This is to certify that the

dissertation entitled

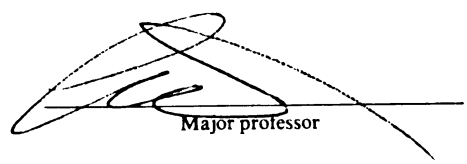
EXPRESSION OF THE SUBUNITS OF RIBULOSE 1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE IN E. coli, AND MUTATIONAL ANALYSIS OF THE SMALL SUBUNIT

presented by

JOHN FITCHEN

has been accepted towards fulfillment
of the requirements for

Ph. D. degree in Genetics



Major professor

Date July 18, 1989

EXPRESSION OF THE SUBUNITS OF RIBULOSE 1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE IN *E. coli*, AND MUTATIONAL ANALYSIS OF THE SMALL SUBUNIT

by

John Fitchen

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Program in Genetics

1989

ABSTRACT

EXPRESSION OF THE SUBUNITS OF RIBULOSE 1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE IN Escherichia coli, AND MUTATIONAL ANALYSIS OF THE SMALL SUBUNIT

by

John Fitchen

Ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO) from plants, cyanobacteria, and most other photosynthetic organisms is composed of eight identical "large subunits" ($\approx$55 kDa) and eight identical "small subunits" ($\approx$15 kDa). The large subunits contribute all residues directly participating in the catalytic mechanism. The function of the small subunits is unclear.

Plasmids were constructed to direct the expression of RuBisCO subunits in Escherichia coli using the genes encoding the large and small subunits (rbcL and rbcS), from Zea mays and from the cyanobacterium Anabaena sp. PCC 7120. Both of the subunits accumulated to high levels in E. coli carrying the expression plasmids. Only in those combinations in which the large subunits were from the cyanobacterial rbcL did the subunits assemble into a functional holoenzyme.

The simultaneous expression of both cyanobacterial subunits in E. coli was used to investigate the role of conserved residues in the small subunit. Twenty-two directed mutations were made in the

three regions of broadest sequence conservation. The extent to which the mutant small subunit proteins were associated with large subunits was assessed, and the carboxylase activity of the mutant enzymes was assayed. At least one mutation in each of the three conserved regions was found to completely prevent holoenzyme assembly. Other mutations were found to result in decreased levels of assembly. None of the mutations prevented catalysis except by interfering with holoenzyme assembly.

Examination of homologous residues in a model for the three-dimensional structure of spinach RuBisCO suggested that some of the mutants did not assemble because of mis-pairing or non-pairing of charged residues on the interfacing surfaces of the large and small subunits. In others, the failure to assemble correctly may have resulted from disruption of intersubunit hydrophobic pockets. The molecular interactions impaired by the substitutions that reduced but did not prevent assembly appeared to be similar to those disrupted in the mutants in which assembly was completely blocked.

To all of my teachers

ACKNOWLEDGMENTS

I would like to thank Lee McIntosh for his help and guidance and support for the experiments described in this thesis. Erin Bell and Mickey Gurevitz and other members of the McIntosh laboratory contributed directly by supplying plasmids and antisera, and I would like to thank them and all other members, past and present, of the McIntosh Lab for creating an environment that could be both fun and challenging. Chris Somerville, Ed Tolbert, and Hans Kende are also deserving of thanks for their interest and advice, and for serving on my guidance committee.

Carl-Ivar Branden and the members of his laboratory provided access to the unpublished coordinates of their model for the 3-D structure of RuBisCO and were gracious hosts during my brief stay in Uppsala.

Finally, I would like to thank the many friends who have shared their enthusiasm and wisdom and contributed substantially to my education.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

$RbuP_2$	ribulose 1,5-bisphosphate
RuBisCO	$RbuP_2$ carboxylase/oxygenase
L	large subunit of RuBisCO
rbcL	gene encoding L
S	small subunit of RuBisCO
rbcS	gene encoding S
SDS	sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
kDa	kilodalton
bp	nucleotide base pair
kb	kilobase

Chapter 1 GENERAL INTRODUCTION

LITERATURE REVIEW

RuBisCO (Ribulose 1,5-bisphosphate carboxylase/oxygenase, E. C. 4.1.1.39) catalyzes both the oxygenation and the carboxylation of ribulosebisphosphate (RbuP2). The reaction of RbuP2 with CO2 to form a six carbon intermediate which is then hydrated and cleaved to form two molecules of 3-P-glycerate represents the first committed step of the reductive photosynthetic carbon cycle (C-3 photosynthetic pathway). The reaction of RbuP2 with O2 to form 2-P-glycollate and 3-P-glycerate represents the first committed step of the oxidative photosynthetic carbon cycle (photorespiration). Because of this central, decisive role in partitioning between photosynthesis and photorespiration RuBisCO and its two competing activities are important not only in determining plant growth, but also in modulating global biomass production and atmospheric CO2 levels.

RuBisCO Structure

The protein now known to be RuBisCO was initially identified in

spinach leaf extracts on the basis of its abundance and electrophoretic mobility (1). Subsequently isolated from scores, if not hundreds, of organisms, the enzyme has been extensively characterized.

The simplest form of RuBisCO, designated form II, is a homodimer of 57 kDa subunits (2,3). This form is found in purple, non-sulfur, photosynthetic bacteria (3,4,5). The primary sequence of the enzyme from Rhodospirillum rubrum was determined by protein sequencing (6). Genes for the form II enzymes of R. rubrum and Rhodobacter sphaeroides have been cloned and the R. rubrum gene has been sequenced (7,8). The crystal structure of the R. rubrum enzyme has been determined to 2.9 Å resolution (9).

The subunit structure of RuBisCO present in plants, cyanobacteria, and some other photosynthetic bacteria is more complex and is referred to as form I RuBisCO. The enzyme from these organisms is composed of eight identical subunits of molecular weight 50-55 kDa and eight identical subunits of molecular weight 12-18 kDa (10,11). Genes from many organisms encoding both types of subunits have been cloned and sequenced (for examples 12,13). Partial amino acid sequences have been determined by protein sequencing of subunits isolated from several organisms and at least one complete sequence for each type of subunit has been determined chemically (14,15). The crystal structures of the enzyme from tobacco and spinach have been determined to 2.6 Å and 2.8 Å resolution, respectively (16,17,18,19).

The 50-55 kDa subunit of the heteromultimeric form I enzyme is referred to as the large subunit (L). The deduced amino acid sequences of the large subunits are highly conserved between most

organisms. (For example the deduced amino acid sequences of spinach (20) and Anabaena (21), a cyanobacterium, are 85 % identical although a few known sequences, such as the one deduced for Alkaligenes eutrophus (22), a hydrogen bacterium, are more divergent.) The three-dimensional structure of the large subunits from spinach and tobacco are very similar to the three-dimensional structure of the form II R. rubrum subunit (16,18) and the arrangement of the 55 kDa subunits in the plant enzyme crystal structures is a pattern of four dimers (essentially a tetramer of form II enzymes) with the 15 kDa subunits attached peripherally (16,19). The sequences of the typical large subunits, although only about 25 to 40 % identical overall to the sequences of the 57 kDa subunits of the form II enzyme (6,7), contain short regions of almost complete identity to the form II sequence near residues of defined function (6,7). These aspects of similarity in both structure and sequence are strong enough to suggest a common origin (7).

The amino acid sequences of the 12-18 kDa subunits, referred to as the small subunits (S), are much more divergent between organisms. Within higher plants there is at least 55% identity between known sequences (23), but the sequences of S from two cyanobacteria are only about 40% identical to representative plant sequences (21,24). The amino acid sequence of the eukaryotic alga Chlamydomonas reinhardii is about 50% identical to the higher plant S sequences (25). The deduced amino acid sequences reported for the eukaryotic alga Olisthodiscus luteus and the prokaryote Alkaligenes eutrophus are 57% identical to each other but are only roughly 25% identical to higher plant sequences (R. A. Cattolico, personal communication; 22).

As with comparisons between form I large subunit and the form II subunit, conservation in some regions of S is much more striking than indicated by figures for overall identity (23).

RuBisCO Biosynthesis

The gene encoding the large subunit (rbcL) is part of the chloroplast genome (26,27), and transcription and translation of the large subunits is effected by chloroplast RNA polymerase and chloroplast ribosomes (28,29). The expression of rbcL in vivo is under environmental (e. g. light) and developmental control in plants (30,31). At some time after synthesis the large subunits are post-translationally processed (32,33). The two amino-terminal amino acids are removed; the next residue, a proline, is acetylated (33); and, at least in some species, one or more lysine residues are tri-methylated (34). It is not known whether these modifications are made before or after assembly with small subunits.

The gene for the small subunit (rbcS) is located in the nuclear genome (35). Its expression is also under environmental (e. g. light) and developmental control (30,36), although, in cyanobacteria, expression does not appear to be regulated by light (24). The protein is encoded in precursor form with an amino-terminal (transit peptide) extension of about 45 residues (35,37). The precursor is translated on free ribosomes in the cytosol (38,39), recognizes a protein on the surface of the chloroplast (40,41), is transported through the double-membrane chloroplast envelope (35), and is processed to its mature form (35). The processing probably occurs in two steps (42) with at least the final step being catalyzed by a

soluble stromal enzyme (35). Individual residues and groups of residues in the transit peptide, at the site of proteolytic cleavage between the transit peptide and the mature protein, and within the mature small subunit have been implicated in various aspects of uptake into chloroplasts, processing, and assembly into holoenzyme in studies utilizing directed mutagenesis and uptake of mutant small subunit precursors with isolated chloroplasts in vitro (43,44,45,46,47).

Current evidence suggests that individual subunits, however, do not spontaneously assemble to form the holoenzyme, but rather require the action of yet another protein (for review, see 48). This protein, comprised of two nuclear-encoded subunits, is observed to associate with newly synthesized large subunits and has been named the RuBisCO subunit binding protein (49,50). Antibodies directed against this protein can inhibit in vitro RuBisCO assembly in chloroplast extracts (51). The gene encoding one of the subunits has been cloned from wheat and castor bean (52). Comparison of their sequences with other known sequences suggests that the RuBisCO subunit binding protein is a member of a class of proteins called chaperonins which are believed to function in macromolecular assembly in bacteria, chloroplasts, and mitochondria (reviewed in 53). It has been postulated that the stage of RuBisCO assembly for which the binding protein is required is the formation of L subunit dimers or octomers (54).

The biosynthesis of RuBisCO in cyanobacteria is considerably less complicated because both subunits are encoded in the bacterial chromosome (55) and because the lack of internal compartmentation

obviates the need for either subunit to be transported across a membrane. The presence of a chaperonin in cyanobacteria has not been demonstrated (56) but assembly of cyanobacterial rubisco subunits expressed in a heterologous system has been shown to be dependent on the endogenous chaperonin (57). Whether the large subunit, or even the small subunit, undergo post-translational processing in cyanobacteria has not been addressed.

Structure/Function Relationships of RuBisCO

The activation of the enzyme by CO_2 and Mg^{++} and the catalytic mechanism of the carboxylase reaction have been studied in extensive detail. The oxygenase reaction is less well understood, but shares many of the features (e.g. activation and $RbuP_2$ binding) of the carboxylase reaction. Both are extensively covered in a recent review (58).

Large subunit residues

Affinity labels, side-chain-specific modifying reagents, and, more recently, site-directed mutagenesis have been used to identify individual residues of the protein participating in each of the binding steps and partial reactions of catalysis. The R. rubrum enzyme was used for most of the studies because of its simpler structure.

The lysine which is carbamylated in the activation of the enzyme was localized to the large subunit by carboxymethylation with ^{3}H-diazomethane (59) and identified to be lysine 191 of the R. rubrum enzyme by sequencing of the labeled peptide (60). This residue was

replaced with a glutamate by *in vitro* mutagenesis by Estelle *et al.* (61) creating an enzyme incapable of catalysis, and confirming the essential role of this residue.

Methionine 330 was identified as a target of affinity labels that inactivated the enzyme (62). Mutagenesis of the Met codon to a Leu codon resulted in an enzyme with a much higher Km for all substrates (63). The reversibility of binding of 2-carboxyribitol 1,5-bisphosphate (a transition state analog which binds the wild-type enzyme irreversibly) by the mutant enzyme suggested that Met 330 is important for binding or stability of the ene-diol intermediate formed prior to attack by CO_2 or O_2 (63).

A region near the amino terminus of the protein is thought to be part of the active site because several residues in the region can be crosslinked to known active-site residues (64). Glutamate 48 of the *R. rubrum* enzyme is the only absolutely conserved residue in the region (64). Creation of a glutamine-48 mutant of the *R. rubrum* enzyme, and demonstration that it was catalytically inactive, established that glutamate 48 is an essential residue, but a function has not been assigned (64).

Affinity labeling and comparitive sequence analysis placed Lys at or near the active site (65). Because the epsilon-amino group of the residue has an unusually low pKa (7.9) it was proposed that this residue could be the nucleophile which removes the C-3 proton from $RbuP_2$ (65). The properties of a series of mutant enzymes in which Lys 166 was replaced by other amino acids are consistant with this role (66).

Several of the tentative functional assignments made to individual residues on the basis of biochemical evidence, however, have not withstood the genetic test made possible by in vitro mutagenesis. It was originally proposed that aspartate 188, because of its proximity to lysine 191, could be one of the acidic residues coordinating the Mg^{++} (67). Mutagenesis of the aspartate codon to a Glu codon, and characterization of the substituted enzyme showed that this residue is not required for activation (67). The residues responsible for coordination of Mg^{++} have not been identified, but should be discernable from the three dimensional structure of the enzyme. Histidine 291 was also proposed to be the base responsible for proton extraction from the C-3 of $RbuP_2$ (68). The evidence supporting this assignment came from studies demonstrating that RuBisCO activity is sensitive to derivatization with histidine-specific modifying agents such as diethylpyrocarbonate (69). Protection against inactivation, which can be effected by competitive inhibitors which resemble $RbuP_2$, also prevents reaction of diethylpyrocarbonate with histidine 291 (68), implying that the residue is located in the active site near the $RbuP_2$ binding pocket. Any essential role for this residue, however, was disproven by substituting an alanine residue at this position by in vitro mutagenesis and demonstrating that catalysis was not impaired (70).

