to the recombinant enzyme. With the genes in the opposite orientation in pUC9 (pSynRES2) there are only low levels of enzyme activity, indicating that the *lac* promoter controls expression. Two other groups have found similar results with the same construct (Tabita & Small 1985; Christeller *et al.* 1985).
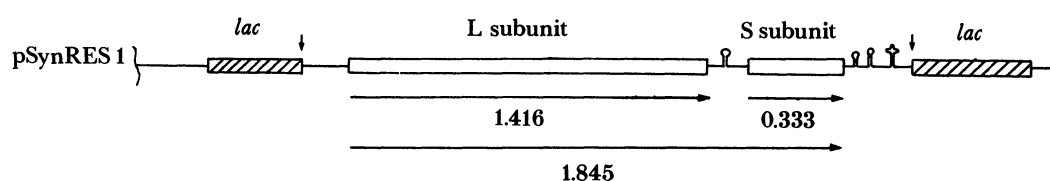

FIGURE 2. The organization of the structural genes encoding the L and S-subunits of Rubisco from *Synechoccocus* PCC6301. The *Pst*I fragment has been shown in the correct orientation in pSynRES1 to be controlled from the *lac* promoter of pUC9. The secondary structure identified within the intergenic region and the 3'-flanking region of the S subunit are shown and might act as termination loops. Both subunits are translated from one large polycistronic message (Shinozaki & Sugiura 1985).

We have also constructed pANtac containing the *Synechococcus* genes under the control of the *tac* promoter in pKK223-3; this also produces active carboxylase.

Batches of *E. coli* JM83, transformed with pSynRES1, were grown in a supplemented L-broth (LB) in a 5-litre fermenter. The cells were sonicated or treated with lysozyme to release the carboxylase into an activation buffer. After purification by ammonium sulphate precipitation, sucrose density centrifugation and ion-exchange chromatography, the enzyme was considered to be approximately 50 % pure, based on a carboxylase specific activity of 2.2 μmol CO_2 per minute per milligram of protein. SDS-gradient gel electrophoresis indicated that one other protein, slightly larger in size than the L subunit, was the major contaminant in these solutions. Further ion-exchange chromatography on a MonoQ FPLC column (figure 3) resolved the enzyme from this other contaminant and resulted in a carboxylase preparation with a specific activity close to the predicted maximum of *ca.* 3.6 units per milligram of protein (Andrews & Abel 1981). A second peak of carboxylase activity was eluted from the column in a region that contained the major contaminant.

SDS-gel electrophoresis indicated that the component of the second peak of activity is a protein of $M_r \sim 63000$ compared to the 59 kDa protein of the normal L subunit (see figure 4). Scrutiny of the plasmid containing the Rubisco genes indicates a possible source for this protein (see figure 5). The *Pst*I fragment ligated into the *Pst*I site of pUC9 would result in a fusion 'in-frame' with the β-galactosidase gene, which is the normal product of *lac*-directed transcription. The amino acids inherited from β-galactosidase are shown in figure 5, along with
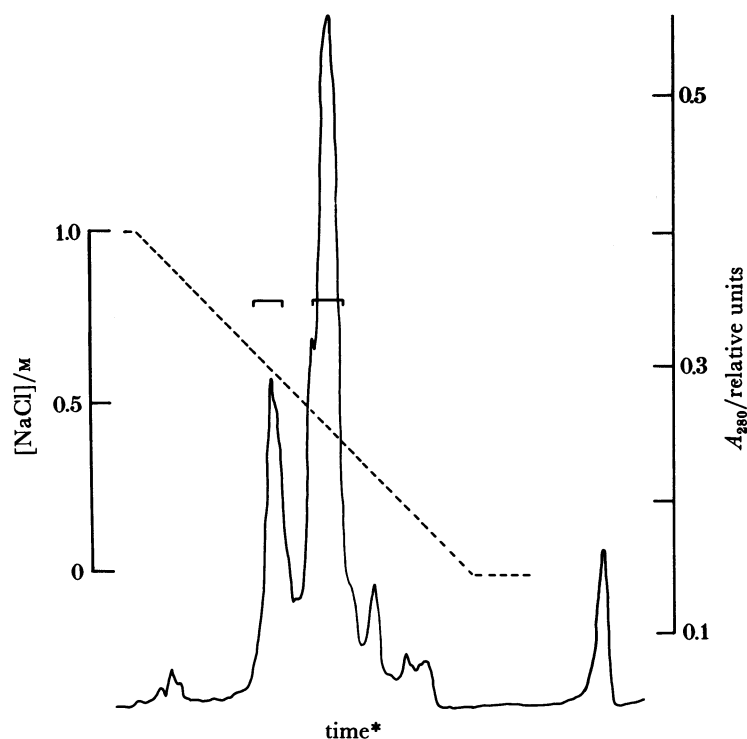
440 S. GUTTERIDGE AND OTHERS

