

## The role of the N-terminus of the large subunit of ribulose-bisphosphate carboxylase investigated by construction and expression of chimaeric genes

Catherine A. KETTLEBOROUGH, Martin A. J. PARRY, Steven BURTON, Steven GUTTERIDGE, Alfred J. KEYS and Andrew L. PHILLIPS

Department of Biochemistry, Rothamsted Experimental Station, Harpenden

(Received May 20/August 10, 1987) – EJB 87 0585

The genes for the large and small subunits of ribulose biphosphate carboxylase/oxygenase (Rubisco) from *Anacystis nidulans* have been expressed in *Escherichia coli* under the control of the *lac* promoter to produce active enzyme. The enzyme can be purified from the cells to yield up to 200 mg Rubisco/l cultured bacteria, and is indistinguishable from the enzyme extracted from *A. nidulans*. In order to investigate the role of the N-terminus of the large subunit in catalysis, chimaeric genes were constructed where the DNA coding for the 12 N-terminal amino acids in *A. nidulans* was replaced by DNA encoding the equivalent, but poorly conserved, region of either the wheat or maize large subunit. These genes, in constructs also containing the gene for the *A. nidulans* small subunit, were expressed in *E. coli* and produced enzymes with similar catalytic properties to the wild-type Rubisco of *A. nidulans*. In contrast, when the N-terminal region of the large subunit was replaced by unrelated amino acids encoded by the pUC8 polylinker, enzyme activity of the expressed protein was reduced by 90% under standard assay conditions, due to an approximately tenfold rise in the  $K_m$  for ribulose 1,5-bisphosphate. This confirms that the N-terminus of the large subunit has a function in catalysis, either directly in substrate binding or in maintaining the integrity of the active site.

Ribulose biphosphate carboxylase/oxygenase (Rubisco) catalyses the initial step in the fixation of atmospheric CO<sub>2</sub> by plants and photosynthetic bacteria. It also catalyses the oxygenation of the ribulose 1,5-bisphosphate (ribulose-*P*<sub>2</sub>) substrate in the first step of the photorespiratory cycle. This latter process is wasteful, since it removes ribulose-*P*<sub>2</sub> from the Calvin cycle, and also because recovery of the products of oxygenation result in the further loss of CO<sub>2</sub>. The relative rates of the two reactions therefore affect the efficiency of carbon assimilation. Consequently, there is considerable interest in attempts to alter the structure of Rubisco from crop plants by genetic manipulation in order to increase the proportion of carboxylation relative to oxygenation.

A prerequisite of rational genetic engineering is a three-dimensional map of the active site and knowledge of both the reaction mechanism and the roles of specific amino acid residues in the reaction. Unfortunately, little is known about the detailed mechanism of ribulose-*P*<sub>2</sub> carboxylation and oxygenation by Rubisco. The 2-carboxy-3-keto-arabinitol 1,5-bisphosphate intermediate of carboxylation has been identified [1], but a 2-peroxy-3-keto-intermediate of oxygenation is, as yet, only postulated [2]. Furthermore, only a single amino acid residue has been definitely identified as having a role in the reaction: a lysine at position 201 in the spinach large subunit (198 in *Anacystis nidulans*; 191 in *Rhodospirillum*

*rubrum*) which is carbamylated by CO<sub>2</sub> in the conversion of inactive to active enzyme [3]. Several other amino acid side chains have been implicated in the reaction by their labelling with active-site affinity probes [4], but their exact role is uncertain.

The N-terminus of the large subunit of Rubisco has also been shown to be important in maintaining catalytic activity. When wheat Rubisco was subjected to mild proteolysis, a 90% loss of activity was detected which correlated with the loss of amino acid residues 9–14 from the large subunit [5]; the small subunit was unaffected by this treatment. Rubisco from spinach, maize [5] and *A. nidulans* (M. A. J. Parry and A. L. Phillips, unpublished) similarly show a loss of activity on protease treatment. The N-terminal region of the large subunit is therefore important in catalysis by the L<sub>8</sub>S<sub>8</sub> enzyme, yet there is little homology between wheat and *A. nidulans* Rubisco in this region.

The ultimate aim of our work is to improve the efficiency of photosynthesis in crop plants by genetic engineering of Rubisco. The ideal model system for investigation of Rubisco enzymology by mutagenesis would therefore involve the genes for a crop plant Rubisco expressed in bacteria to yield assembled and active enzyme. However, present evidence indicates that the large subunit of wheat Rubisco fails to assemble correctly when synthesised in *Escherichia coli* [6], possibly due to the absence of a tentatively identified assembly protein [7, 8]. Several groups have therefore chosen Rubisco from cyanobacteria as a model system, since the large and small subunits can be separated and reassembled into an active complex [9, 10]. It has now been shown that genes for large and small subunits of Rubisco from several cyanobacteria can

Correspondence to A. L. Phillips, Department of Biochemistry, Rothamsted Experimental Station, Harpenden, England AL5 2JQ

Abbreviations. Rubisco, ribulose biphosphate carboxylase/oxygenase; ribulose-*P*<sub>2</sub>, ribulose 1,5-bisphosphate.

Enzyme. Ribulose biphosphate carboxylase/oxygenase (EC 4.1.1.39).

be successfully expressed in *E. coli* using a range of plasmid vectors and promoters [11–16].

In this paper we describe experiments designed to investigate the role of the N-terminus of the large subunit in catalysis by Rubisco. Plasmids were constructed consisting of the genes for the large and small subunits of Rubisco from *Anacystis nidulans* fused behind the *lac* promoter in pUC vectors. The DNA encoding the N-terminus of the large subunit was replaced with the equivalent DNA from maize and wheat large subunit genes, and also with DNA encoding an unrelated N-terminus. The chimaeric genes were expressed in *E. coli* and the enzyme produced purified and analysed with respect to their assembly and catalytic activity.