A missense mutation of residue Gly 171 to Asp identified this residue as essential for activity (71). Whether the residue functions in catalysis or is merely important for folding or structure is not known. Interestingly, mutations of Leu 290 and Thr 342, recovered as a temperature sensitive mutant and its intragenic

suppressor, have been reported to have altered substrate specificity (72,73).

Small subunits.

The function of the small subunits in the form I enzyme has been the subject of much discussion. Reports as early as 1970 suggested that the large subunits contained most or all of the residues associated with catalytic activity (11,74,75). Later, demonstration that the form II enzyme functions with only subunits similar to the large subunits of the form I enzyme suggested that the small subunits of the form I enzyme might play only a minor role. The complications the nuclear encoded small subunits introduce into RuBisCO biosynthesis, however, seemed, teleologically, to be justifiable only if they had, or conferred, a function which could not be accomplished by large subunits alone (58).

Early experiments to assign functions to individual subunits were hindered by problems with insolubility of dissociated subunits (75). Progress resumed when Andrews and Abel (76) found that they could reassociate dissociated small subunits with large subunit-core-complexes derived from some cyanobacterial RuBisCOs to reform holoenzyme. Andrews and Lorimer (77) used <u>in vitro</u> reassociation and subunits from various organisms to demonstrate that the relative substrate specificities (for CO_2 and O_2), which vary significantly between organisms, are determined solely by the large subunits.

In 1988 Andrews showed that large subunits do have detectable catalytic activity in the complete absence of small subunits (78). Although this confirmed that the small subunit is not directly

required in the catalytic mechanism, the level of activity in the large subunit cores was very low, namely less than 1% of the activity of holoenzymes (78). The restoration of full activity by *in vitro* addition of small subunits, however, also verified the claim that the small subunits do have an important role in RuBisCO activity (78). The mechanism by which the small subunits perform this role has not been demonstrated, but a likely explanation is that they cause a conformational shift in the large subunits, altering the arrangement of large subunit residues in the active site in such a way as to favorably affect catalysis (78).

In vitro mutagenesis and *in vitro* gene reconstruction, to make enzymes with single-residue-substituted or hybrid small subunits, have been used in attempts to identify individual residues or regions of the small subunit protein that are essential for enzyme function. Because of the lack of biochemical and, until recently, structural (three-dimensional) information on which to base hypotheses about function of individual residues, experiments have targeted conserved regions.

Wasmann *et al*. studied a conserved region near the middle of the small subunit sequence (79). They found that the region, which encompasses residues 49 to 54, confers the ability to assemble with pea large subunits in isolated pea chloroplasts (79). This region is not present in small subunits from cyanobacteria and represents the most striking divergence between cyanobacterial and higher-plant small subunits.

Voordouw *et al*. (80) mutated the codons for Trp 67 and Trp 70 in the *rbcS* gene of *Anacystis nidulans* and expressed the mutant genes

concommitantly with wild-type large subunits in Escherichia coli. According to most criteria, the substituted enzymes were indistinguishable from native RuBisCO, except that the $VmaxCO_2$ was reduced 2.5-fold (80). McFadden and Small (81) have reported on work in progress using a similar system. Preliminary analysis indicated that mutations in the conserved region of the small subunit between residues 12 and 22 were "deleterious to RuBisCo function and lead to a 39-94% reduction in Vmax", although the specific cause of this drop in activity was not determined (81). Neither of these experiments allowed the assignment of specific function to individual small subunit residues.

In general, the experiments to investigate structure-function relationships in the small subunit have been rather inconclusive. The absence, until recently, of a three-dimensional structure, and lack of data on the function of individual residues, have prevented the formulation of detailed hypotheses for experimental testing.

INTRODUCTION TO THE THESIS

As described above, the small subunit of type I RuBisCO, at least in some cases, is not absolutely required for enzymatic activity (77). Catalysis by large subunit octamers without small subunits, however, is very slow (77). This suggests that the small subunits in some way alter the conformation of the large subunits and

indirectly change the position of large subunit residues in the active site in a manner favorable for catalysis (77). The goal of the experiments described in this thesis was to better understand the nature and mechanism of this (and other possible) contribution(s) of the small subunit to RuBisCO function. Specifically, I postulated that there must be at least two requirements for the small subunit: to assemble with, and bind to, the large subunit, and to alter large subunit conformation in such a way as to effect catalysis. I wanted to identify individual residues in the small subunit which were important for one or the other (or both) of these properties and to understand on a structural level how these residues were functioning in the holoenzyme. I also hoped that if there were other requirements for the residues of the small subunit (e. g. a more direct role in catalysis) that they might be discovered if they could be disrupted by mutation.

My choice of experimental system reflected the realization that directed mutants can only (realistically) be produced <u>in vitro</u> and that transformation of organisms with mutant genes is of limited usefulness if a background of wild-type gene-product obscures the product of the introduced gene. Because gene replacement, which could theoretically be used to overcome this problem, is extremely inefficient in plants and because no eukaryotic mutants lacking RuBisCO small subunits have been isolated (probably due to multiple genes and/or conditional lethality of potential mutants) it was not possible to conduct the experiments in plants. Because none of the readily transformable cyanobacteria, in which gene replacement is efficient, have been shown to be able to grow without carboxylase

activity, effectively limiting the study of mutant RuBisCO in these organisms to mutants in which at least some activity remains, cyanobacteria were also deemed unsuitable for these experiments. The system eventually chosen and used in the experiments which are the culmination of this thesis, the expression of cyanobacterial RuBisCO genes in E. coli, proved advantageous in that the small subunit gene could be readily manipulated and the products studied without interference from an endogenous RuBisCO.

My initial experiments were directed toward expression of Zea mays RuBisCO large and small subunits, individually, in E. coli and partial characterization of the expression products. These experiments are presented in chapter two. In the experiments described in chapter three, E. coli were made to express both Z. mays subunits, simultaneously, and in combination with large and small subunits of the cyanobacterium Anabaena PCC 7120. Chapters four and five describe the creation of a collection of directed mutations in the Anabaena rbcS gene and analysis of the resulting altered small subunits when the mutant genes were expressed with wild-type Anabaena large subunits in E. coli.

REFERENCES

1. Wildman, S. G. & Bonner, J. (1947) Arch. Biochem. 14, 381-413.
2. Anderson, L. E., Price, G. B. & Fuller, R. C. (1968) Science 161, 482-484.

3. Tabita, F. R. & McFadden, B. A. (1974) J. Biol. Chem. 249, 3459-3464.

4. Gibson, J. L. & Tabita, F. R. (1977) J. Biol. Chem. 252, 943-949.

5. Gibson, J. L. & Tabita, F. R. (1977) J. Bacteriol. 132, 812-823.

6. Hartman, F. C., Stringer, C. D. & Lee, E. H. (1984) Arch. Biochem. Biophys. 232, 280-295.

7. Nargang, F., McIntosh, L. & Somerville, C. (1984) Mol. Gen. Genet. 193, 220-224.

8. Quivey, R. G. & Tabita, F. R. (1984) Gene 31, 91-101.

9. Schneider, G., Lindqvist, Y., Branden, C. I. & Lorimer, G. (1986) EMBO J. 5, 3409-3415.

10. Rutner, A. C. & Lane, M. D. (1967) Biochem. Biophys. Res. Comm. 28, 531-537.

11. Akazawa, T., Kondo, H., Shimazue, T., Nishimura, M. & Sugiyama, T. (1972) Biochemistry 11, 1298-1303.

12. McIntosh, L., Poulsen, C. & Bogorad, L. (1980) Nature 288, 556-560.

13. Bedbrook, J. R., Smith, S. M. & Ellis, R. J. (1980) Nature 287, 692-697.

14. Martin, P. G. (1979) Aust. J. Plant Physiol. 6, 401-408.

15. Amiri, I., Salnikow, J. & Vater, J. (1984) Biochim. Biophys. Acta 784, 116-123.

16. Chapman, M. S., Suh, S. W., Cascio, D., Smith, W. W. & Eisenberg, D. (1987) Nature 329, 354-356.

17. Chapman, M. S., Suh, S. W., Curmi, P. M. G., Cascio, D., Smith, W. W. & Eisenberg, D. S. (1988) Science 241, 71-74.

18. Andersson, I., Knight, S., Schneider, G., Lindqvist, Y., Lundqvist, T., Branden, C. I. & Lorimer, G. H. (1989) Nature 337, 229-234.

19. Knight, S., Andersson, I. & Branden, C. I. (1989) Science 244, 702-705.

20. Zurawski, G., Perrot, B., Bottomley, W. & Whitfeld, P. R. (1981) Nucleic Acids Res. 9, 3251-3270.

21. Curtis, S. E. & Haselkorn, R. (1983) Proc. Natl. Acad. Sci. USA 80, 1835-1839.

22. Andersen, K. & Caton, J. (1987) J. Bacteriol. 169, 4547-4558.

23. Nagy, F., Fluhr, R., Morelli, G., Kuhlemeier, C., Poulsen, C., Keith, B., Boutry, M. & Chua, N. H. (1986) Phil. Trans. R. Soc. Lond. B 313, 409-417.

24. Nierzwicki-Bauer, S. A., Curtis, S. E. & Haselkorn, R. (1984) Proc. Natl. Acad. Sci. USA 81, 5961-5965.

25. Goldschmidt-Clermont, M. & Rahire, M. (1986) J. Mol. Biol. 191, 421-432.

26. Chan, P. H. & Wildman, S. G. (1972) Biochem. Biophys. Acta 277, 677-680.

27. Coen, D. M., Bedbrook, J. R., Bogorad, L. & Rich, A. (1977) Proc. Natl. Acad. Sci. USA 74, 5487-5491.

28. Blair, G. E. & Ellis, R. J. (1973) Biochim. Biophys. Acta 319, 223-234.

29. Gooding, L. R., Roy, H. & Jagendorf, A. T. (1973) Arch. Biochem. Biophys. 159, 324-335.

30. Smith, S. M. & Ellis, R. J. (1981) J. Molecular. Applied Genet. 1, 127-137.

31. Nelson, T., Harpster, M. H., Mayfield, S. P. & Taylor, W. C. (1984) J. Cell Biol. 98, 558-564.

32. Langridge, P. (1981) FEBS Lett. 123, 85-89.

33. Mulligan, R. M., Houtz, R. L. & Tolbert, N. E. (1988) Proc. Natl. Acad. Sci. USA 85, 1513-1517.

34. Houtz, R. L., Stults, J. T., Mulligan, R. M. & Tolbert, N. E. (1989) Proc. Natl. Acad. Sci. USA, in press.

35. Highfield, P. E. & Ellis, R. J. (1978) Nature 271, 420-424.

36. Tobin, E. M. (1981) Plant Molecular Biol. 1, 35-51.

37. Dobberstein, B., Blobel, G. & Chua, N. H. (1977) Proc. Natl. Acad. Sci. USA 74, 1082-1085.

38. Gray, J. C. & Kekwick, R. G. O. (1974) Eur. J. Biochem. 44, 491-500.

39. Roy, H., Terenna, B. & Cheong, L. C. (1977) Plant Physiol. 60, 532-537.

40. Pfisterer, J., Lachmann, P. & Kloppstech, K. (1982) Eur. J. Biochem 126, 143-148.

41. Pain, D., Kanwar, Y. S. & Blobel, G. (1988) Nature 331, 232-237.

42. Robinson, C. & Ellis, R. J. (1984) Eur. J. Biochem. 142, 343-346.

43. Wasmann, C. C., Reiss, B., Bartlett, S. G. & Bohnert, H. J. (1986) Mol. Gen. Genet. 205, 446-453.

44. Kuntz, M., Simons, A., Schell, J. & Schreier, P. H. (1986) Mol. Gen. Genet. 205, 454-460.

45. Reiss, B., Wasmann, C. C. & Bohnert, H. J. (1987) Mol. Gen. Genet. 209, 116-121.

46. Wasmann, C. C., Reiss, B. & Bohnert, H. J. (1988) J. Biol. Chem. 263, 617-619.

47. Reiss, B., Wasmann, C. C., Schell, J. & Bohnert, H. J. (1989) Proc. Natl. Acad. Sci. USA 86, 886-890.

48. Ellis, R. J. (1987) Nature 328, 378-379.

49. Barraclough, R. & Ellis, R. J. (1980) Biochim. Biophys. Acta 608, 19-31.

50. Musgrove, J. E., Johnson, R. A. & Ellis, R. J. (1987) Eur. J. Biochem. 163, 529-534.

51. Cannon, S., Wang, P. & Roy, H. (1986) J. Cell Biol. 103, 1327-1335.

52. Hemmingsen, S. M., Woolford, C., Van der Vies, S. M., Tilly, K., Dennis, D. T., Georgopoulos, C. P., Hendrix, R. W. & Ellis, R. J. (1988) Nature 333, 330-334.

53. Ellis, R. J. & Hemmingsen, S. M. (1989) Trends. Biochem. Sci., in press.

54. Ellis, R. J. & Van der Vies, S. M. (1988) Photosynthesis Res. 16, 101-115.

55. Shinozaki, K. & Sugiura, M. (1983) Nucleic Acids Res. 11, 6957-6964.

56. Torres-Ruiz, J. A. & McFadden, B. A. (1988) Arch. Biochem. Biophys. 261, 196-204.

57. Goloubinoff, P., Gatenby, A. A. & Lorimer, G. H. (1989) Nature 337, 44-47.

58. Andrews, T. J. & Lorimer, G. H. (1987) Rubisco: Structure, Mechanisms, and Prospects for Improvement in: The Biochemistry of Plants v. 10, Stumpf, P. K. & Conn, E. E., eds., Academic Press, New York.

59. Lorimer, G. H. & Miziorko, H. M. (1980) Biochemistry 19, 5321-5328.

60. Lorimer, G. H. (1981) Biochemistry 20, 1236-1240.

61. Estelle, M., Hanks, J. F., McIntosh, L. & Somerville, C. (1985) J. Biol. Chem. 260, 9523-9526.

62. Fraij, B. & Hartman, F. C. (1983) Biochemistry 22, 1515-1520.

63. Terzaghi, B. E., Laing, W. A., Christeller, J. T., Petersen, G. B. & Hill, D. F. (1986) Biochem. J. 235, 839-846.

64. Hartman, F. C., Larimer, F. W., Mural, R. J., Machanoff, R. & Soper, T. S. (1987) Biochem. Biophys. Res. Comm. 145, 1158-1163.