FIGURE 3. The separation of the two major proteins in the *Synechococcus* preparation that are isolated after standard purification procedures. The procedure involves FPLC with a Mono Q cation column, and developing with a linear NaCl gradient from 0 to 1.0 M at pH 7.6. Two peaks exhibiting carboxylase activity were isolated. (*Time on arbitrary scale.)
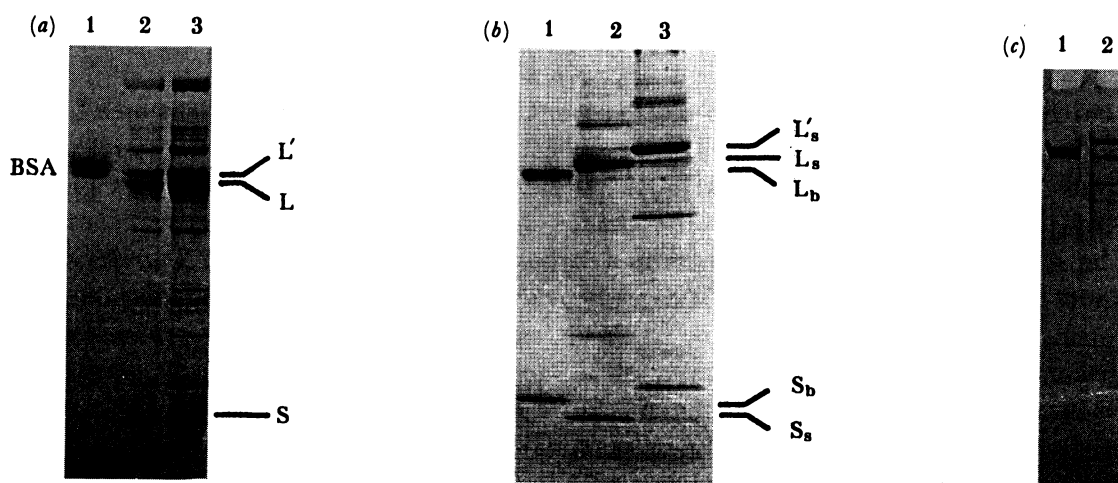

FIGURE 4. Polyacrylamide gel electrophoresis of the two protein peaks resolved by FPLC. (*a*), SDS–polyacrylamide gradient gels (10–20%), containing the samples of partially purified *E. coli* cell extract, are compared with the bovine serum albumin (BSA) standard (track 1). Two major protein bands with approximately the mobility of the Rubisco L subunit (L and L′) are clearly evident. (*b*) A similar gel of the two proteins resolved by FPLC methods. In track 1 is pure barley Rubisco L_bS_b, track 2 was the major peak from the ion-exchange column with the highest carboxylase activity (L_s) and track 3 is the peak of minor activity (L'_s) but major contaminant. The gel shown in (*c*) is a non-denaturing gradient gel (4–20%) of the two proteins. Track 1 is the main *Synechococcus* Rubisco, track 2 the second peak of activity. S_s, small subunit of *Synechococcus*; S_b, small subunit of barley.