## MATERIALS AND METHODS

### Materials

*E. coli* strain NF1 carrying plasmid pSV55, consisting of a 2.2-kb *Pst*I fragment containing the genes for the large and small subunits of *Anacystis nidulans* (= *Synechococcus* PCC6301) Rubisco [17] inserted in pLa2311 [11], were obtained from A. A. Gatenby. *E. coli* strain C600 carrying plasmid pZMB1B, consisting of a 4.3-kb *Bam*HI fragment of *Zea mays* chloroplast DNA containing the gene for the large subunit of Rubisco [18], inserted in pBR322 [19] was also a gift from A. A. Gatenby. Plasmid pTac39, consisting of a 9.7-kb *Bam*HI fragment of wheat chloroplast DNA containing the wheat chloroplast gene for the large subunit of Rubisco, inserted in pBR322 [20], was obtained from T. A. Dyer. *E. coli* strain JM83 (*ara*, *Δlac-pro*, *strA*, *thi*, *φ89dlacZΔM15*) was used for propagation of all pUC-type plasmids.

Restriction and DNA modification enzymes were obtained from Bethesda Research Laboratories, New England Biolabs or Pharmacia;  $\text{NaH}^{14}\text{CO}_3$  and  $[\alpha\text{-}^{32}\text{P}]\text{dATP}$  were from Amersham International; deoxy- and dideoxynucleotide triphosphates and 5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactoside were from Boehringer, Mannheim; ribulose 1,5-bisphosphate was from Sigma.

### DNA manipulations

Most routine DNA manipulations were carried out essentially as described by Maniatis et al. [21]. Sequencing of double-stranded plasmid DNA was carried out by the dideoxynucleotide chain-termination method [22] from the universal or reverse-sequencing primers (New England Biolabs) by a combination of the methods of Chen and Seeburg [23] and of Hattori and Sakaki [24].

### Rubisco purification

Bacteria expressing wild-type and manipulated genes for Rubisco were grown to late-log phase in L-broth (1% bactotryptone, 0.5% yeast extract, 1% NaCl) containing glycerol (1–2%) and ampicillin (100  $\mu\text{g}/\text{ml}$ ) in a 7-l fermenter (LH Fermentation Ltd) and harvested by centrifugation. Cells were frozen immediately in liquid nitrogen and stored at  $-20^\circ\text{C}$ . Rubisco was purified from bacteria by the following procedure. Frozen cells (40 g) were thawed and resuspended in 200 ml of extraction buffer (50 mM Hepes pH 7.0, 0.1 mM EDTA, 10 mM  $\text{MgCl}_2$ , 2 mM phenylmethylsulphonyl fluoride, 50 mM  $\text{NaHCO}_3$ , 1 mM dithiothreitol) and sonicated on ice in batches of 50–75 ml, each for four 1-min periods. Care was taken to ensure that there was no significant rise in

temperature during sonication. Cell debris was removed by centrifugation for 20 min at  $4^\circ\text{C}$  at  $33000 \times g(\text{av})$  in an MSE HS21  $8 \times 50\text{-ml}$  rotor, and other insoluble material by centrifugation for 40 min at  $4^\circ\text{C}$  at  $120000 \times g(\text{av})$  in a Beckman 70Ti  $8 \times 30\text{-ml}$  rotor. The supernatant was made to a final concentration of 10% poly(ethyleneglycol) ( $M_r = 3500$ ) and left for 20 min at  $4^\circ\text{C}$ . Precipitated material was pelleted by centrifugation for 40 min at  $4^\circ\text{C}$  at  $120000 \times g(\text{av})$  in a Beckman Ti70 rotor. The supernatant was discarded and the pellet thoroughly drained, then resuspended in a minimum volume of 20 mM Tris/HCl pH 8.0, 1 mM  $\text{MgCl}_2$ , 1 mM dithiothreitol, 1 mM EDTA (column buffer A). This was then loaded onto a column (2.5 cm  $\times$  33 cm) of Fast-Flow Sepharose Q (Pharmacia), preequilibrated with column buffer A. The column was washed with column buffer A to elute unbound material and the bound protein was eluted with a gradient of 0–0.5 M NaCl in column buffer A. Fractions containing Rubisco were pooled and desalted by passage down a column (5 cm  $\times$  34 cm) of Sephadex G-25 (Pharmacia) in 5 mM Bicine/NaOH pH 8.0. Peak protein-containing fractions were identified by their  $A_{280}$ , pooled and loaded onto a second column of Fast-Flow Sepharose Q (2.5 cm  $\times$  36 cm) pre-equilibrated with 20 mM Hepes/NaOH pH 7.0, 1 mM  $\text{MgCl}_2$  (column buffer B). The column was washed with column buffer B and bound protein eluted with a gradient of 0–0.5 M NaCl in column buffer B. Fractions containing Rubisco were identified by enzyme assay, pooled and desalted by passage down a Sephadex G-25 column in 5 mM Bicine/NaOH pH 8.0. The purified enzyme was freeze-dried and stored at  $4^\circ\text{C}$  over desiccant.

### Rubisco assay

Rubisco activity in *E. coli* cells was measured by an adaptation of the method of Tabita et al. [25]. Bacteria were grown overnight in L broth and the absorbance of suitable diluted samples measured at 650 nm. Samples of 1 ml were then pelleted and resuspended in 100  $\mu\text{l}$  100 mM Bicine/NaOH pH 8.2, 20 mM  $\text{MgCl}_2$ , 50 mM  $\text{NaHCO}_3$  and mixed with 20  $\mu\text{l}$  toluene. Aliquots of 100  $\mu\text{l}$  were assayed as described by Gutteridge et al. [26] except that samples for estimation of  $^{14}\text{C}$  fixed in acid-stable form were removed at intervals, up to 3 min.

Rubisco activity in large numbers of samples such as column fractions were assayed by a simplified, bench-top assay carried out in scintillation vials. The vials contained 450  $\mu\text{l}$  100 mM Bicine/NaOH pH 8.2, 20 mM  $\text{MgCl}_2$ , 50 mM  $\text{NaH}^{14}\text{CO}_3$ , 0.33 mM ribulose- $P_2$  and the reactions were started by the addition of 50  $\mu\text{l}$  enzyme and shaken by hand for a few seconds. Reactions were carried out at room temperature for 5 min and stopped by the addition of 200  $\mu\text{l}$  10 M formic acid and  $^{14}\text{C}$  fixed in acid-stable form determined as previously [26].