65. Hartman, F. C., Milanez, S. & Lee, E. H. (1985) J. Biol. Chem. 260, 13968-13975.

66. Hartman, F. C., Soper, T. S., Niyogi, S. K., Mural, R. J., Foote, R. S., Mitra, S., Lee, E. H., Machanoff, R. & Larimer, F. W. (1987) J. Biol. Chem. 262, 3496-3501.

67. Gutteridge, S., Sigal, I., Thomas, B., Arentzen, R., Cordova, A. & Lorimer, G. (1984) EMBO J. 3, 2737-2743.

68. Igarashi, Y., McFadden, B. A. & El-Gul, T. (1985) Biochemistry 24, 3957-3962.

69. Saluja, A. K. & McFadden, B. A. (1982) Biochemistry 21, 89-95.

70. Niyogi, S. K., Foote, R. S., Mural, R. J., Larimer, F. W., Mitra, S., Soper, T. S., Machanoff, R. & Hartman, F. C. (1986) J. Biol. Chem. 261, 10087-10092.

71. Dron, M., Rahire, M., Rochaix, J. D. & Mets, L. (1983) Plasmid 9, 321-324.

72. Chen, Z., Chastain, C. J., Al-Abed, S. R., Chollet, R. & Spreitzer, R. J. (1988) Proc. Natl. Acad. Sci. USA 85, 4696-4699.

73. Chen, Z. & Spreitzer, R. J. (1989) J. Biol. Chem. 264, 3051-3053.

74. Sugiyama, T., Matsumoto, C. & Akazawa, T. (1970) J. Biochem. 68, 821-831.

75. Nishimura, M., Takabe, T., Sugiyama, T. & Akazawa, T. (1973) J. Biochem. 74, 945-954.

76. Andrews, T. J. & Abel, K. M. (1981) J. Biol. Chem. 256, 8445-8451.

77. Andrews, T. J. & Lorimer, G. H. (1985) J. Biol. Chem. 260, 4632-4636.

78. Andrews, T. J. (1988) J. Biol. Chem. 263, 12213-12219.

79. Wasmann, C. C., Ramage, R. T., Bohnert, H. J. & Ostrem, J. A. (1989) Proc. Natl. Acad. Sci. USA 86, 1198-1202.

80. Voordouw, G., deVries, P. A., van den Berg, W. A. M. & deClerck, E. P. J. (1987) Eur. J. Biochem. 163, 591-598.

81. McFadden, B. A. & Small, C. L. (1988) Photosynthesis Res. 18, 245-260.

Chapter 2

EXPRESSION OF *Z. mays* RuBisCO SUBUNITS IN *E. coli*

INTRODUCTION

Foreign proteins expressed in *E. coli* have included proteins of prokaryotic and eukaryotic (both nuclear and organellar) origin. They have ranged from proteins consisting of a single, low molecular weight subunit (1), to large proteins containing multiple, heterogeneous subunits (2). Levels of expression of the foreign proteins have ranged from barely detectable to about 50% of the total bacterial protein (3).

Several general characteristics are required for plasmids to be useful for expression of foreign gene products in *E. coli*. An expression plasmid must have a promoter recognized by *E. coli* RNA polymerase in the correct orientation for transcription of the coding sequence of the foreign gene. If high levels of expression are desired, it is necessary to use a strong promotor. It may also be helpful if expression from the promotor can be experimentally manipulated, as constitutive high levels of expression of foreign gene products may be deleterious (4). The plasmid must also contain a recognition site for ribosome binding and translation initiation at

the 5' end of the protein coding region of the message. This ribosome binding site should include a Shine-Dalgarno sequence (5) complementary to the 16s RNA of the ribosome and should be located 5 bp to 9 bp upstream of the AUG initiation codon (6). Additionally, the coding region used for the plasmid must be uninterrupted. This necessitates the use of a synthetic gene or cDNA if the original foreign gene contained introns.

As an initial step in establishing a system for genetic analysis of RuBisCO, I constructed plasmids directing the expression of the subunits of *Zea mays* RuBisCO in *E. coli*. This chapter contains the details of the construction of these plasmids and the results of preliminary characterization of the expression products. The conclusions from my experiments are essentially similar to those from other, published work and will be discussed in this context.

MATERIALS AND METHODS

Bacterial Strains and Plasmids.

The *E. coli* strains used in the experiments described in this chapter were HB101 (F- *pro leu thi lacY hsdR andA recA rpsL ara galK xyl mtl supE44*) and JM83 [F (*lac-pro*) *ara rpsL* 80d*lacZM15*]. Plasmids pUC9, pUC19, pBR322, pCQV2, and pZmc37 have been described (7, 8, 9, 10, 11). Plasmid pBT76-2 was a gift from W. Taylor (CSIRO, Canberra). Plasmids p71a, pBC12, p855, and p77 are described under "results".

Antisera

Rabbit antiserum raised against tomato RuBisCO holoenzyme was a gift from Barbara Wilson (Biochemistry Department, MSU). Rabbit antiserum raised against Zea mays small subunit was a gift from Erin Bell (MSU-DOE-PRL, MSU)

DNA Manipulation and Analysis

Plasmid DNA preparations, restriction digests, electrophoretic separations, ligations, transformations, and other DNA manipulations were performed according to standard procedures (12). DNA sequence analysis was carried out using the chemical cleavage method of Maxam and Gilbert (13).

Electron Microscopy

Fixation, embedding, and electron microscopy of bacteria carrying expression plasmids was according to standard protocols (14, 15). Embedding, sectioning, and photographing of images was performed by Jon Kemp (Cornell University).

Gel Electrophoresis

Proteins were resolved in 10% - 17.5% acrylamide gradient SDS-polyacrylamide gels using the buffer system described by Laemmli (16). Samples were suspended in SDS-sample buffer to give final concentrations of 10% glycerol, 2% SDS, 2% β-mercaptoethanol, and 62mM Tris-HCl pH 6.8, with bromophenol blue included as a tracking dye. Samples were boiled for two minutes prior to loading.

Western Blotting

Transfer of proteins from SDS-PAGE gels to nitrocellulose filters was as described (17). Immunodetection with antiserum and protein A-alkaline phosphatase was performed according to standard procedures (18).

Preparation of Bacterial Extracts

Cultures of E. coli expressing RuBisCO subunits were usually grown to stationary phase at 37°C in L-broth (19) without glucose with vigorous aeration and harvested by centrifugation. In the case of cultures bearing plasmids with the temperature sensitive CI857 allele (10), growth was at 30°C to mid-log phase and induction was at 42°C. Pelleted bacteria were resuspended in assay buffer (20) (typically about twice the volume of the pellet) and disrupted by ultrasonic treatment as described (20). The lysate was centrifuged at 30,000 x g for 15 minutes. When the pellet fraction was to be analyzed, it was resuspended in a volume of assay buffer equal to the volume of the supernatant fraction. When the soluble fraction was to be assayed for carboxylase activity, it was dialyzed against 500 volumes of assay buffer for twelve hours, beforehand.

Carboxylase Assay

The assay used measured Rbup2-dependent incorporation of radioactivity from 14CO2 into acid-stable forms (20 based on 21).

Purification of Protein Granules

Cultures of cells expressing high levels of large or small subunit protein were grown to late stationary phase , harvested by

centrifugation, and resuspended in half the original volume of 1/2X STET (1X STET is 8% sucrose, 5% Triton X-100, 50 mM EDTA, 50 mM Tris-HCl pH 8). Lysozyme was added to a concentration of 300 ug/ml and the mixture was incubated at 37 C for 15 minutes. Granules were pelleted by centrifugation at 650 x g for three minutes, washed once with 1/4 original volume of 1/2X STET and washed several times with 10 mM Tris-HCl pH 8.0.

Isolation of RuBisCO from *Zea mays*

RuBisCO was isolated from *Z. mays* seedlings using the procedure of Hall and Tolbert (22).

RESULTS

Plasmids Directing Expression of Large Subunits

The *Zea mays* gene (*rbcL*) for the large subunit of RuBisCO was obtained from the plasmid pZmc37 which contains BamHI fragment 9 of the *Zea mays* chloroplast genome (11, 23), encompassing the entire gene for the large subunit (*rbcL*) as well as the genes for the β and ϵ subunits of coupling factor F . A plasmid (p71a) in which the *lac* promotor is upstream of the rbcL coding sequence was made by subcloning a HincII-BamHI fragment from pZmc37 into pUC9. The HincII-BamHI fragment includes about 300 bp upstream of the *rbcL* coding sequence and about 1.2 kbp between the 3' end of the coding region and the BamHI recognition site. Because the *lac* promotor and the *rbcL* coding region were in an orientation and arrangement which

could theoretically give *lac* promoted transcription, I tried to detect expression of the large subunit by comparing the protein profiles of transformed and control cells and by assaying for catalytic activity. No prominent band corresponding to the large subunit was observed when proteins in extracts prepared from *E. coli* (strain JM83) harboring this clone were resolved by SDS-PAGE, stained with coomassie blue, and compared to similarly resolved proteins from control cultures. Additionally, no RbuP2 carboxylase activity could be detected in the extracts.

It was hypothesized that some feature of the sequence between the *lac* promotor and the *rbcL* coding region could be interfering with transcription or translation, so a series of clones were made in which various portions of this region had been deleted. The plasmid p71a was linearized by cleavage with the restriction enzyme *HincII*, for which the only recognition site is between the *lac* promotor and the *rbcL* structural gene. Bidirectional deletions of varying length were made by limited digestion of the linearized plasmid with the double-strand-specific exonuclease *Bal-31*. Aliquots of the exonuclease digestion were removed and digestion terminated at various times to ensure a range of deletion sizes. The aliquots were pooled and ligated without any additional treatment to increase the proportion of blunt ends.

Randomly selected ampicillin-resistant colonies from the transformation of this mixture of deletion plasmids were streaked in individual patches and allowed to grow into confluent masses of colonies. A sample of cells from each patch was resuspended in SDS-PAGE sample buffer and lysed by boiling for 2 min. Proteins in

the lysate were resolved by SDS-PAGE and the resulting gels were inspected visually after Coomassie-blue staining. Lysates from several clones were found to contain an abundant 55 kDa protein. An example of an SDS-PAGE gel used for the initial screen showing several clones with elevated levels of 55 kilodalton protein is shown in Figure 1.

Of the clones screened, the one with the highest levels of expression of 55 kilodalton protein was designated pBC12. The sequence of the region between the lac promotor and the beginning of the rbcL coding region was determined and is shown in Figure 2 with a reconstruction of the extent of the Bal31 digestion that created the deletion. Analysis of the sequence of pBC12 showed that the fusion of lac and rbcL after exonuclease treatment occurred in such a way as to create a near perfect consensus Shine-Dalgarno (5) sequence. This fortuitous arrangement assured that large subunit translation would begin at the appropriate methionine and that the resulting product would not be a fusion protein.

Extracts prepared by ultrasonic treatment of bacteria carrying pBC12 were assayed for RbuP2 carboxylase activity. The assay used measured RbuP2-dependent incorporation of radioactivity from 14CO2 into acid insoluble material and had a minimal level of detection of about 5 pmol CO2 fixed/min/mg total protein. No activity could be detected in any fractions (crude, 30,000 x g supernatant, or 30,000 x g pellet) of the bacterial extracts (data not shown).

Because no activity could be detected, two other lines of evidence were developed to establish that the abundant 55 kilodalton protein in the E. coli extracts was large subunit protein.

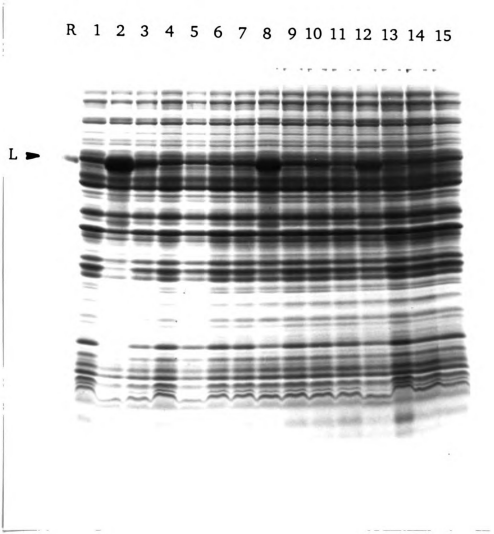


Figure 1. Example of Coomassie-stained SDS-PAGE gel used for screening clones for expression of large subunit. The sample in the leftmost lane (R) was maize RuBisCO. Lane 1 was loaded with extract from control cells bearing pUC9. Lane 2 was loaded with extracts from cells bearing pBC12. Lanes 3 - 15 were loaded with extracts from selected clones from the initial screen.