[136]

an extended *N*-terminus of the L subunit, which is due to the translation of the 5′ flanking region to the normal ATG start of the L subunit. The increase in size of the L subunit as a result of the fusion is about 3 kDa, due to 10 amino acid residues from the *N*-terminus of β-galactosidase and 24 from the translated 5′ region of the L-subunit gene. The amount of the fusion peptide of carboxylase expressed by the plasmid is about 50% of the normal L subunit. The activity of this form of the enzyme is significantly lower than that of the normal enzyme because, as revealed by the gel electrophoresis, there is little S subunit associated with this fusion protein and in non-denaturing gels different aggregated forms can be resolved. Clearly, the extension of the L subunit by 34 amino acids interferes with the assembly of the protein, which involves both L or L and S subunits coming together correctly. The values for the molecular weights are shown in table 3 for the holoenzyme and the extended species in both denaturing and non-denaturing conditions.
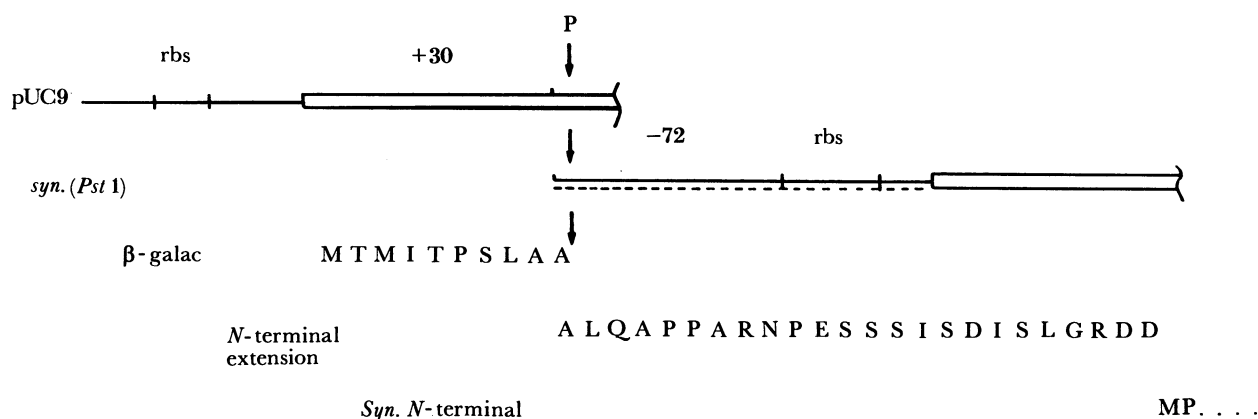

FIGURE 5. The source of the second form of the *Synechococcus* L subunit L′. The ligation of the *Pst*1 fragment in pSynRES1 is in frame with the *lacZ* gene product, β-galactosidase, resulting in a protein that is about 3 kDa larger due to ten amino acid residues of β-galactosidase and 24 of the 5′ flanking region of the *Synechococcus* gene; rbs indicates ribosome binding site. (Single-letter abbreviations for amino acid residues as follows: A, Ala; D, Asp; G, Gly; I, Ile; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr.)

TABLE 3. THE RELATIVE SIZE OF THE INDIVIDUAL SUBUNITS OF CARBOXYLASE AND THE HOLOENZYME, DETERMINED BY GEL ELECTROPHORESIS

subunit*	$10^{-3}M_r$	
	SDS	non-SDS (holoenzyme)
L	59	630
L′	63	260
		437
		620
S	12	

* L′ is the L subunit extended by some 34 amino acids at the *N*-terminus. Estimations of M_r were made by reference to standard proteins of known M_r electrophoresed under the same conditions.

Manipulation of the structure of the form I enzyme

The construction of an acceptable expression vector containing the structural genes of the form I enzyme will allow mutagenesis *in vitro* to be used to change the primary sequence, as previously applied to the form II enzyme. In the case of the form I carboxylase there are two subunits that could be manipulated. It is known from reversible dissociation studies that the S subunit is essential for the catalytic activity of the enzyme (Andrews & Ballment 1984). However, it has also recently been shown that the S subunit is unlikely to affect the kinetic constants that dictate the specificity factor or partition coefficient of the carboxylase and oxygenase reactions (Andrews & Lorimer 1985). Unfortunately, no detailed kinetic analysis has been achieved with recombinant versions of a hybrid enzyme, composed of a mixture of one species of L and a different species of S subunit, to confirm these conclusions. The constructs containing the *Synechoccocus* genes might be manipulated in this fashion.