To determine the kinetic constants the freeze-dried, purified enzyme was dissolved in  $\text{CO}_2$ -free buffer (100 mM Bicine/NaOH pH 8.2, 20 mM  $\text{MgCl}_2$ ) and rapidly desalted in the same buffer to remove any  $\text{CO}_2$  contaminating the enzyme [27]. The enzyme was then incubated in the presence of  $\text{NaH}^{14}\text{CO}_3$  and  $P_i$  to 50 mM and 10 mM respectively. The enzyme was fully activated within 3 min and could be stored at  $25^\circ\text{C}$  for at least 4 h without significant loss of activity.

For the  $K_m$  determinations, 10- $\mu\text{l}$  aliquots of activated enzyme (1–2 mg/ml) were assayed as described by Gutteridge et al. [26] except that the  $\text{HCO}_3^-$  and ribulose- $P_2$  concentrations defined in the legend to Table 1 were used. Kinetic

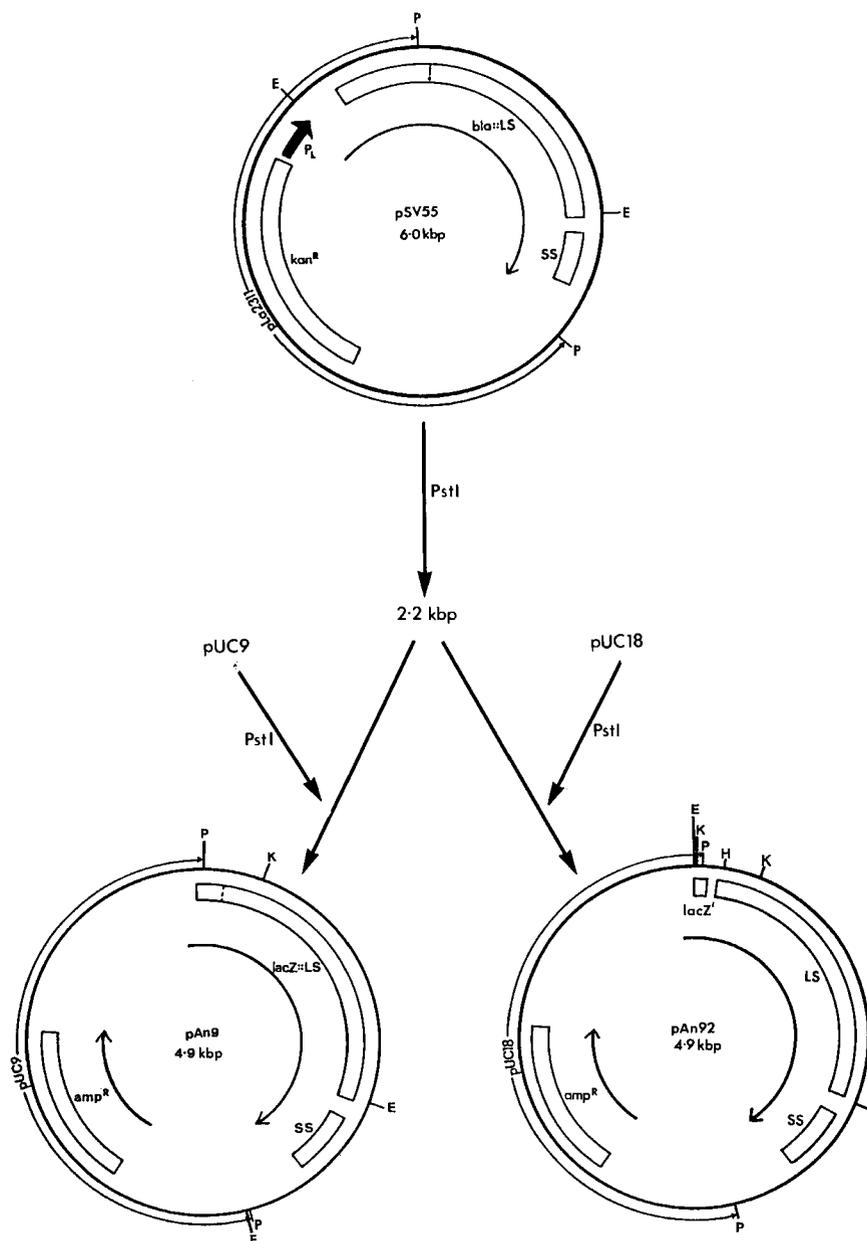


Fig. 1. Construction of plasmids expressing the large- and small-subunit genes of Rubisco from *Anacystis nidulans*. Bacteria containing pAn9 express the large subunit as a *lacZ::LS* fusion polypeptide; bacteria containing pAn92 express the wild-type large subunit. Restriction endonuclease sites are shown for *EcoRI* (E), *PstI* (P) and *KpnI* (K). The *HaeIII* (H) site involved in large-subunit gene fusions is also shown

constants were determined using a median method of analysis [28] with a microcomputer (Kontron Ltd, St Albans, UK). For the determination of specificity factors, the total consumption of ribulose- $P_2$  method [29] was used with a 50- $\mu$ l aliquot of activated enzyme (25 mg/ml).

## RESULTS

### *Expression of Anacystis nidulans Rubisco genes in Escherichia coli*

The 2.2-kb *PstI* fragment containing the genes for the large and small subunit of *A. nidulans* Rubisco [17] was excised from plasmid pSV55 [11] and inserted in the *PstI* site of pUC9 (Fig. 1). This construct was introduced into *E. coli* strain JM83 by transformation and ampicillin-resistant colonies iso-

lated. Suspension cultures of these clones were examined for ribulose- $P_2$ -dependent  $CO_2$  fixation and all were shown to have detectable levels of Rubisco activity. However, SDS/polyacrylamide gel electrophoresis followed by transfer to nitrocellulose and detection with antisera raised against wheat Rubisco (data not presented) showed that the major antigenic polypeptide synthesised by these bacteria was significantly larger than authentic Rubisco large subunit, as reported by van der Vies et al. [14] for a similar construct. Examination of the DNA sequence of the 2.2-kb *PstI* fragment [17, 30] indicated that this polypeptide was the result of a coincidence of reading frames between the Rubisco large subunit gene and the N-terminus of the *lacZ* reporter gene into which the *PstI* fragment had been inserted, resulting in a *lacZ::LS* fusion. In order to prevent this, the *PstI* fragment was inserted in the *PstI* site of pUC18, in which the reading frames no longer

coincide and where the *lacZ* translation product is terminated by a UGA codon upstream of the Rubisco large-subunit-initiating AUG codon. This construct was introduced into *E. coli* and yielded colonies which in liquid culture had Rubisco activity of 5–10-fold higher than the above construct in pUC9. One clone, pAn92 (Fig. 1), was selected as the basis for further experiments.