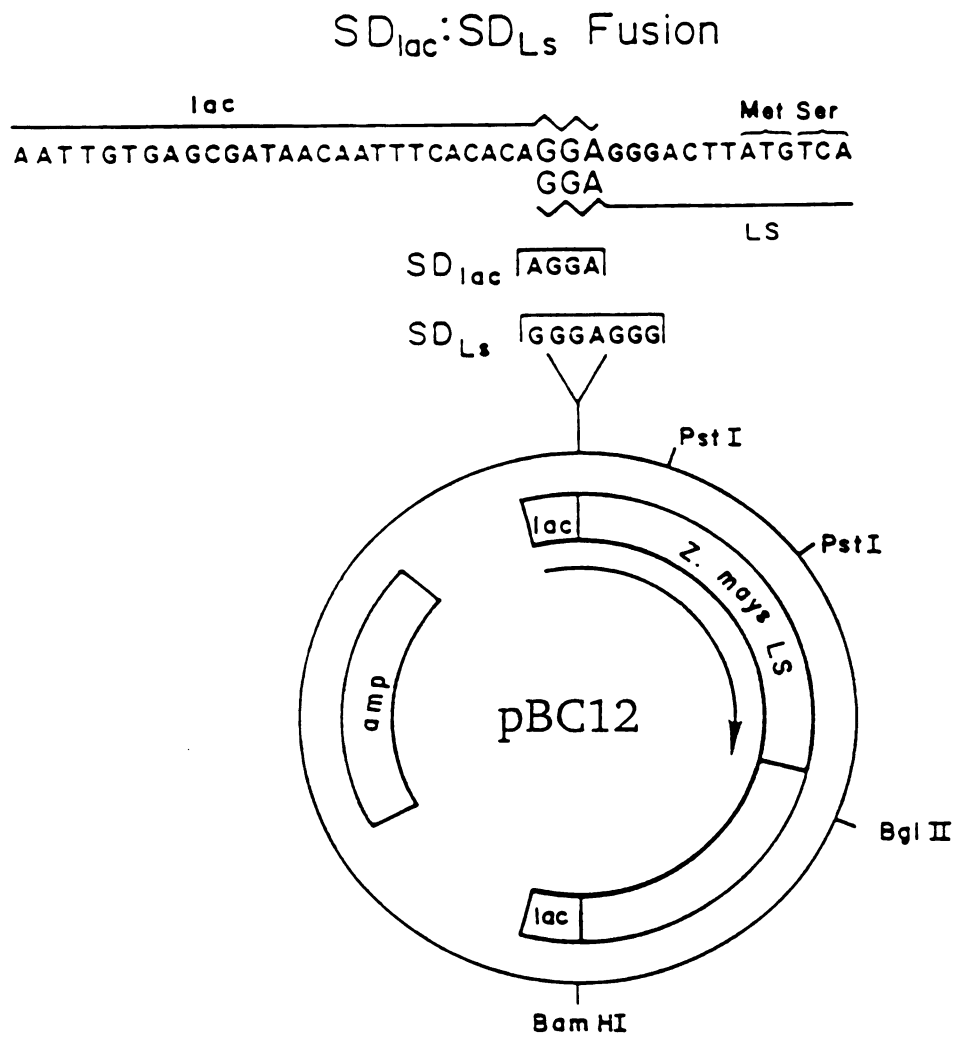


Figure 2. Expression plasmid pBC12. This construction, made as described in the text, resulted in a hybrid Shine-Dalgarno sequence for translation initiation. The fusion site is at one of the three bases indicated, but cannot be distinguished because they are identical in both the bacterial and plant sequences. (Figure modified from reference 24)

Immunological identity was shown after resolution by SDS-PAGE and transfer to nitrocellulose, using rabbit antiserum containing antibodies raised against purified tomato RuBisCO. The immunoblot from this experiment is shown in Figure 3. With extended color development, many bands became visible, but by far the most intense reaction corresponded to the 55 kilodalton protein. Direct proof that the abundant 55 kilodalton protein, rather than a minor comigrating species, was the large subunit was provided by the observation that the apparent molecular weight of the abundant band decreased when part of the rbcL coding region was deleted. The deletion was made by removing a 597 bp PstI fragment from the rbcL coding region, maintaining the original reading frame. The predicted Mr of the product of the remaining coding region is 32.5 kilodaltons. The observed Mr of the new abundant band in the gel which is shown in Figure 4 is 34 kDa. The difference between these two values is well within the observed deviation from linearity for migration of proteins in SDS-PAGE (25).

Because the large subunit protein synthesized in E. coli lacked catalytic activity, the RuBisCO properties which could be studied with the system were limited to characterization of the structure of the bacterially produced protein and comparison with the protein from higher plants. E. coli (strain JM83, constitutive for lac expression) bearing pBC12 were disrupted by ultrasonic treatment. Lysates were fractionated into "supernatant" and "pellet" fractions after centrifugation at 30,000 x g for 10 min. Fractions were resolved by SDS-PAGE. The bulk of the 55 kDa protein was in an insoluble or rapidly sedimenting fraction (Figure 5). This

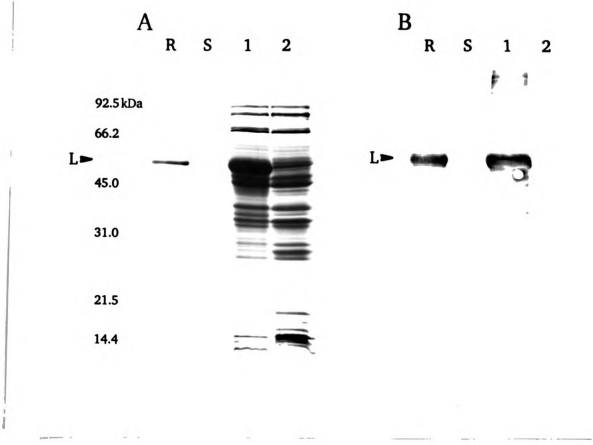


Figure 3. Expression of large subunit by cultures bearing expression plasmid pBC12. (A) A Coomassie-stained gel. Lanes were loaded as follows: R, maize RuBisCO; S, molecular weight standards; 1, extract of cells bearing pBC12; 2, extract of control cells bearing pUC18. (B) A western blot of a gel identical to (A) developed with antibodies against RuBisCO.

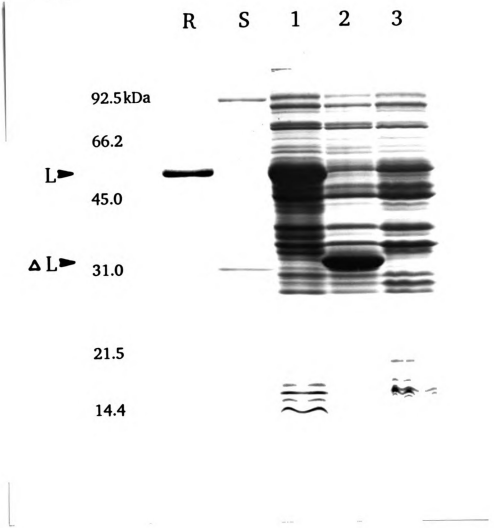


Figure 4. Coomassie-stained SDS-PAGE gel showing lower molecular weight of prominant band in cultures bearing expression plasmid with in-frame deletion. Lanes were loaded with: R, maize RuBisCO; S, molecular weight standards; 1, extract of cells bearing pBC12; 2, extract of cells bearing the expression plasmid with the in-frame deletion; 3, extract of control cells bearing the vector pUC18.

Figure 5. Fractionation of extracts of cells expressing large subunits. (A) A coomassie stained gel. The lanes were loaded with samples as follows: R, Maize RuBisCO; S, molecular weight standards; 1, supernatant fraction after centrifugation of cell lysates from cultures bearing expression plasmid pBC12; 2, pellet fraction from above centrifugation; 3, supernatant fraction from control lysates; 4, pellet fraction from control lysates. (B) A western blot of a gel identical to the one shown in (A). The blot was developed with antibody to tomato RuBisCO.

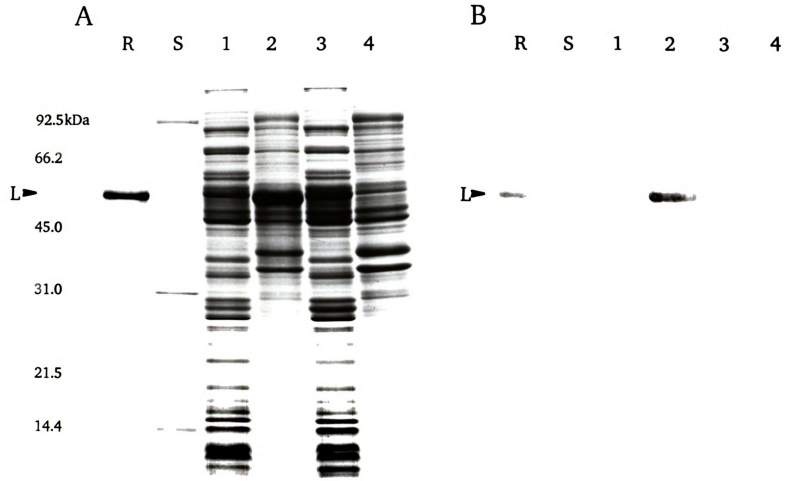


Figure 5

localization was consistent in extracts of bacteria transformed with the plasmids directing lower levels of expression (only partially characterized, these plasmids were identified in the screen described above).

Electron microscopy of thin sections prepared from pBC12-bearing cells provided further information on the nature of the large subunit protein as synthesized in E. coli. Large (about 1 um diameter) electron dense granules were observed in many cells. These granules were located terminally (Figure 6) and were not present in non-transformed E. coli. These granules could be isolated by treating the bacteria with lysozyme and 0.1% Triton X-100 to degrade cell walls and membranes, and centrifugation to remove the granules from the lysate. The granules could then be dissolved by bringing the suspension to 8 molar urea, or by boiling in a buffer containing 2% SDS. Considering that this preparation may have contained wall debris or other subcellular structures, the purity of the resulting fraction (shown in Figure 7) suggests that the predominant, or perhaps even the sole, component of these granules is the large subunit.

The Mr of the large subunit protein was compared to the Mr of authentic large subunit isolated from Z. mays seedlings to see if I could detect a difference in molecular weight which could result from in vivo processing of the large subunit in plants. No difference in mobility could be detected when large subunit protein purified as above was electrophoresed in lanes adjacent to, or mixed with, Z. mays RuBisCO. A large-format SDS-PAGE gel showing such a comparison is shown in Figure 8.

Figure 6. Electron micrograph of RuBisCO granule in cells of E. coli bearing expression plasmid pBC12. (Figure reprinted from reference 26)

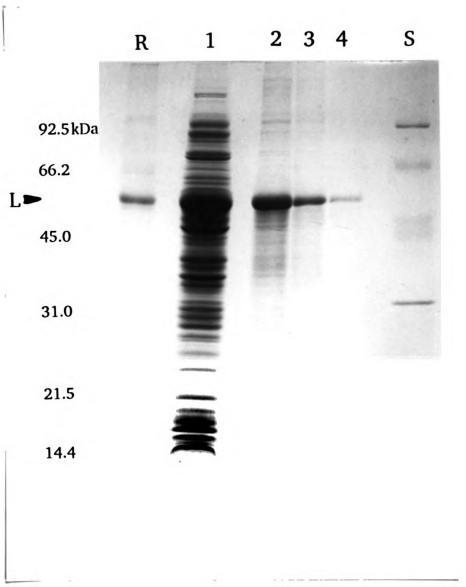


Figure 7. Coomassie-stained SDS-PAGE gel showing purity of protein in RuBisCO granules. Proteins were resolved by SDS-PAGE and detected by staining with coomassie blue. Lanes are: R, maize RuBisCO; 1, total protein extracted from cells bearing expression plasmid pBC12; 2, 3 and 4 , partially purified granules; S, molecular weight markers.

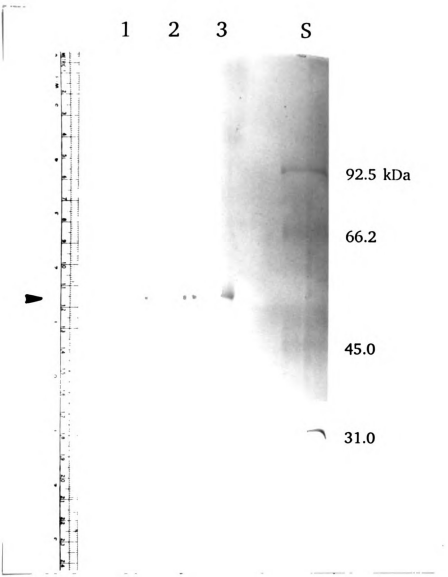


Figure 8. Comparison of gel-mobilities of authentic RuBisCO large subunit and the large subunit protein synthesized in E. coli. Lanes are: 1, maize RuBisCO; 2, equimolar mixture of authentic and bacterially-synthesized large subunits; 3, bacterially synthesized large subunits; S, molecular weight markers. Gel was stained with Coomassie blue.

Plasmids Directing Expression of Small Subunits

Sequence encoding Z. mays RuBisCO small subunit. A cDNA clone (pBT76-2) corresponding to one of the Z. mays genes for the small subunit of RuBisCO was obtained from Dr. W. Taylor (CSIRO, Canberra). The sequence of the cDNA insert was determined by the method of Maxam and Gilbert (13). The sequence is shown with the deduced amino acid sequence in Figure 9. The clone did not include the region corresponding to the 5' end of the message, encoding part of the transit peptide, but the region encoding the entire mature protein was represented in the clone. The deduced sequence of the part of the protein represented in the clone was found to be identical to the deduced sequence from a Z. mays cDNA sequence reported by Lebrun et al. (27), except for a leucine residue which, on the basis of comparison with other known small subunit sequences, was probably overlooked in their sequencing.