The particular mutations that we are constructing have involved only the L subunit with a view to exchanging different regions of the primary sequence of the *Synechococcus* L subunit
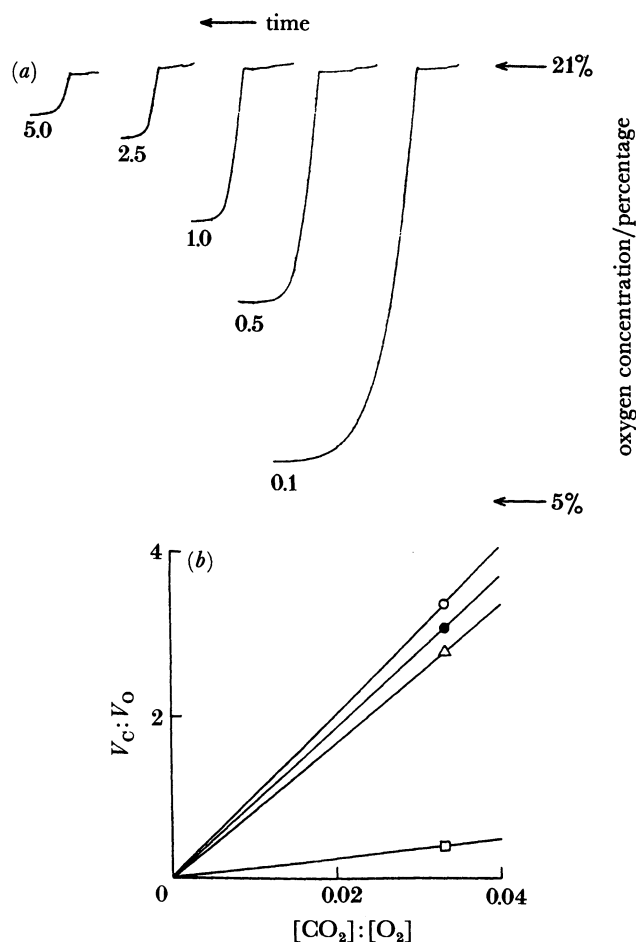

Figure 6. The total consumption of RuBP by simultaneous oxygenation and carboxylation. (*a*) The trace is a record of the oxygen consumed at different concentrations of HCO_3^- (mM). These curves provide a measure of the carboxylase:oxygenase ratio ($V_C:V_O$) which can be plotted against the $CO_2:O_2$ ratio to obtain the specificity factor as shown in (*b*). Symbols: ○, wheat; ●, maize; △, tobacco; □, *Rhodospirillum rubrum*.

with sequences from a higher-plant enzyme. In this way we hope to identify selected regions of the L sequences that dictate the kinetic parameters.

It has become clear from this approach that large numbers of mutant forms of the enzyme will be generated, which require screening for any change in specificity factor. We have developed a relatively rapid method for determining the specificity factor from a single oxygen-electrode trace of the consumption of RuBP by the enzyme. Figure 6*a* shows the form of the trace that is obtained when the enzyme is allowed to oxygenate the bisphosphate in the presence of various concentrations of CO_2 (as HCO_3^-). The majority of the substrate is carboxylated rather than oxygenated and thus the electrode trace is a record of how much bisphosphate is consumed by oxygenation. The RuBP substrate is quantitated accurately by driving it all through a carboxylation reaction and then, by difference, calculating a carboxylase:oxygenase ratio for that particular concentration of CO_2. Division of the activity ratio by the substrate ratio $CO_2:O_2$, i.e. the slope of the line in figure 6*b*, provides the specificity factor τ. Confirmation of the result can be obtained by repeating the assay at different CO_2 concentrations to give further points that coincide with the line in the figure. The generation of an oxygen-electrode trace for the total consumption of the bisphosphate with different species of carboxylase allows a direct visual comparison of the specificity factors for these species. Some values of specificity factors of different higher plant carboxylases, compared with the wheat enzyme and the recombinant forms of the *Synechococcus* and *Rhodospirillum rubrum* enzymes, are shown in table 4.