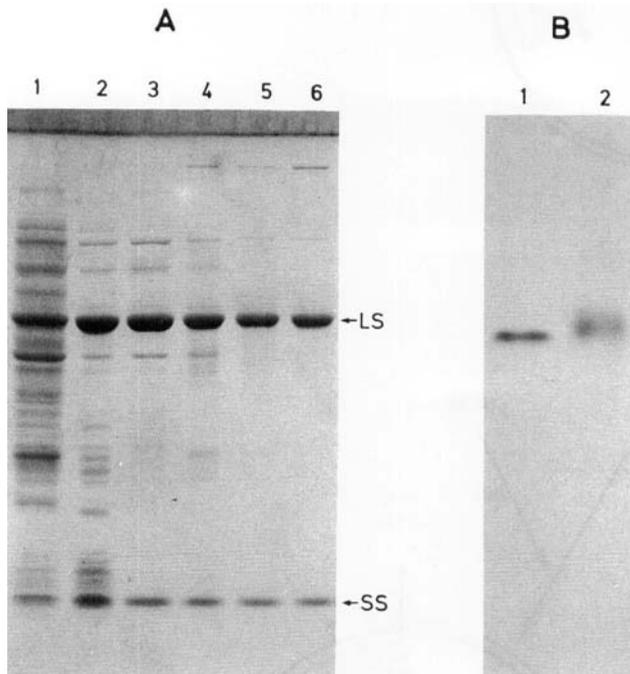


Fig. 2. Electrophoresis of Rubisco (A) during purification and (B) purified. (A) SDS/polyacrylamide gel electrophoresis of sequential steps in the purification of Rubisco from *E. coli* cells containing pAn92. (1) Crude extract from cell paste after sonication and centrifugation; (2) resuspended pellet following poly(ethyleneglycol) precipitation; (3) peak fractions after ion-exchange chromatography on Fast-Flow Sepharose Q at pH 8.0; (4) desalted eluate from Sephadex G-25; (5) peak fractions after ion-exchange chromatography on Fast-Flow Sepharose Q at pH 7.0; (6) eluate from Sephadex G-25. (B) Non-dissociating polyacrylamide gel electrophoresis of Rubisco purified from (1) bacteria expressing pAn92 and (2) from wheat leaves (*Triticum aestivum* cv. Sicco)

#### Purification and properties of Rubisco synthesized in *E. coli*

Initial attempts to purify Rubisco from bacteria expressing the *A. nidulans* genes involved sonication of the suspended cells, poly(ethyleneglycol) precipitation and sucrose density gradient centrifugation followed by FPLC ion-exchange chromatography on Mono-Q resin (Pharmacia) at room temperature. This, however, gave a low recovery and was a relatively slow procedure, because of the low capacity of the Mono-Q column. This resulted in a degree of proteolysis of the large subunit which was unaffected by the addition of protease inhibitors. A more rapid method was therefore developed which is outlined above, resulting in pure protein within two days, with little evidence of proteolytic attack. The yield of purified protein ranged between 50–200 mg/l bacterial culture.

Electrophoresis of the Rubisco synthesised in *E. coli* in a non-denaturing polyacrylamide gel showed a single protein which migrated slightly faster in the gel than Rubisco isolated from wheat leaves (Fig. 2B). When the recombinant Rubisco was denatured and examined by electrophoresis in polyacrylamide gels containing sodium dodecyl sulphate as denaturant, the large and small subunits of the complex were resolved (Fig. 2A). Analysis of the staining intensities of the large and small polypeptides were consistent with a 1:1 ratio of subunits, suggesting that the enzyme had the normal  $L_8S_8$  stoichiometry. The recombinant Rubisco-catalysed carboxylation of ribulose- $P_2$  with a  $V_{max}$  of  $5.0 \mu\text{mol CO}_2 \text{ min}^{-1} (\text{mg protein})^{-1}$  at  $25^\circ\text{C}$ . The  $K_m$  for ribulose- $P_2$ ,  $\text{CO}_2$  and  $\text{O}_2$ , and the specificity factor, a measure of the relative proportions of carboxylase and oxygenase activities, were also determined and are shown in Table 1.

#### Construction and expression of chimaeric genes for Rubisco

A low level of homology in the DNA between the gene for the *A. nidulans* and wheat/maize large subunits of Rubisco allowed the construction of chimaeric genes in which the DNA encoding the N-terminal 12 amino acids of the large subunit of *A. nidulans* were replaced by the equivalent region from wheat and from maize, which encodes 15 amino acids in each case (Fig. 3). These constructions consisted of a 74-bp *EcoRI*–*AluI* fragment from the wheat large-subunit gene contained within pTac39 (T. A. Dyer and R. Barker, personal

Source of N-terminus	Clone	N-terminal amino acid sequence
<i>A. nidulans</i>	pAn92	MPKQTOSAAGYKA
Maize	pZmAn135	MSPQTETKRASVGFKA
Wheat	pTaAn158	MSPQTETKAGVGFKA
<i>lacZ</i>	placAn50	MITNSP

GVKDYKLTYY.....