Plasmid directing expression of a protein identical to mature (processed) small subunit. A recognition site for the restriction endonuclease SphI was found at precisely the location in the rbcS gene corresponding to the site in the polypeptide at which proteolytic processing occurs to generate the mature small subunit (see Figure 9). Cleavage at this restriction site allowed isolation of a fragment containing only sequence coding for the amino acids of the mature small subunit. This fragment was subcloned into the expression vector pCQV2, as shown in Figure 10a. The spacing between the λ PR promotor and ribosome binding site of the vector and the

```
                                        30                                  60
                  Val Ser Asn Gly Gly Arg Ile Arg Cys Met Gln Val Trp Pro Ala Tyr
GGG GGG GGG GGG GTC AGC AAC GGC GGA AGG ATC CGG TGC ATG CAG GTG TGG CCG GCC TAC

                                        90                                 120
Gly Asn Lys Lys Phe Glu Thr Leu Ser Tyr Leu Pro Pro Leu Ser Thr Asp Asp Leu Leu
GGC AAC AAG AAG TTC GAG ACG CTG TCG TAC CTG CCG CCG CTG TCG ACG GAC GAC CTG CTG

                                       150                                 180
Lys Gln Val Asp Tyr Leu Leu Arg Asn Gly Trp Ile Pro Cys Leu Glu Phe Ser Lys Val
AAG CAG GTG GAC TAC CTG CTG CGC AAC GGC TGG ATA CCC TGC CTC GAG TTC AGC AAG GTC

                                       210                                 240
Gly Phe Val Tyr Arg Glu Asn Ser Thr Ser Pro Cys Tyr Tyr Asp Gly Arg Tyr Trp Thr
GGC TTC GTG TAC CGC GAG AAC TCC ACC TCC CCG TGC TAC TAC GAC GGC CGC TAC TGG ACC

                                       270                                 300
Met Trp Lys Leu Pro Met Phe Gly Cys Asn Asp Ala Thr Gln Val Tyr Lys Glu Leu Gln
ATG TGG AAG CTG CCC ATG TTC GGC TGC AAC GAC GCC ACC CAG GTG TAC AAG GAG CTG CAG

                                       330                                 360
Glu Ala Ile Lys Ser Tyr Pro Asp Ala Phe His Arg Val Ile Gly Phe Asp Asn Ile Lys
GAG GCC ATC AAA TCC TAC CCG GAC GCC TTC CAC CGC GTC ATC GGC TTC GAC AAC ATC AAG

                                       390                                 420
Gln Thr Gln Cys Val Ser Phe Ile Ala Tyr Lys Pro Pro Gly Ser Asp  *
CAG ACG CAG TGC GTC AGC TTC ATC GCC TAC AAG CCC CCG GGC AGC GAC TAG ACC GCG CCC

                                       450                                 480
GCC GGC CGC CCC CCG CCG GCT AGC TAT AGC TCC TGC GTG AGC TAG TAG CTA GCT AGT GCC

                                       510                                 540
ATG CGT CGT CTC TGT CGT TCG GTT TTG CTT CGG GGT CAC CGT ACC CTT TGC TTG CTT GGT

                                       570                                 600
TTC TTC TTT CCT TTT TCC TTT TTT TTT TCC TTT CCC CGG CCA TGG TTC CTT TGC TTT CCA

                                       630                                 660
GTT TTC TCT GCT GGA TGT GAT GTA TCC ATT GTT GCA AGC ATG GCC TTG CAT TGG CTA CCT

                                       690                                 720
CTA TAC CTG CTG CTA CAA AAC TAC TGC AAC GCC TAT ATA TAC TTG GGG TGA GGA ACA TGT

                                       750                                 780
GAA TGC AAG CTC CGG CTA TCA TAT ACA TGT AAT ATG GAC TAC AAA ACT ATA TAT ATA AAT

                                       810                                 840
CCG CCG AGG CGC CGA CAA TAC TAT AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA
```

Figure 9. The sequence of the cDNA insert of clone pBT76-2. The twelve G residues at the 5' end are the result of oligo-dG tailing of the cDNA in the cloning procedure and are not derived from the mRNA. The site corresponding to the final proteolytic cleavage is indicated (▼). The SphI restriction enzyme recognition site used for construction of expression plasmids is underlined.

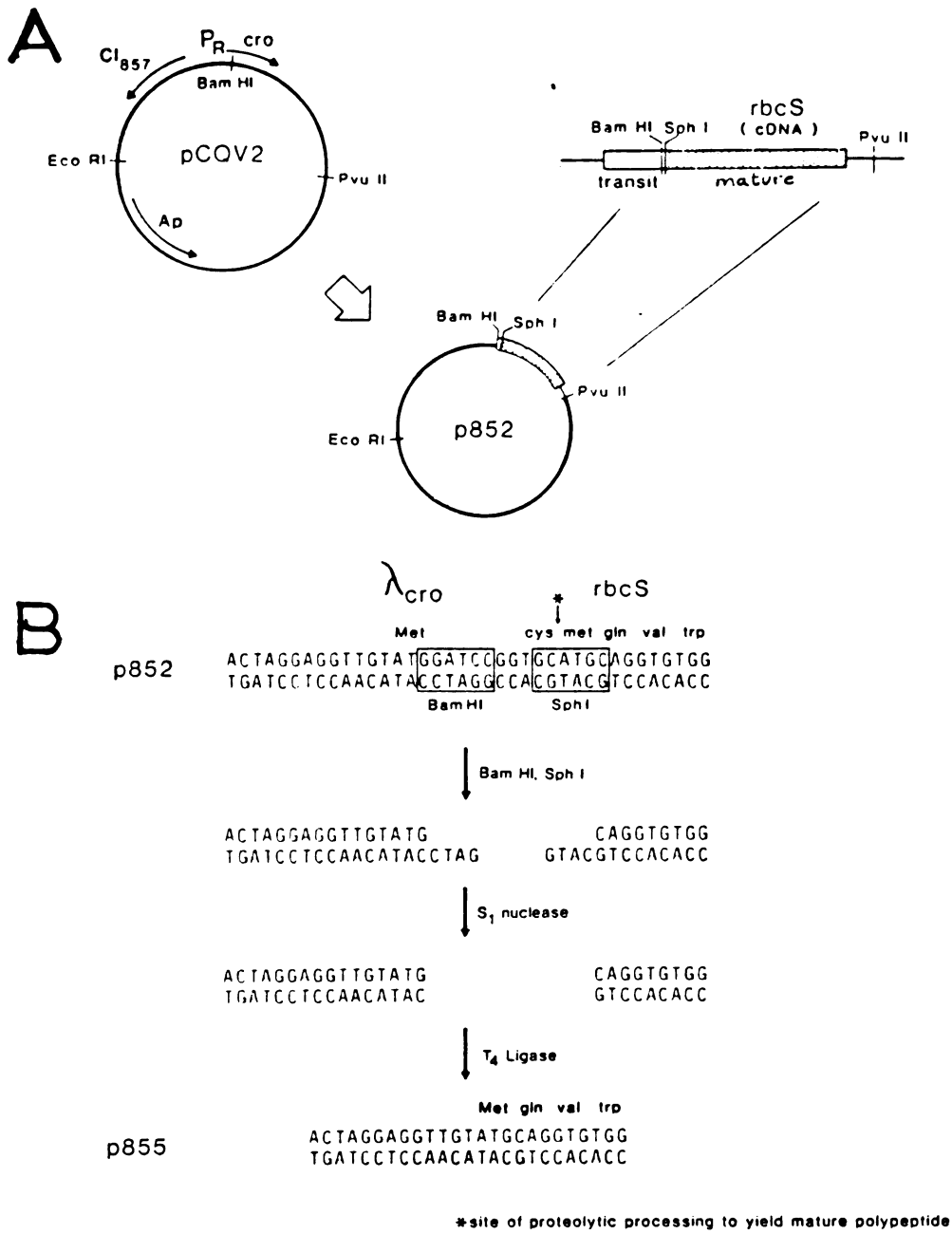


Figure 10. Construction of expression plasmid p855. The details of the construction are given in the text. (A) diagrams the subcloning of a fragment of the small subunit gene encoding only the residues of the mature protein into an expression vector. (B) shows how the nucleotide spacing immediately upstream of the rbcS coding sequence was optimized for expression.

rbcS coding sequence was manipulated by cleavage with BamHI and SphI and treatment with single-strand-specific nuclease as shown in Figure 10b. In the resulting plasmid, p855, the spacing was such that translation must initiate at the AUG codon for the amino-terminal methionine of authentic mature small subunit. Transcription could be controlled experimentally because the plasmid also encodes the temperature sensitive CI857 allele of the lambda repressor CI (10).

Extracts of E. coli bearing p855 were analyzed by SDS-PAGE. No differences in patterns of coomassie staining of resolved proteins from p855-bearing and control extracts could be detected (Figure 11a). Immunodetection of nitrocellulose replicas of the gels using an antiserum raised against Z. mays small subunit, gave a weak reaction at the predicted molecular weight (13 kDa)(Figure 11b). This reaction was detected in lanes derived from the pellet fraction after centrifugation of extracts at 30,000 x g for 15 min. No reaction was detected in lanes derived from the supernatant fraction.

Plasmid directing the expression of a fusion protein similar to the small subunit. The level of small subunit protein synthesized in cells bearing p855 was lower than I had expected and it was not clear whether they would be sufficient for the experiments planned. Because of my earlier success with expression of the large subunit from the lac promotor, I designed and constructed an rbcS-expressing plasmid in which transcriptional control and translation initiation would use lac rather than bacteriophage lambda signals. A derivative of pUC19 was made by linearizing pUC19 with HindIII and removing the resulting four unpaired bases at each end with exonuclease S1, and

Figure 11. Expression of small subunit by cultures bearing p855. (A) A coomassie-stained SDS-PAGE gel comparing extracts of p855-bearing and control cells. The lanes were loaded with: R, maize RuBisCO; S, molecular weight standards; 1, total cellular protein extracted from control cultures bearing the vector pUC18; 2, total cellular protein extracted from cultures bearing p855. (B) A western blot of a gel similar to the one shown in (A), developed with antiserum recognizing maize small subunit.

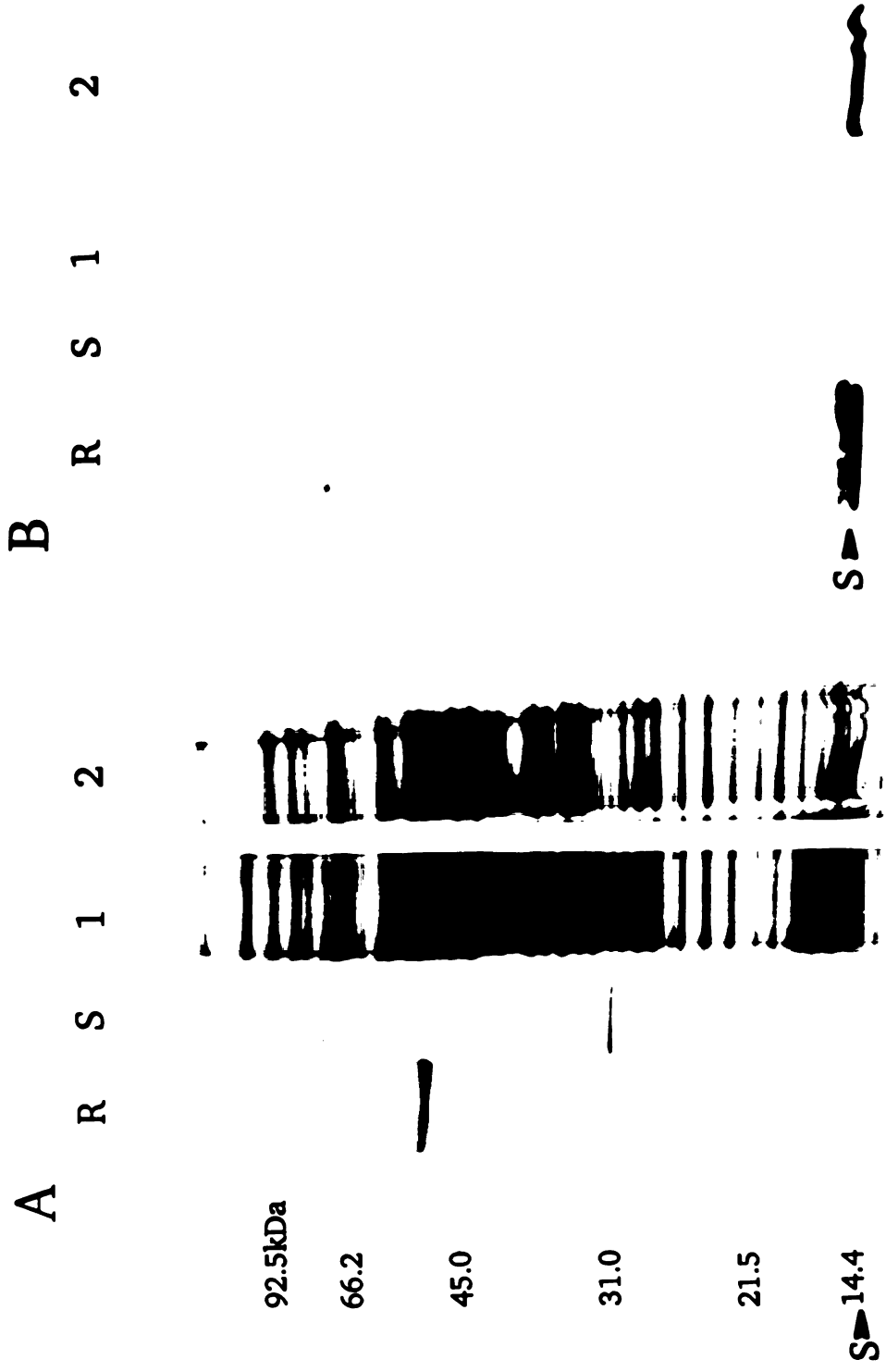


Figure 11

religation, thereby bringing the SphI site in the "multiple cloning site" into the same translational reading frame as the SphI site in the rbcS sequence. The rbcS coding region and part of the 3' untranslated region were isolated as an SphI PvuII fragment of pBT76-2 and subcloned into the modified vector. The predicted protein product of this plasmid would contain seven amino acids derived from lacZ appended to the amino terminus of the mature small subunit.

Cells bearing this lac-based expression plasmid, p77, contained much higher levels of small subunit protein than cells bearing the lambda-based expression plasmid p857 described above. A Coomassie-blue stained SDS-PAGE gel of extracts form E. coli HB101 is shown in Figure 12. An intensely stained band at 13.5 kilodaltons is present in extracts of p77 but not control extracts. Fractionation into supernatant and pellet fractions after centrifugation of ultrasonically disrupted cells at 30,000 x g for 15 min showed that the abundant protein was localized in the rapidly sedimenting fraction (Figure 13). The band was shown to be small subunit protein, as expected, by western blotting and immunodetection with small-subunit-specific polyclonal antiserum. An immunoreactive band detected at 18 kDa is probably due to read through of the stop codon at the end of rbcS, as the E. coli strain used in this experiment (HB101) carries the suppressor mutation SupE.

The rapidly-sedimenting nature of the small subunit fusion protein produced by cells bearing p77 suggested that by dissolving walls and membranes and centrifugation at low speed, as was described earlier for partial purification of the large subunit produced from

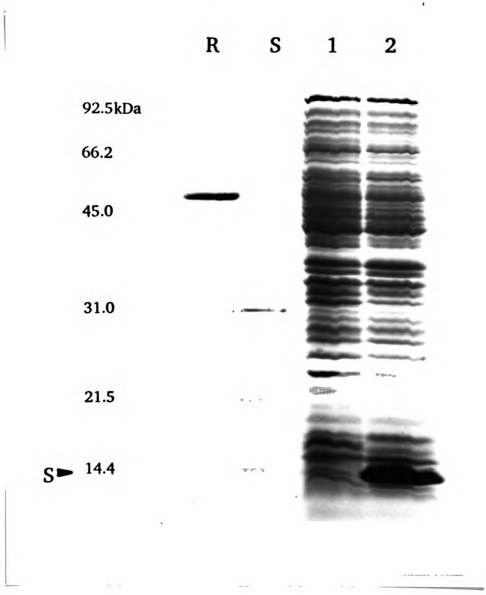


Figure 12. Expression of small subunit protein in cells bearing p77. Proteins were resolved on an SDS-PAGE gel and stained with coomassie blue. Lane 1 was loaded with total cellular protein from a control culture. Lane 2 was loaded with total cellular protein from cells of a culture bearing p77. Lane R was loaded with maize RuBisCO. Molecular weight standards were loaded in S.