TABLE 4. SPECIFICITY FACTORS MEASURED FROM THE TOTAL CONSUMPTION OF RuBP

source	structure	specificty factor τ
higher plants:		
wheat	L_8S_8	100
tobacco	L_8S_8	90
bacteria:		
Synechococcus	L_8S_8	52
Rhodospirillum rubrum	L_2	11

CONCLUDING REMARKS

Without the help of a detailed structural model of the active site of Rubisco, the approach to engineering the enzyme by site-directed mutagenesis has been limited. Although we know more about the function of specific amino acid residues at the active site, it is not clear what contribution they make to the kinetic parameters of the carboxylase, and thus what changes might produce a 'better' enzyme. One particular aspect that this site-directed approach has highlighted is that multiple mutation might be necessary so that a specific change in one region of structure is correctly compensated for in another. In the absence of a detailed map, this would be feasible only by non-specific mutagenesis, e.g. using chemical mutagens, or alternatively, domain-replacement with the DNA sequences from different L subunits, as outlined for the *Synechococcus* construct above. Both approaches would benefit from the development of a rapid screen for the effect of the mutations of carboxylase at the level of the colony. The recent successes of transforming photosynthetic bacteria with shuttle vectors may provide the appropriate host–vector system for such a screen.

The construction of two recombinant systems that produce acceptable amounts of both form I and II Rubisco with quite distinct specificity factors means that enough protein is available, either of the wild type or mutated enzyme species, for crystallographic studies. Such information for the inferior carboxylases of bacterial sources, compared with the superior higher-plant-type Rubisco, will allow us, by extrapolation, to define what is required through structural manipulation to produce a more efficient carboxylase.

REFERENCES

Andrews, T. J. & Abel, K. M. 1981 Kinetics and subunit interactions of ribulose bisphosphate carboxylase/oxygenase from the cyanobacterium *Synechococcus* sp. *J. biol. Chem.* **256**, 8445–8451.

Andrews, T. J. & Ballment, B. 1984 Active site carbamate formation and reaction intermediate analogue binding by ribulose bisphosphate carboxylase/oxygenase in the absence of its small subunits. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3660–3664.

Andrews, T. J. & Lorimer, G. H. 1985 Catalytic properties of a hybrid between cyanobacterial large subunits and higher plant small subunits of ribulose bisphosphate carboxylase/oxygenase. *J. biol. Chem.* **260**, 4632–4636.

Botstein, D. & Shortle, D. 1985 Strategies and applications of *in vitro* mutagenesis. *Science, Wash.* **229**, 1193–1201.

Christeller, J. T., Terzaghi, 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.

Estelle, M., Hanks, J., McIntosh, L. & Somerville, C. R. 1985 Site-specific mutagenesis of ribulose-1,5-bisphosphate carboxylase/oxygenase. *J. biol. Chem.* **260**, 9523–9526.

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

Gatenby, A. A., Van der 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.

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

Gutteridge, S., Millard, B. N. & Parry, M. A. J. 1986 Inactivation of ribulose-P_2 carboxylase by limited proteolysis: loss of catalytic activity without disruption of bisphosphate binding or carbamylation. *FEBS Lett.* **196**, 263–268.

Gutteridge, S., Parry, M. A. Schmidt, C. N. G. & Feeney, J. 1984a An investigation of ribulose bisphosphate carboxylase activity by high resolution ^{1}H NMR. *FEBS Lett.* **170**, 355–359.

Gutteridge, S., Sigal, I., Thomas, B., Arentzen, R., Cordova, A. & Lorimer, G. H. 1984b A site-specific mutation within the active site of ribulose-1,5-bisphosphate carboxylase of *Rhodospirillum rubrum*. *EMBO J.* **3**, 2737–2743.