Fig. 3. Predicted amino acid sequence of the N-terminal of wild-type and chimaeric large subunits of Rubisco produced by gene fusions. In the gene fusions the 5' region of the *A. nidulans* large-subunit gene [17] encoding 12 N-terminal amino acids was replaced with DNA encoding wheat (T. A. Dyer and R. Barker, personal communication) or maize [31] large-subunit N-terminus, or with DNA from the pUC8 *lacZ* gene polylinker. Sequence identity between species is indicated; also shown are the sites of trypsin cleavage (arrow heads) which abolish catalytic activity of wheat Rubisco [5]

Fig. 4. Construction of plasmids containing chimaeric genes designed to express *A. nidulans* Rubisco large subunit with changed N-termini, together with small subunit. pTaAn158, pZmAn135 and placAn50 express large subunit with the N-terminal 12 amino acids replaced by residues from wheat, maize and pUC8 polylinker (*lacZ*) respectively. Restriction endonuclease sites shown: *EcoRI* (E), *KpnI* (K), *PstI* (P), *BamHI* (B), *HincII* (H2), *AluI* (A), *HaeIII* (H). Only the *AluI* and *HaeIII* sites involved in construction of LS chimaeras are shown, denoted by A/H (*AluI*/*HaeIII* ligation) and S/H (*SmaI*/*HaeIII* ligation)