Figure 13. Fractionation of lysates prepared from cultures bearing p77. (A) A coomassie stained SDS-PAGE gel of the different fractions. The lanes are: R, maize RuBisCO; S, molecular weight standards; 1, supernatant fraction after centrifugation of cell lysate; 2, pellet fraction after centrifugation; 3 and 4, supernatant and pellet fractions of lysates from a control culture bearing pUC18. (B) A western blot of a gel identical th the one shown in (A). The blot was developed with antibodies raised against maize RuBisCO small subunit. The higher molecular weight reacting material in lane R is probably small subunit still associated with large subunit. The unexpected bands in lane 2 are discussed in the text.

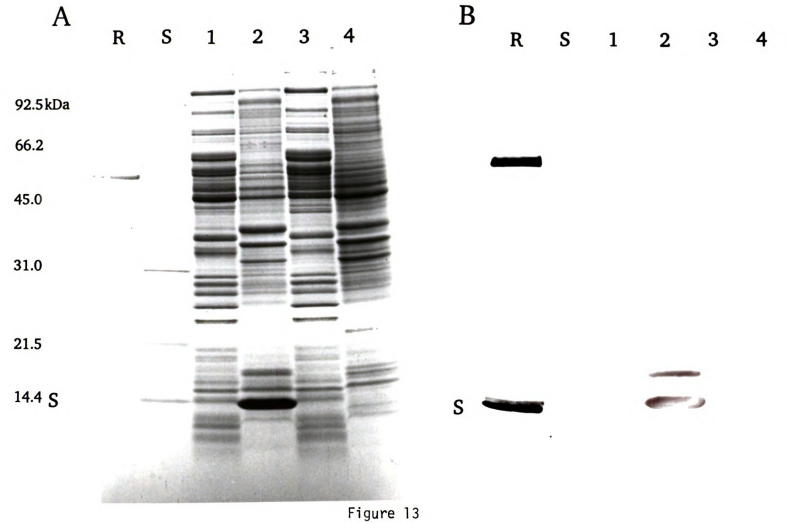


Figure 13

pBC12, could provide an easy procedure for partial purification. An SDS-PAGE gel of the initial extract and the end product of this simple purification is shown in Figure 14.

DISCUSSION

The plasmids described in this chapter were designed to direct the expression of individual subunits of Zea mays RuBisCO in E. coli. The level of RuBisCO subunit was analyzed in cultures carrying each of the plasmids to evaluate their efficacy as expression plasmids. Extracts of large subunit-expressing and small subunit-expressing cultures were prepared to allow rudimentary characterization of the expression products.

Large subunit protein accounted for 20-30% of the total cellular protein in E. coli carrying pBC12. This high level may have been due to the construction of a consensus Shine-Dalgarno sequence (5) and its perfect spacing with respect to the initiation codon of rbcL (6). Other nucleotides surrounding the translational start are also characteristic of those found in highly expressed genes (28). Additionally, the accumulation of the high levels of large subunit found in the bacteria may have been due, in part, to its observed insolubility. Aggregation into granules would be expected to minimize interference with cellular processes, and could cause the protein to be inaccessible to intracellular proteases. Other plasmids expressing large subunits of Zea mays and wheat RuBisCO have been reported by Gatenby et al. (29, 30). These plasmids are capable of directing synthesis of large subunit protein at levels up to 2% of

Figu

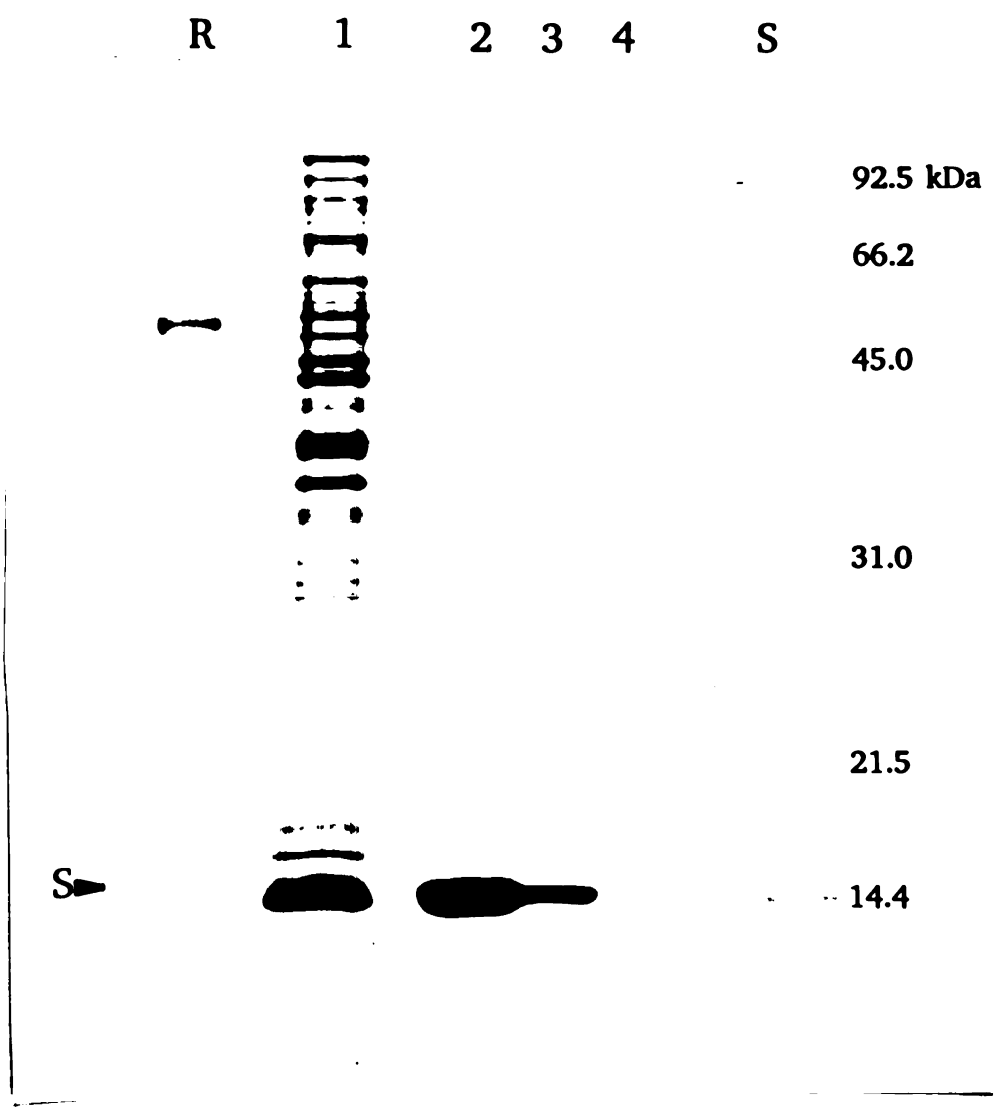


Figure 14. Coomassie blue-stained SDS-PAGE gel showing the purity of small subunit-fusion protein in granules isolated from E. coli bearing expression plasmid p77. The lanes were loaded with: R, maize RuBisCO; 1, total protein extracted from cells bearing p77; 2, 3 and 4, partially purified granules from culture used for lane 1; S, molecular weight standards.

the total bacterial protein (31).

The levels of expression in cultures carrying the fusion-small subunit-expressing plasmid were much higher than levels in cultures carrying the non-fusion small subunit-expressing plasmid. The reason for the higher level of expression from the fusion-small subunit expression plasmid is not clear from my experiments because there were several differences between it and the non-fusion-small subunit expression plasmid. That the promotor in p77 was E. coli lac while the promotor in p855 was λ PR, was probably not responsible for the difference, as both promotors are capable of directing high levels of transcription. It is unlikely that lower levels in cells bearing p855 were a result of the higher temperature at which the cells were grown (to fully induce PR) because no more small subunit was detected in extracts from cells grown at 37°C, a temperature at which substantial transcription of PR occurs in CI857 backgrounds. A more likely explanation for the difference in levels of expression is the that only one of the proteins is synthesized as a fusion protein. Examples have been reported of fusion proteins accumulating to higher levels than non-fusion proteins, when expressed in E. coli, including cases in which the fusion represents the addition of only a few amino acids (32) as is the case here. A plasmid directing expression of a fusion protein consisting primarily of wheat small subunit sequence has been described (33). The fusion protein product has ten amino acids derived from B-galactosidase at the amino terminus. Accumulation was reported to be up to 2% of the total bacterial protein (31).

The insolubility of the expression products of these plasmids

appears to be a common phenomenon for foreign proteins synthesized in E. coli. Many proteins, some of them quite soluble in the tissue in which they are normally found, accumulate in large granules similar to the ones observed for the large subunit when expressed at high levels in E. coli (34, 35). These granules often contain almost exclusively the foreign protein (36) as was observed for granules from large subunit-expressing and fusion-small subunit-expressing cells. Specifically, the insolubility of RuBisCO subunits expressed in E. coli which we observed is in agreement with the results of Gatenby for large subunits of Z. mays (29) and wheat (30) expressed in E. coli.

We assayed extracts containing large subunits for activity because there were conflicting reports in the literature about whether the large subunit alone was capable of catalysis (37, 38), and because it could not be determined in those cases where no activity was detected in isolated large subunits, whether the lack of activity could have been a result of the harsh conditions used to separate the subunits of the holoenzyme. Our failure to detect activity was duplicated by other investigators studying expression of higher plant large subunits in E. coli (29, 30). More recently, it has been shown that cyanobacterial large subunits expressed in E. coli do have catalytic activity, but the level of activity is less than 1% of the activity of the holoenzyme (39).

The amino terminus of the large subunit is blocked to Edman degradation, preventing direct sequencing (40), and it had been reported that the large subunit was made in the form of a precursor 10-20 amino acids larger than the mature size (41). The electrophoretic mobility of large subunit protein produced in E. coli

was compared to the mobility of large subunits isolated from Zea mays seedlings. I was unable to detect a difference in electrophoretic mobility in the gel system employed. Gatenby (30) reported that in the gels used in his experiments, E. coli-synthesized large subunits migrated as a protein one kilodalton larger than the plant large subunits. Subsequent studies have determined the structure of the amino terminus in plants, and shown that proteolytic cleavage is only one of several post-translational modifications that occur (40). Only two amino acids are removed, representing a loss of about 0.2 kilodaltons.

In conclusion, my results show that either subunit, when synthesized at high concentration in E. coli, accumulates in a rapidly sedimenting fraction. The large subunit, expressed alone, was not catalytically active, suggesting that simultaneous expression of both subunits would be required for enzymatic activity.

REFERENCES

1. Guarente, L., Roberts, T. M. & Ptashne, M. (1980) Science 209, 1428-1430.
2. Boss, M. A., Kenten, J. H., Wood, C. R. & Emtage, J. S. (1984) Nucleic Acids Res. 12, 3791-3806.
3. Paul, D. C., van Frank, R. M., Muth, W. L., Ross, J. W. & Williams, D. C. (1983) European J. Cell Biol. 31, 171-174.
4. von Meyenburg, K., Jorgensen, B. B. & van Deurs, B. (1984) EMBO J. 3, 1791-1797.
5. Shine, J. & Dalgarno, L. (1974) Proc. Natl. Acad. Sci. USA 71, 1342-1346.
6. Kozak, M. (1983) Microbiol. Rev. 47, 1-45.
7. Vieira, J. & Messing, J. (1982) Gene 19, 259-268.
8. Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) Gene 33, 103-119.
9. Bolivar, F., Rodriguez, R. L., Betlach, M. C. & Boyer, H. W. (1977) Gene 2, 95-113.
10. Queen, C. (1983) J. Molec. Applied Genet. 2, 1-10.
11. Coen, D. M., Bedbrook, J. R., Bogorad, L. & Rich, A. (1977) Proc. Natl. Acad. Sci. USA 74, 5487-5491.
12. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
13. Maxam, A. M. & Gilbert, W. (1980) Methods Enzymol. 65, 499-559.
14. Ryter, A. & Kellenberger, E. (1958) Z. Naturforsch. B 13, 597-605.

15. Meek, G. A. (1976) Practical Electron Microscopy for Biologists 2nd ed. Wiley, Chichester.

16. Laemmli, U. K. (1970) Nature 227, 680-685.

17. Towbin, H., Staehelin, T. & Gordon, J. (1979) Proc. Natl. Acad. Sci. USA 76, 4350-4354.

18. Blake, M. S., Johnston, K. H., Russell-Jones, G. J. & Gotschlich, E. C. (1984) Anal. Biochem. 136, 175-179.

19. Lennox, E. S. (1955) Virology 1, 190-206.

20. Gurevitz, M., Somerville, C. R. & McIntosh, L. (1985) Proc. Natl. Acad. Sci. USA 82, 6546-6550.

21. Pierce, J. W., McCurry, S. D., Mulligan, R. M. & Tolbert, N. E. (1982) Methods Enzymol. 89, 47-55.

22. Hall, N. P. & Tolbert, N. E. (1978) FEBS Lett. 96, 167-169.

23. McIntosh, L., Poulsen, C. & Bogorad, L. (1980) Nature 288, 556-560.

24. Mcintosh, L., Hirschberg, J., Somerville, C. R., & Fitchen, J. H. (1984) in Sybesma, C. (ed.) Advances in Photosynthesis Research Martinus Nijhoff/Junk, The Hague.