Herndon, C. S. & Hartman, F. C. 1984 2-(4-bromoacetamide)anilino-2-deoxypentitol 1,5-bisphosphate, a new affinity label for ribulose bisphosphate carboxylase/oxygenase from *Rhodospirillum rubrum*. *J. biol. Chem.* **259**, 3102–3110.

Jordan, D. B. & Ogren. W. L. 1981 Species variation in the specificity of ribulose bisphosphate carboxylase/oxygenase. *Nature, Lond.* **291**, 513–515.

Jordan, D. B. & Ogren, W. L. 1983 Species variation in kinetic properties of ribulose bisphosphate carboxylase/oxygenase. *Archs Biochem. Biophys.* **227**, 425–433.

Nargang, F., McIntosh, L. & Somerville, C. R. 1984 Nucleotide sequences of the ribulose bisphosphate carboxylase gene from *Rhodospirillum rubrum*. *Molec. gen. Genet.* **193**, 220–224.

Pierce, J. & Gutteridge, S. 1985 Large scale preparation of ribulose bisphosphate carboxylase from a recombinant system in *E. coli* characterised by extreme plasmid instability. *Appl. Env. Microbiol.* **49**, 1094–1100.

Quivey, R. G. & Tabita, F. R. 1985 Cloning and expression in *E. coli* of the form II ribulose, 1,5-bisphosphate carboxylase/oxygenase from *Rhodopseudomonas sphaeroides*. *Gene* **31**, 91–101.

Robey, E. A. & Schachman, H. K. 1985 Regeneration of active enzyme by formation of hybrids from inactive derivatives: Implications for active sites shared between polypeptide chains of aspartate transcarbamylase. *Proc. natn. Acad. Sci. U.S.A.* **82**, 361–365.

Saver, B. C. & Knowles, J. R. 1982 Ribulose-1,5-bisphosphate carboxylase: Enzyme catalysed appearance of solvent tritium at carbon 3 of ribulose bisphosphate reisolated after partial reaction. *Biochemistry, Wash.* **21**, 5398–5403.

Shinozaki, K. & Sugiura, M. 1983 The gene for the small subunit of ribulose bisphosphate carboxylase/oxygenase is located close to the gene for the large subunits in the cyanobacterium *Anacystis nidulans* 6301. *Nucl. Acids Res.* **11**, 6957–6964.

Shinozaki, K. & Sugiura, M. 1985 Genes for the large and small subunits of ribulose 1,5-bisphosphate carboxylase/oxygenase constitute a single operon in a cyanobacterium *Anacystis nidulans* 6301. *Molec. gen. Genet.* **200**, 27–32.

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

Smith, M. 1986 *A. Rev. Genet.* (In the press.)

Somerville, C. R. & Somerville, S. C. 1984 Cloning and expression of the *Rhodospirillum rubrum* ribulose bisphosphate carboxylase gene in *E. coli*. *Molec. gen. Genet.* **193**, 214–219.

Sue, J. M. & Knowles, J. R. 1982 Ribulose-1,5-bisphosphate carboxylase: fate of the tritium label in [3 – ³H]ribulose 1,5-bisphosphate during the enzyme-catalysed reaction. *Biochemistry, Wash.* **21**, 5404–5409.

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

Terzaghi, B. E., Laing, W. A., Christeller, J. T., Peterson, G. & Hill, D. F. 1986 Ribulose-1,5-bisphosphate carboxylase: The effect on the catalytic properties of changing methionine-330 to leucine in *Rhodospirillum rubrum* enzyme. *Biochem. J.* **235**, 839–846.

Wells, J. A., Cunningham, B. C., Graycar, T. P. & Estell, D. A. 1986 Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin. *Phil. Trans. R. Soc. Lond.* A **317**, 415–423.