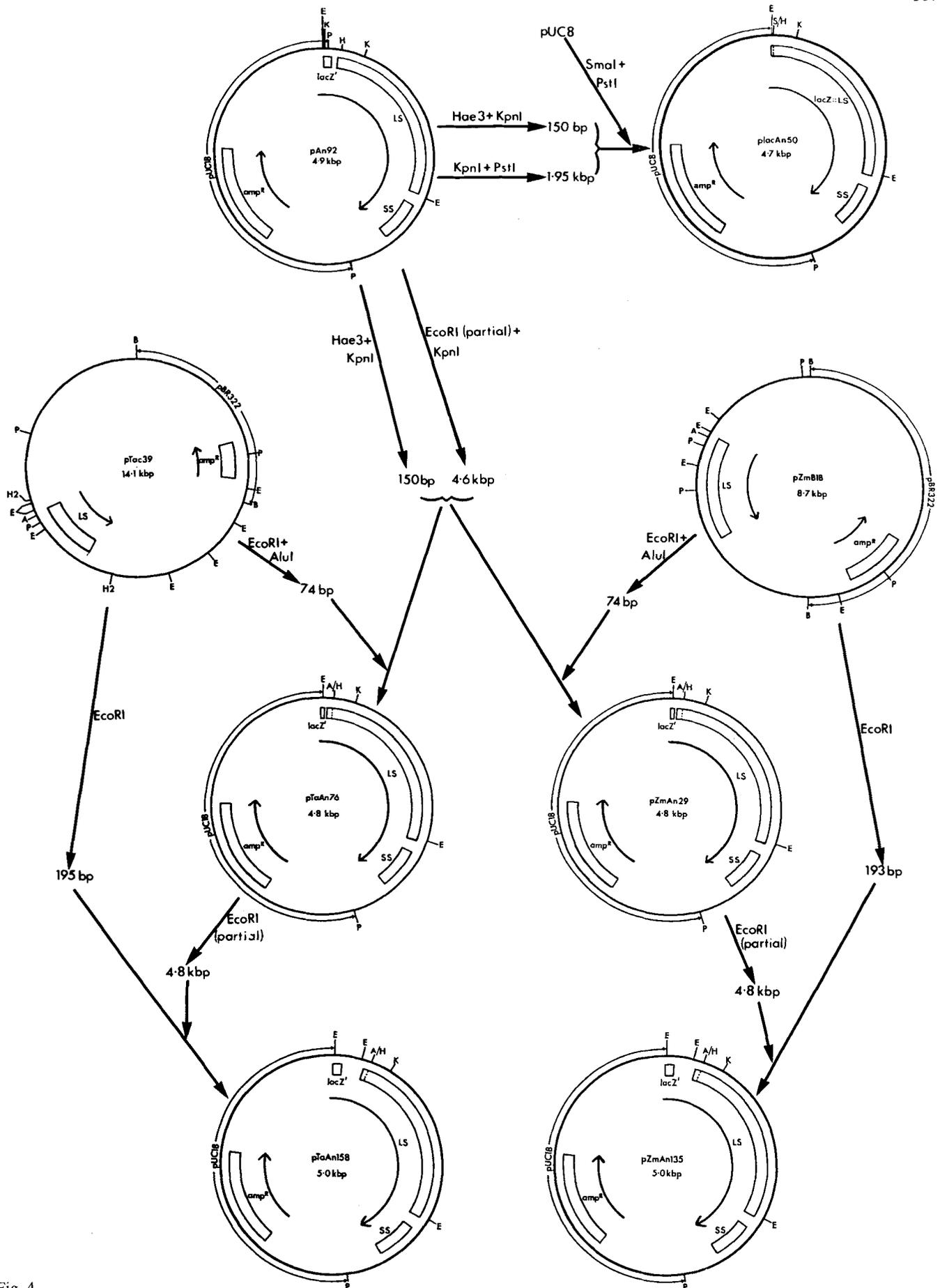


Fig. 4

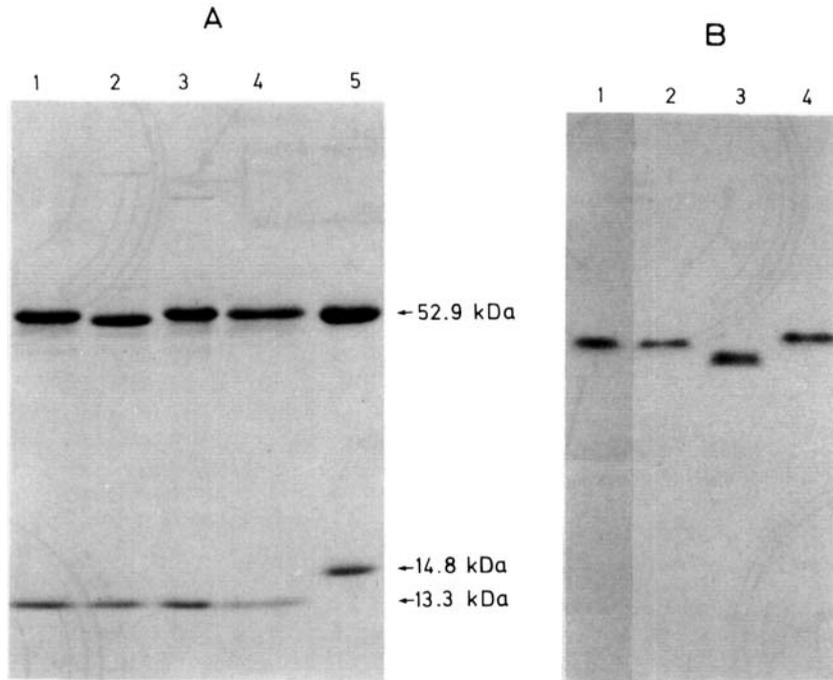


Fig. 5. Electrophoresis of purified wild-type and chimaeric Rubiscos. (A) SDS/polyacrylamide gel electrophoresis of purified wild-type and chimaeric Rubiscos from bacteria containing (1) pZmAn135 (LS with maize N-terminus), (2) placAn50 (*lac*::LS truncated fusion), (3) pAn92 (wild type *A. nidulans* Rubisco), (4) pTaAn158 (LS with wheat N-terminus), (5) purified wheat Rubisco. (B) Non-denaturing polyacrylamide gel electrophoresis of purified wild-type and chimaeric Rubiscos from bacteria containing (1) pTaAn158 (LS with wheat N-terminus), (2) pZmAn135 (LS with maize N-terminus), (3) placAn50 (*lac*::LS truncated fusion), (4) pAn92 (wild-type *A. nidulans* Rubisco)

communication), or a 74-bp *EcoRI*–*AluI* fragment from the maize large-subunit gene contained within pZmB1B [18, 31], ligated to a 150-bp *HaeIII*–*KpnI* fragment from the *A. nidulans* large subunit gene. These were fused to the remainder of the *A. nidulans* Rubisco large subunit and small subunit genes carried on a 4.6-kb *KpnI*–*EcoRI* fragment also including pUC18 (Fig. 4). This yielded pTaAn76, containing the wheat N-terminus, and pZmAn29, containing the maize N-terminus (Figs 3 and 4). These constructs were introduced into *E. coli* JM83 by transformation, but problems in achieving high expression were encountered. This was possibly due to the comparatively short distance of 24 bp between a UAA codon terminating *lacZ* translation and the AUG codon initiating large-subunit translation in each case. In order to test this, pTaAn76 and pZmAn29 were linearized by partial digestion with *EcoRI* and ligated to the *EcoRI* fragment which lies upstream of the large subunit in wheat (195 bp) and maize (193 bp), yielding plasmids pTaAn158 and pZmAn135 (Fig. 4). The identity of these plasmids was confirmed by plasmid isolation and restriction endonuclease digestion and by sequencing of the double-stranded DNA from the reverse primer by the dideoxy nucleotide chain-termination method [22]. The constructs were introduced into *E. coli* JM83 and ampicillin-resistant colonies examined for ribulose- $P_2$  dependent  $CO_2$  fixation. Both chimaeric constructs gave rise to colonies which in suspension culture had Rubisco activity similar to pAn92.

A third plasmid was also constructed where the *HaeIII*–*KpnI*–*PstI* assembly was ligated into pUC8 cut with *SmaI* and *PstI*, which yielded a *lacZ*::LS fusion such that the DNA encoding the 12 N-terminal residues of *A. nidulans* large-subunit gene were replaced with *lacZ* DNA encoding six unrelated amino acids (Figs 3 and 4). This construct,

placAn50, was introduced into JM83 and showed a low level of ribulose- $P_2$  dependent  $CO_2$  fixation.

#### Assembly and enzymology of Rubisco from chimaeric genes

Bacteria containing pTaAn158, pZmAn135 and placAn50 were used to inoculate L broth in a 7-l fermenter and the mutant enzymes purified from the cells as described in Methods. Samples of each enzyme with an altered N-terminus were electrophoresed in the absence (Fig. 5B) and presence (Fig. 5A) of sodium dodecyl sulphate. In the presence of SDS, all purified enzymes separated into large and small subunits of the expected sizes and stoichiometries. Bacteria containing placAn50, the *lac*::LS fusion, produced a slightly smaller large subunit, as predicted from the amino acid sequence (Fig. 3). Under non-denaturing conditions, the enzymes all ran in similar positions to the wild-type enzyme isolated from bacteria transformed with pAn92, indicating that the substitution of N-terminal residues in the large subunit had no effect on assembly of the  $L_8S_8$  complex.

Preliminary experiments had demonstrated that there was no difference between the wild-type and chimaeric enzymes with respect to rate of activation or allosteric response to  $NaHCO_3$ ,  $MgCl_2$  and  $P_i$ . A standard activation scheme was adopted, where enzyme was incubated for 5 min at 25°C in 50 mM Bicine/NaOH pH 8.2 containing 20 mM  $MgCl_2$ , 50 mM  $NaHCO_3$  and 10 mM  $P_i$ . Unlike freeze-dried preparations of purified wheat and spinach Rubisco [26], the enzyme preparations of wild-type and chimaeric *Anacystis* Rubisco did not require a heat-activation step.