25 Weber, K., & Osborn, M. (1969) J. Biol. Chem. 244, 4406-4412.

26 Somerville, C. R., McIntosh, L., Fitchen, J. & Gurevitz, M. (1986) Methods in Enzymol. 118, 419-433.

27. Lebrun, M., Waksman, G. & Freyssinet, G. (1987) Nucleic Acids Res. 15, 4360.

28. Stormo, G. D., Schneider, T. D. & Gold, L. M. (1982) Nucleic Acids Res. 10, 2971-2996.

29. Gatenby, A. A., Castleton, J. A. & Saul, M. W. (1981) Nature 291, 117-121.

30. Gatenby, A. A. (1984) Eur. J. Biochem. 144, 361-366.

31. Bradley, D., van der Vies, S. M. & Gatenby, A. A. (1986) Phil. Trans. R. Soc. Lond. B 313, 447-458.

32. Sung, W. L., Yao, F.-L. & Narang, S. A. (1987) Methods Enzymol. 153, 385-389.

33. van der Vies, S. M., Bradley, D. & Gatenby, A. A. (1986) EMBO J. 5, 2439-2444.

34. Williams, D. C., van Frank, R. M., Muth, W. L. & Burnett, J. P. (1982) Science 215, 687-688.

35. Schoner, R. G., Ellis, L. F. & Schoner, B. E. (1985) Bio/Technology 3, 151-154.

36. Goeddel, D. V., Kleid, D. G., Bolivar, F., Heyneker, H. L., Yansura, D. G., Crea, R., Hirose, T., Kraszewski, A., Itakura, K. & Riggs, A. D. (1979) Proc. Natl. Acad. Sci. USA 76, 106-110.

37. Nishimura, M. & Akazawa, T. (1973) Biochem. Biophys. Res. Commun. 54, 842-848.

38. Andrews, T. J. & Ballment, B. (1983) J. Biol. Chem. 258, 7514-7518.

39. Andrews, T. J. (1988) J. Biol. Chem. 263, 12213-12219.

40. Mulligan, R. M., Houtz, R. L. & Tolbert, N. E. (1988) Proc. Natl. Acad. Sci. USA 85, 1513-1517.

41. Langridge, P. (1981) FEBS Lett. 123, 85-89.

Chapter 3

CONCURRENT EXPRESSION OF *Z. mays* RuBisCO LARGE AND SMALL SUBUNITS IN *E. coli*, AND EXPRESSION IN COMBINATION WITH *Anabaena* 7120 RuBisCO SUBUNITS.

INTRODUCTION

In the previous chapter the construction of plasmids directing the expression of the large subunit of *Z. mays* RuBisCO and two slightly different forms of the small subunit were described, and the results of experiments to characterize the expression products were presented. The large subunit, which bears the active sites (1,2,3), was found to be in an insoluble form and was not catalytically active. The next experiment was to combine the expression systems so that both large and small subunits would be produced simultaneously in the same bacterium. It was hypothesized that if the small subunit were required for solubility and activity, then, by co-expressing the two subunits, it might be possible to obtain an active form I RuBisCO in *E. coli*. The experiments described in this chapter were designed to explore this hypothesis.

While these experiments were being performed, preliminary results became available from experiments being performed by Michael Gurevitz that suggested that transcription of a cyanobacterial *rbc*

operon (including both rbcL and rbcS) in E. coli, under the control of the E. coli lac promotor, resulted in the synthesis of both RuBisCO subunits and assembly of active holoenzyme (reported in 4). Similar results were also reported by other groups (5,6,7). In light of the new results, I designed and conducted additional experiments to investigate how the co-expression of Z. mays RuBisCO subunits compared and contrasted with the co-expression of cyanobacterial RuBisCO subunits.

As in chapter two, many of the experiments described here are similar to experiments performed simultaneously or subsequently in other laboratories. The results of my experiments will be discussed in the context of the results of others.

MATERIALS AND METHODS

Antisera, molecular weight standards, and other specialty reagents were as described in chapter two. Purified spinach RuBisCO was a gift from Dr. S. Somerville.

Plasmids pBC12 and p77 were described in chapter two. Plasmid pZS13, directing the expression of a fusion protein similar to the precursor form of the Z. mays small subunit, was constructed by Erin Bell for experiments unrelated to this thesis. Its construction and characterization are described in an appendix to her thesis (8). Plasmid pANX105 was from Gurevitz et al. (4). Cosmid cloning vector pWH4 (9) was a gift from Dr. C. Peter Wolk.

DNA manipulation and analysis, SDS-PAGE, and western blotting were as described in chapter two.

Dual transformation of E. coli was performed serially. Cells containing a large subunit expression plasmid were made competent and transformed with a small subunit expression plasmid by standard procedures (10).

Growth of bacterial cultures, harvesting of cells, preparation of bacterial extracts, and carboxylase activity assays were as described in chapter two.

Sucrose density gradient centrifugation was performed as described (4).

RESULTS

Plasmid Constructions.

Incompatibility considerations.

Two plasmids utilizing the same mechanism for control of replication and control of plasmid copy number are said to belong to the same incompatibility group (for review see 11). Only one representative of any incompatibility group can be stably maintained in a given cell and its progeny. To allow stable maintenance of large subunit expression plasmids with small subunit expression plasmids, all of the small subunit plasmids were of one incompatibility group (colE1) and both of the large subunit expression plasmids were constructed in a vector of a second incompatibility group (lambda).

Subcloning of engineered gene expressing Z. mays large subunit to a plasmid with bacteriophage lambda replicon.

The Z. mays rbcL coding sequence, intact with the upstream lac promotor as constructed in the experiments described in the previous chapter, was isolated as a PvuII fragment from pBC12. Sal I linkers were used to create suitable "ends" and the fragment was ligated into the Sal I site of the cosmid vector pWH4 (9) to create pL53. Compatibility of this bacteriophage lambda replicon plasmid with the colE1 replicon small subunit expression plasmid p77 was demonstrated by scoring for both plasmid-encoded drug resistance markers and re-isolation of both plasmids from dually transformed cells (results not shown).

Construction of a plasmid compatible with colE1 replicons to direct the expression of Anabaena RuBisCO large subunits.

pL56, a plasmid to direct the expression of Anabaena 7120 large subunit in E. coli, was constructed using the engineered arrangement of lac promotor and rbcL made by Gurevitz et al. (4). A Pvu II - Sca I fragment including the lac promotor and rbcL coding sequence of pANX105 (4) was isolated and subcloned with the aid of Sal I linkers into the Sal I site of pWH4 (9). As with pL53, pL56 could be stably maintained with colE1 replicons (data not shown).

Construction of a plasmid to direct the expression of Anabaena RuBisCO small subunits in E. coli.

pSewt, a plasmid to direct the expression of Anabaena 7120 small subunit was made in two stages. A 0.5 kb ScaI-DraI fragment

Figure 3.1. Analysis of extracts prepared from E. coli expressing Z. mays RuBisCO large subunits and Z. mays RuBisCO small subunits. A, coomassie-stained SDS-PAGE gel. The lanes were loaded with: R, spinach RuBisCO; S, molecular weight standards; 1, pellet fraction after centrifugation of lysate from cells carrying expression plasmids pL53 and p77, centrifugation was for 10 min at 10,000 x g; 2, supernatant fraction from lysate and centrifugation used to prepare sample for lane 1; 3, pellet fraction after centrifugation of lysate from control culture at 10,000 x g for 10 min; 4, supernatant fraction after centrifugation of lysate from control culture at 10,000 x g for 10 min. B, western blot of SDS-PAGE gel identical to the one shown in A. The blot was developed with a mixture of antisera containing antibodies against tomato RuBisCO holoenzyme and Z. mays RuBisCO small subunit.

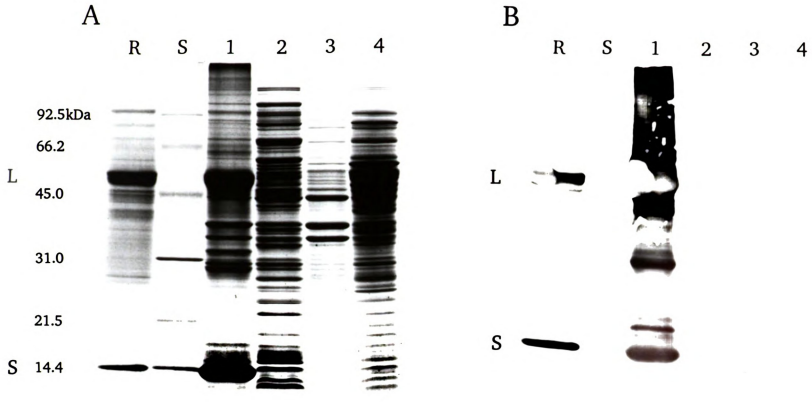


Figure 3.1

containing the rbcS protein coding region was isolated from pANX105 (4) and subcloned into pGC1 (12) which had been cleaved with BamHI and treated with S1 nuclease to create appropriate ends for ligation. The rbcS region was then isolated from this intermediate plasmid as an XbaI-EcoRI fragment and cloned into the XbaI and EcoRI sites of pUC19 (13). The rbcS coding region is not in the reading frame of the lacZ fragment of the vector and translation of rbcS message must begin at the initial methionine of the authentic small subunit. Further characterization of this plasmid and its expression products is described in chapter four.

Expression of RuBisCO Subunits in E. coli.

Co-expression of Z. mays large and mature small subunits.

Analysis of extracts of E. coli carrying both pL53 and p77, directing the expression of Z. mays RuBisCO large subunits and Z. mays RuBisCO small subunits, respectively, clearly showed that both subunits were accumulating to high levels in the bacteria. (A coomassie-stained SDS-PAGE gel and a western blot of an identical gel are shown in Figure 3.1). Fractionation of extracts by centrifugation at 10,000 x g for 10 min showed that the vast majority of both subunits was in a rapidly sedimenting fraction. Very low levels of both subunits were detected in the supernatant fraction in one experiment, but no soluble subunits observed in subsequent trials of the experiment (data not shown).

Several proteins of molecular weights different from both the large and small subunits were also observed to vary in abundance between the extracts of RuBisCO-expressing cells and extracts of control cells. The 28 kDa and the 33 kDa abundant proteins in the pellet fraction of the RuBisCO-expressing cells, not observed in the pellet fraction of the control cells, reacted with anti-RuBisCO antibodies and were probably products of discrete premature translational termination or site-specific proteolysis. The general background "smear" of antibody-cross-reactive material is probably due to products of random premature termination and non-specific degradation. The 43 kDa protein in the pellet fraction of the control cells that appears to be absent in the RuBisCO-expressing cells was probably encoded by the plasmid (uncharacterized, but not containing *rbcL* or *rbcS* coding sequence) carried by the control cells.

Carboxylase assays were performed on extracts from cultures bearing both expression plasmids as described (4, modified from 14). No carboxylase activity could be detected in either total lysate or pellet or supernatant fractions (data not shown).

Co-expression of the *Z. mays* large subunit and the precursor form of the *Z. mays* small subunit.

Plasmid pZS13 was constructed by Erin Bell for experiments in her thesis (8). It was shown to direct the synthesis in *E. coli* of high levels of a fusion protein consisting of seven residues of *E. coli* B-galactisidase fused to the amino terminus of the precursor form of the *Z. mays* small subunit (8). This product was localized

almost exclusively to an insoluble or rapidly sedimenting fraction (8).

An SDS-PAGE gel showing the analysis of extracts from E.coli carrying plasmids pL53 and pZS13 is presented in Figure 3.2. Both the Z.mays large subunit of RuBisCO and this precursor form of the Z. mays small subunit were very abundant in the insoluble fraction, with the small subunit precursor present in greater than equimolar amounts when compared to the large subunit. No evidence of soluble enzyme was detected with the SDS-PAGE gels and no catalytic activity could be detected in either the soluble or the insoluble fraction (data not shown).

Co-expression of Anabaena 7120 large and small subunits.

An SDS-PAGE gel and a western blot showing the analysis of extracts of E. coli bearing plasmids pL56 and pSewt are presented in Figure 3.3. Large subunit protein could be readily detected in the coomassie-stained SDS-PAGE gel, and probably represented 1-2% of the total bacterial protein. Only a small proportion of the large subunit protein was soluble, as was most clearly demonstrated by the relative intensities of the bands on the western blot. Small subunit protein could not be detected above the background of other proteins in the SDS-PAGE gel and was probably present in substoichiometric amounts. The Anabaena small subunit was not recognized by the antibody raised against tomato holoenzyme.

Both the soluble fraction and the insoluble fraction of the extract were analyzed for carboxylase activity. The soluble fraction was found to have a specific activity of 45 nmol CO_2 fixed/min/mg

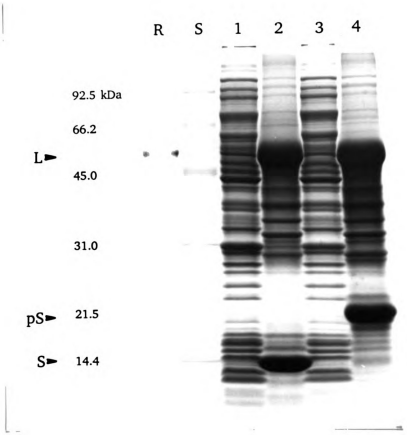


Figure 3.2. Analysis of extracts prepared from E. coli expressing Z. mays RuBisCO large subunits and the precursor form of Z. mays RuBisCO small subunits. The SDS-PAGE gel shown here was stained with coomassie blue. The lanes were loaded with: R, spinach RuBisCO; S, molecular weight standards; 1, supernatant fraction after centrifugation at 30,000 x g for 10 min of lysate from cells expressing Z. mays large and mature small subunits; 2, pellet fraction from the lysate and centrifugation used for lane 1; 3, supernatant fraction, after centrifugation at 30,000 x g of lysate from cells carrying plasmids pL53 and pZS13; 4, pellet from lysate and centrifugation used in preparation of sample for lane 3.

Figure 3.3. Analysis of extracts prepared from E. coli expressing Anabaena RuBisCO large subunits and Anabaena RuBisCO small subunits. A, coomassie-stained SDS-PAGE gel. The lanes were loaded with: R, spinach RuBisCO; S, molecular weight standards; 1, pellet fraction after centrifugation of lysate from cells expressing both Anabaena RuBisCO subunits (centrifugation was at 10,000 x g for 10 min); 2, supernatant fraction after centrifugation of lysate from cells expressing both Anabaena RuBisCO subunits (centrifugation was the same as for lane 1); 3, pellet fraction after centrifugation of lysate from control culture at 10,000 x g for 10 min; 4, supernatant fraction after centrifugation of lysate from control culture at 10,000 x g for 10 min. B, western blot of SDS-PAGE gel identical to the one shown in A. The blot was developed with a mixture of antisera containing antibodies against tomato RuBisCO holoenzyme and Z. mays RuBisCO small subunit.