Wilkinson, R. J., Fersht, A. R., Blow, D. M., Carter, P. & Winter, G. 1984 A large increase in enzyme-substrate affinity by protein engineering. *Nature, Lond.* **307**, 187–188.
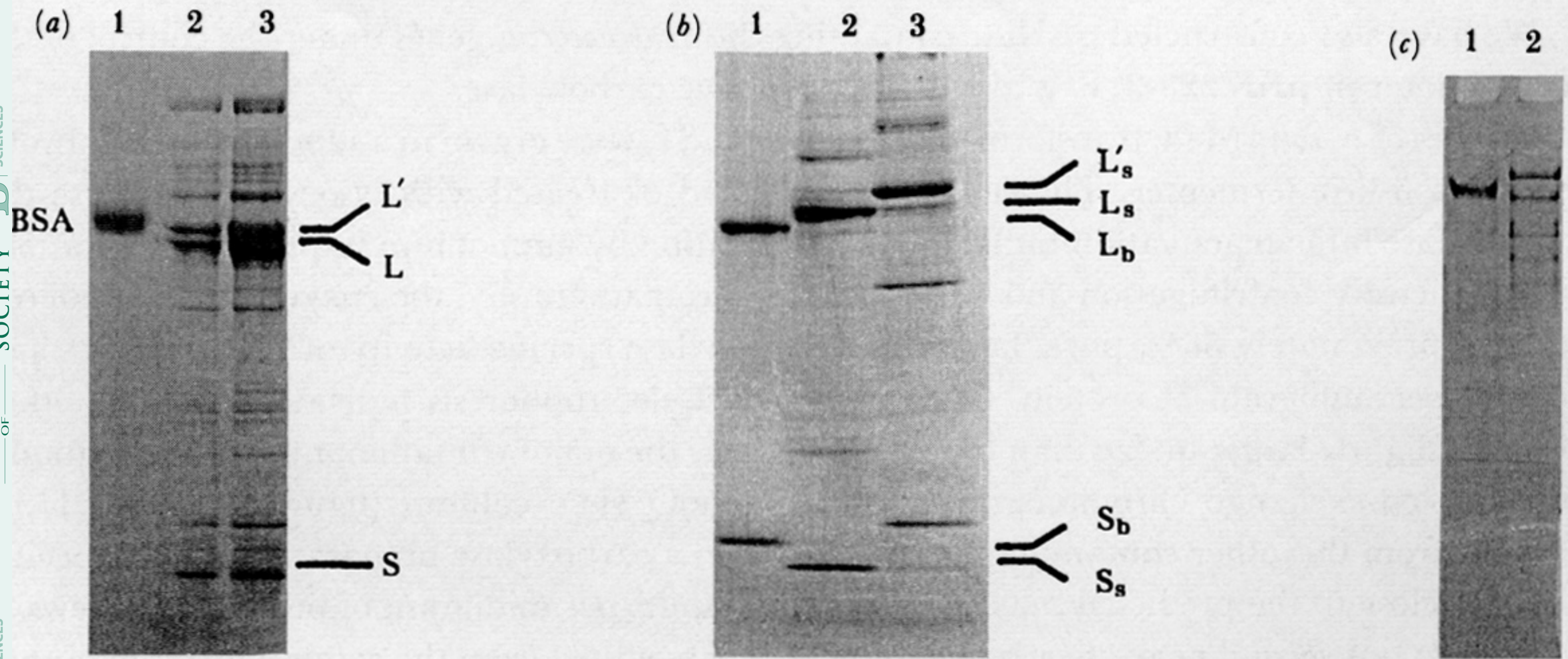
FIGURE 4. Polyacrylamide gel electrophoresis of the two protein peaks resolved by FPLC. (a), SDS–polyacrylamide gradient gels (10–20%), containing the samples of partially purified *E. coli* cell extract, are compared with the bovine serum albumin (BSA) standard (track 1). Two major protein bands with approximately the mobility of the Rubisco L subunit (L and L′) are clearly evident. (b) A similar gel of the two proteins resolved by FPLC methods. In track 1 is pure barley Rubisco L_bS_b, track 2 was the major peak from the ion-exchange column with the highest carboxylase activity (L_s) and track 3 is the peak of minor activity (L_s') but major contaminant. The gel shown in (c) is a non-denaturing gradient gel (4–20%) of the two proteins. Track 1 is the main *Synechococcus* Rubisco, track 2 the second peak of activity. S_s, small subunit of *Synechococcus*; S_b, small subunit of barley.