The carboxylase activity of the wild-type enzyme was within the range of values previously reported for the native enzyme extracted from *A. nidulans* [9, 32]. The catalytic ac-

Table 1. Kinetic parameters of wild-type and mutant Rubisco

For determining the  $K_m$  of  $\text{CO}_2$ , a ribulose- $P_2$  concentration of 0.46 mM was used (1.66 mM for placAn50) and  $\text{HCO}_3^-$  concentrations of 2.7–46 mM. For determining the  $K_m$  of ribulose- $P_2$ , an  $\text{HCO}_3^-$  concentration of 46 mM was used and ribulose- $P_2$  concentrations of 0.023–0.46 mM (0.05–2.29 mM for placAn50). The ribulose- $P_2$  concentration was quantified enzymatically. Figures in parentheses are 95% confidence limits of the determined values. The specificity factor ( $\tau$ ) is defined as  $\tau = [k_{\text{cat}}(\text{carboxylation}) \times K_m(\text{O}_2)] / [k_{\text{cat}}(\text{oxygenation}) \times K_m(\text{CO}_2)]$

Source of Rubisco	$K_m$ for		$V_{\text{max}}$	Specificity factor
	$\text{CO}_2$	ribulose- $P_2$		
	$\mu\text{M}$		$\mu\text{mol CO}_2 \text{ min}^{-1} \text{ mg}^{-1}$	
Wild-type	101	32	5.4	58
<i>A. nidulans</i> pAn92	(86–113)	(28–35)	(4.2–5.6)	
Chimaeras with:				
wheat N-terminus	100	31	5.7	58
pTaAn158	(83–121)	(26–37)	(5.3–6.3)	
maize N-terminus	93	37	5.0	59
pZmAn135	(87–104)	(33–39)	(4.9–5.4)	
<i>lacZ</i> N-terminus	246	380	2.4	56
placAn50	(180–283)	(353–420)	(2.1–2.9)	

tivity of the chimaeric enzymes with wheat (pTaAn158) and maize (pZmAn135) N-terminal amino acids were comparable with the wild-type enzyme, but under standard assay conditions the catalytic activity of the mutant with six 'nonsense' amino acids at its N-terminus (placAn50) was reduced by 90%. The responses to varying substrate concentrations are given in Table 1. When the ribulose- $P_2$  concentration was varied the protein concentration was adjusted to ensure that even at low concentration only a small percentage of the ribulose- $P_2$  was consumed. Only the mutant with the six 'nonsense' amino acids at the N-terminus, placAn50, differed significantly from the wild type. This mutant had a greatly reduced affinity for ribulose- $P_2$  which increased the  $K_m$  for this compound from 32 mM to 410 mM. The substitution of the wheat or maize N-termini had little effect on the  $K_m$  value.

When the ribulose- $P_2$  concentration was held constant at a saturating concentration and the  $\text{CO}_2$  concentration varied, neither of the chimaeric enzymes with wheat or maize large-subunit N-termini showed any differences in  $K_m$  for  $\text{CO}_2$ , whereas that for the enzyme with the *lacZ*:LS fusion showed an approximate twofold increase compared with Rubisco expressed from the wild-type genes. The specificity factor for the two gaseous substrates was also determined, according to the total consumption of ribulose- $P_2$  method [29]. None of the mutant enzymes produced from chimaeric genes were significantly different from the wild-type value, showing that this N-terminal region has no effect on partitioning of the gaseous substrates.

## DISCUSSION

We have achieved expression of *Anacystis nidulans* large- and small-subunit genes in *E. coli* by inserting a *Pst*I fragment containing both genes as a dicistron [17] into pUC18, where transcription is under the control of the *lac* promoter. This gives an extremely high level of expression of the genes and yields active Rubisco at up to 10% of soluble bacterial protein. A UGA codon upstream of the large-subunit AUG initiation codon terminates *lacZ* translation and thus prevents the syn-

thesis of a *lacZ*:LS fusion polypeptide, which was found when the genes were inserted into pUC9 ([14] and this paper). The mRNA transcribed from the *lac* promoter therefore contains three translated open reading frames: a truncated *lacZ* polypeptide and the large and small Rubisco subunits; the large amount of Rubisco synthesized may be due to translational coupling between the three open reading frames, as is believed to occur in, for example, the *E. coli* operons for ribosomal proteins [33] and ATP synthase [34].

The large and small subunit polypeptides assemble correctly into a complex which could be isolated in high yield. Non-denaturing polyacrylamide gel electrophoresis, enzyme specific activity and staining intensities on denaturing PAGE of this enzyme indicate that it has the normal  $\text{L}_8\text{S}_8$  stoichiometry. This is in contrast to the results of Gatenby and colleagues [11], who estimated that a similar construct in pUC9 produced Rubisco with an approximate  $\text{L}_8\text{S}_{2-3}$  stoichiometry. This may have been due to the *lacZ*:LS fusion which this construct (pDB50) encodes, loss of small subunits during enzyme purification, or possibly to differences in *E. coli* strain or growth conditions.

The N-terminus of the large subunit of higher plant Rubisco is known to be sensitive to proteolysis, but rapid isolation and the inclusion of a protease inhibitor prevented such damage to the recombinant *Anacystis* Rubisco. Treatment of the purified enzyme with trypsin removed peptides of similar size as from the wheat enzyme [5], whereas the enzyme isolated from bacteria containing placAn50, which lacks the lysine residues forming the trypsin-sensitive sites (see Fig. 3), was unaffected by this treatment (data not shown). This indicates that the N-termini of the large subunits of our recombinant enzymes were not subjected to proteolysis during purification.

It has been previously shown that amino acid residues 9–14 from the N-terminus of the large subunit of wheat Rubisco are essential for full activity of the enzyme, and that residues 1–8 are non-essential [5]. These two regions are only poorly conserved between Rubisco of the  $\text{L}_8\text{S}_8$  type (see Fig. 3): only three residues are conserved between wheat and *A. nidulans*. Since higher plant and *Anacystis* Rubisco are also considerably different in their enzymological parameters, it is possible that the N-terminal region is important in determining these differences. However, when the chimaeric constructs described above were expressed in *E. coli*, effectively replacing the 12 amino-terminal residues of the *Anacystis* large subunit with the 15 amino-terminal residues from wheat or maize (Fig. 3), the properties of the resulting Rubisco (Table 1) were indistinguishable from Rubisco isolated from bacteria expressing normal *Anacystis* large- and small-subunit genes. Thus differences in catalytic properties between Rubisco from wheat, maize and *A. nidulans* are not due to this N-terminal region.

The replacement of the DNA encoding the 12 amino-terminal residues in *Anacystis* large subunit by DNA encoding six unrelated residues derived from the pUC8 *lacZ* gene fragment of pUC8 results in the synthesis of a large subunit truncated at the N-terminus (Fig. 3). Although this appears to assemble correctly with *A. nidulans* small subunit, it has only 10% of wild-type activity under standard assay conditions. This confirms the earlier report from this laboratory that amino acids in the N-terminal region are essential for full catalytic activity [5]. Further analysis of this truncated mutant has revealed that the loss of activity is due to an approximate 50% decrease in  $V_{\text{max}}$  combined with a tenfold increase in the  $K_m$  for ribulose- $P_2$  and a twofold increase in the  $K_m$  for  $\text{CO}_2$ .

This implies that the N-terminus is involved in some way in forming the binding site for ribulose- $P_2$ , but there are no residues in this region which are conserved between all species, which suggests that it may be involved in maintaining the geometry of the active site rather than having a direct role in ribulose- $P_2$  binding. It is possible that the changes in  $V_{\max}$  and  $K_m$  for  $CO_2$  are a consequence of this change in the ribulose- $P_2$  binding site. The recent crystal structure of Rubisco from *Rhodospirillum rubrum* [35], which is a simple dimer of two large subunits, shows the N-terminus of each subunit to be relatively close to the C-terminal domain of the other subunit, where most of the residues thought to be involved in catalysis are located, although not close enough to interact directly with the substrate. However, the N-terminus of  $L_8S_8$ -type Rubiscos is some 10–14 residues longer and may extend into the active site. Also the crystal structure of the bacterial enzyme is of the inactive form and activation may alter considerably the positions of many regions of the polypeptides.

The authors thank T. A. Dyer and A. A. Gatenby for gifts of plasmids.

## REFERENCES

- Schloss, J. V. & Lorimer, G. H. (1982) *J. Biol. Chem.* **257**, 4691–4694.
- Miziorko, H. M. & Lorimer, G. H. (1983) *Annu. Rev. Biochem.* **52**, 507–535.
- Lorimer, G. H. & Miziorko, H. M. (1980) *Biochemistry* **19**, 5321–5328.
- Hartman, F. C., Stringer, C. D., Milanez, S. & Lee, E. H. (1986) *Phil. Trans. R. Soc. Lond. Ser. B* **313**, 379–395.
- Gutteridge, S., Millard, B. N. & Parry, M. A. J. (1986) *FEBS Lett.* **196**, 263–268.
- Bradley, D., van der Vies, S. M. & Gatenby, A. A. (1986) *Phil. Trans. R. Soc. Lond. Ser. B* **313**, 447–458.
- Barraclough, R. & Ellis, R. J. (1980) *Biochim. Biophys. Acta* **608**, 19–31.
- Henningsen, S. M. & Ellis, R. J. (1986) *Plant Physiol.* **80**, 269–276.
- Andrews, T. J. & Abel, K. M. (1981) *J. Biol. Chem.* **256**, 8445–8451.
- Andrews, T. J. & Ballment, B. (1983) *J. Biol. Chem.* **258**, 7514–7518.
- Gatenby, A. A., van der Vies, S. M. & Bradley, D. (1985) *Nature (Lond.)* **314**, 617–621.
- Tabita, F. R. & Small, C. L. (1985) *Proc. Natl Acad. Sci. USA* **82**, 6100–6103.
- Christeller, J. T., Terzaghi, B. E., Hill, D. F. & Laing, W. A. (1985) *Plant Mol. Biol.* **5**, 257–263.
- van der Vies, S. M., Bradley, D. & Gatenby, A. A. (1986) *EMBO J.* **5**, 2439–2444.
- Gurevitz, M., Sommerville, C. R. & McIntosh, L. (1985) *Proc. Natl Acad. Sci. USA* **82**, 6546–6550.
- Viale, A. M., Kobayashi, H., Takabe, T. & Akazawa, T. (1985) *FEBS Lett.* **192**, 283–288.
- Shinozaki, K. & Sugiura, M. (1983) *Nucleic Acids Res.* **11**, 6957–6964.
- Bedbrook, J. R., Coen, D. M., Beaton, A. R., Bogorad, L. & Rich, A. (1979) *J. Biol. Chem.* **254**, 905–910.
- Gatenby, A. A., Castleton, J. A. & Saul, M. N. (1981) *Nature (Lond.)* **291**, 117–121.
- Howe, C. J., Bowman, C. M., Dyer, T. A. & Gray, J. C. (1982) *Mol. Gen. Genet.* **186**, 525–530.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular cloning: a laboratory manual*, Gold Spring Harbor Laboratory, New York.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl Acad. Sci. USA* **74**, 5463–5467.
- Chen, E. Y. & Seeburg, P. H. (1985) *DNA* **4**, 165–170.
- Hattori, M. & Sakaki, Y. (1986) *Anal. Biochem.* **152**, 232–238.
- Tabita, F. R., Caruzo, P. & Whitman, W. (1978) *Anal. Biochem.* **84**, 462–472.
- Gutteridge, S., Parry, M. A. J. & Schmidt, C. N. G. (1982) *Eur. J. Biochem.* **126**, 597–602.
- Helmerhorst, E. & Stokes, G. B. (1980) *Anal. Biochem.* **104**, 130–135.
- Cornish-Bowden, A. & Eisenthal, R. (1978) *Biochim. Biophys. Acta* **523**, 268–272.
- Gutteridge, S., Phillips, A. L., Kettleborough, C. A., Parry, M. A. J. & Keys, A. J. (1986) *Phil. Trans. R. Soc. Lond. Ser. B* **313**, 433–445.
- Shinosaki, K., Yamada, C., Takahata, N. & Sugiura, M. (1983) *Proc Natl Acad. Sci. USA* **80**, 4050–4054.
- McIntosh, L., Poulsen, C. & Bogorad, L. (1980) *Nature (Lond.)* **288**, 556–560.
- Andrews, T. J. & Lorimer, G. H. (1985) *J. Biol. Chem.* **260**, 4632–4636.
- Nomura, M., Gourse, R. & Baughman, G. (1984) *Annu. Rev. Biochem.* **53**, 75–117.
- McCarthy, J. E. G., Schairer, H. U. & Sebald, W. (1985) *EMBO J.* **4**, 519–526.
- Schneider, G., Lindqvist, Y., Brändén, C.-I. & Lorimer, G. H. (1986) *EMBO J.* **5**, 3409–3415.