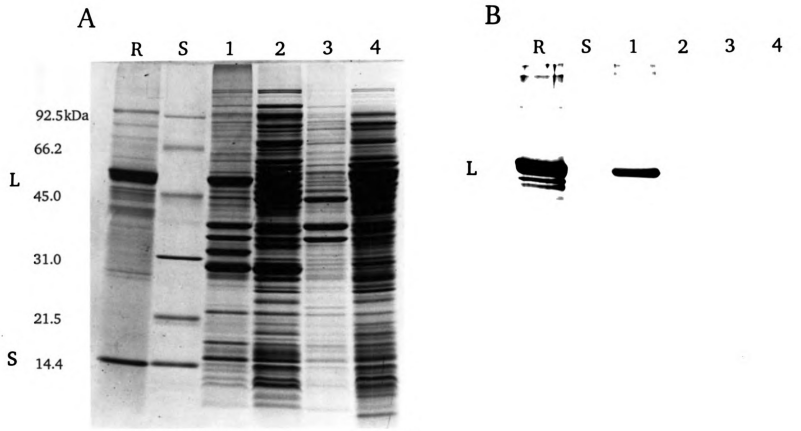


Figure 3.3

total protein. No activity could be detected in the insoluble fraction.

Co-expression of Z. mays large subunits with Anabaena small subunits.

An SDS-PAGE gel and a western blot showing the analysis of extracts of E. coli carrying plasmids pL53 and pSewt are presented in Figure 3.4. Large subunit protein was very abundant but was detected only in the insoluble fraction. Small subunit protein was not detected in either the coomassie-stained gel or the western blot, but a substantial amount would have been required for detection above the background of other proteins in the stained gel, and the antibody used for the western blot did not recognize Anabaena small subunit. No carboxylase activity could be detected in any fraction of extracts from these cells (data not shown).

Co-expression of Anabaena large subunits with Z. mays mature small subunits.

An SDS-PAGE gel and a western blot used for the analysis of extracts prepared from E. coli carrying plasmids pL56 and p77 are shown in Figure 3.5. Large subunit protein could be detected by coomassie-staining of proteins and with the antibody in the insoluble fraction but could not be detected by either method in the soluble fraction. The small subunit was much more abundant than the large subunit. Additionally, the titer of small subunit antibodies used for the immunodetection was much higher than the titer of the large subunit antibodies. Some small subunit could be detected in the soluble fraction on the western blot, but the vast majority of

Figure 3.4. Analysis of extracts prepared from E. coli expressing Z. mays RuBisCO large subunits and Anabaena RuBisCO small subunits. A, coomassie-stained SDS-PAGE gel. The lanes were loaded with: R, spinach RuBisCO; S, molecular weight standards; 1, pellet fraction after centrifugation of lysate from cells carrying expression plasmids pL53 and pSewt (directing the expression of Z. mays large subunits and Anabaena small subunits, respectively), centrifugation was for 10 min at 10,000 x g; 2, supernatant fraction after centrifugation of lysate for lane 1; 3, pellet fraction after centrifugation of lysate from control culture at 10,000 x g for 10 min; 4, supernatant fraction after centrifugation of lysate from control culture at 10,000 x g for 10 min. B, western blot of SDS-PAGE gel identical to the one shown in A. The blot was developed with a mixture of antisera containing antibodies against tomato RuBisCO holoenzyme and Z. mays RuBisCO small subunit.

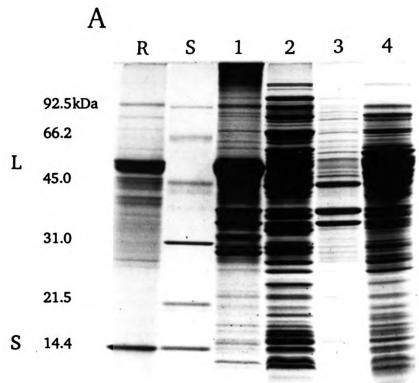


Figure 3.4

Figure 3.5. Analysis of extracts prepared from E. coli expressing Anabaena RuBisCO large subunits and Z. mays RuBisCO small subunits. A, coomassie-stained SDS-PAGE gel. The lanes were loaded with: R, spinach RuBisCO; S, molecular weight standards; 1, pellet fraction after centrifugation for 10 min at 10,000 x g of lysates from E. coli transformed with pL56 and p77; 2, supernatant fraction from lysate and centrifugation used for preparation of sample for lane 1; 3, pellet fraction after centrifugation of lysate from control culture at 10,000 x g for 10 min; 4, supernatant fraction after centrifugation of lysate from control culture at 10,000 x g for 10 min. B, western blot of SDS-PAGE gel identical to the one shown in A. The blot was developed with a mixture of antisera containing antibodies against tomato RuBisCO holoenzyme and Z. mays RuBisCO small subunit.

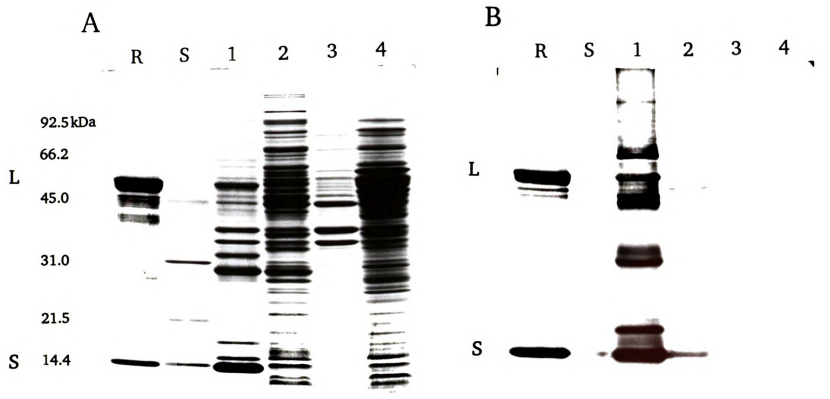


Figure 3.5

the small subunit protein was in the insoluble fraction.

Both the insoluble fraction and the soluble fraction were assayed for carboxylase activity. The soluble fraction was found to have a specific activity of 11 nmol CO_2 fixed/min/mg total bacterial protein. No activity could be detected in the insoluble fraction.

An aliquot of the soluble fraction was loaded on a sucrose density gradient as described (4). Assay of the fractions after centrifugation, and comparison to the activity profile from an identical gradient loaded with purified spinach RuBisCO showed that the activity in this hybrid-expression system had sedimentation properties of assembled holoenzyme. The activity profiles of the gradients loaded with bacterially synthesized hybrid enzyme and spinach RuBisCO are shown in Figure 3.6.

DISCUSSION

The experiments described in this chapter were a logical continuation of the experiments described in chapter two. In those experiments, Z. mays large and small subunits were expressed individually in E. coli. The large subunits, which bear the active sites, were not active, and were localized to an insoluble or rapidly sedimenting fraction. The experiments in this chapter were all based on the hypothesis that simultaneous expression of large and small subunits would be required and sufficient for the assembly of a soluble and active enzyme in E. coli. Expression of Z. mays and Anabaena RuBisCO large and small subunits, simultaneously, in

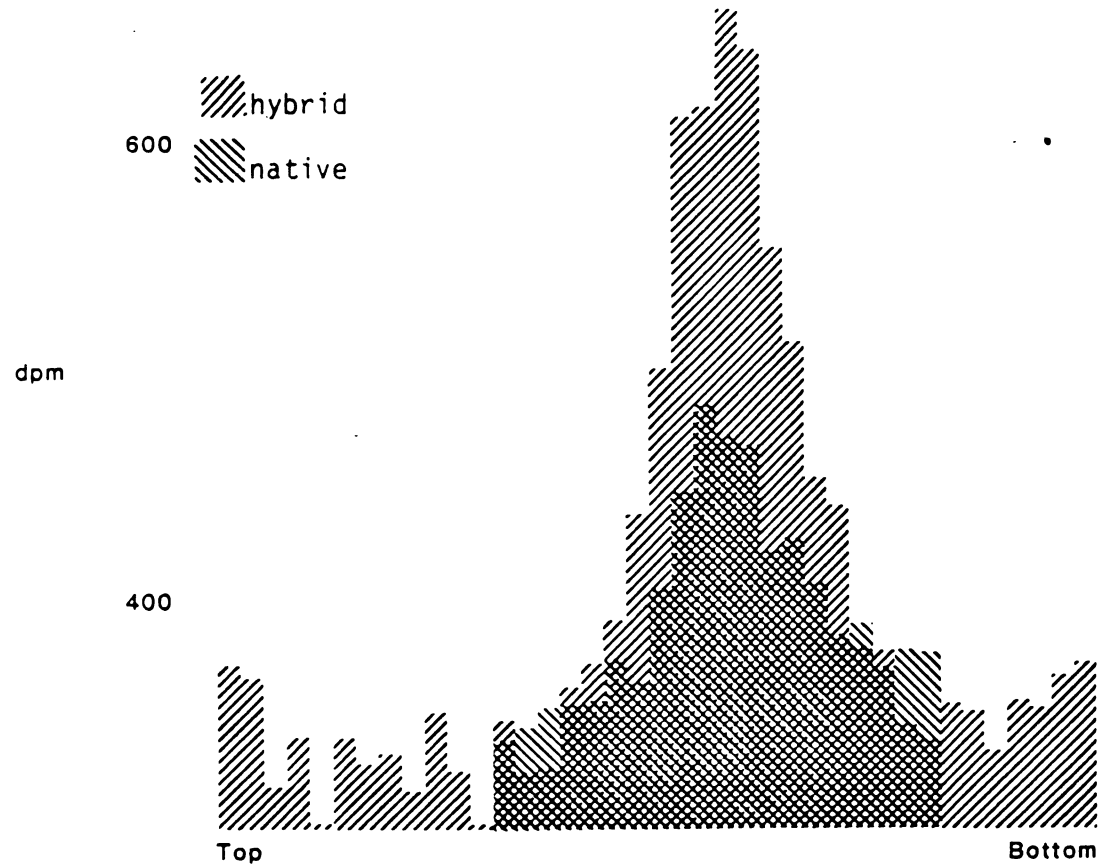


Figure 3.6. Sedimentation profiles of hybrid and native RuBisCO activities in sucrose density gradients. Lysate from cells expressing Anabaena large subunits and Z. mays small subunits was loaded onto a sucrose density gradient and centrifuged as described (4). An identical gradient was loaded with purified spinach RuBisCO and centrifuged opposite the bacterial lysate. Fractions were "dripped" from the bottom and were assayed for carboxylase activity as described (4). Consecutive fractions of the gradient are arrayed along the abscissa. The ordinate represents incorporation of radioactivity from $^{14}CO_2$ into acid stable forms and was not normalized for protein concentration in the fractions. The protein profile of a similar gradient is shown in Gurevitz et al. (4).

homologous and heterologous combinations, was used as an experimental system to test the hypothesis, and showed that the situation was more complicated than anticipated. In some subunit combinations the hypothesis correctly predicted the results, while in others it was not consistent with what was observed. The results of the various combinations are summarized in Figure 3.7.

The results from cells expressing both Z. mays subunits were similar to results subsequently reported by Gatenby et al. (15) and Bradley et al (16). Both of these reports also describe the construction of plasmids to direct the simultaneous expression of higher plant RuBisCO large and small subunits in E. coli. No carboxylase activity was detected in either case. Bradley et al. were able to detect both subunits in the soluble fraction, but analysis showed that the subunits had not assembled properly to form holoenzyme.

There were several differences, however, between the experiments described in this chapter and the experiments from the other laboratories, but in each case assembly and activity would have been predicted to be more likely in the experiments in this thesis for the following reasons. Most importantly, The experiments in Gatenby et al (15) used Z. mays rbcL and wheat rbcS, while my experiments used Z. mays genes for both subunits. Also of possible importance was the fact that Gatenby et al. used a temperature-inducible P_L promotor system, and therefore studied expression and assembly at 41-42°C, whereas my cultures were grown continuously at 37°C. This is important because studies with other foreign proteins expressed in E. coli have shown that the tendency of foreign proteins to form

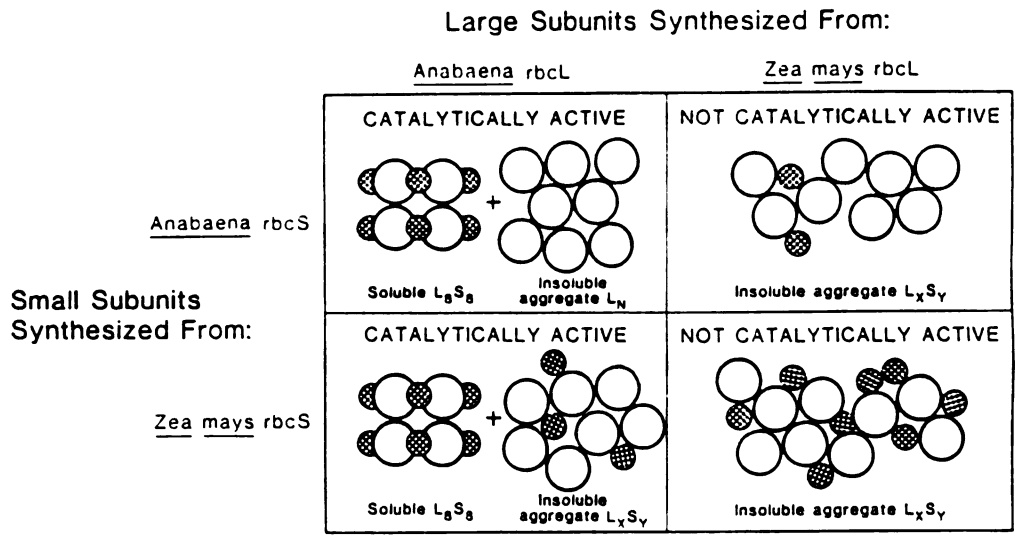


Figure 3.7. Summary of results of co-expression of combinations of Z. mays and Anabaena 7120 RuBisCO large and small subunits in E. coli. Data supporting the presence of Anabaena small subunit in the insoluble fraction when co-expressed with Z. mays large subunits is from western blots (not shown) using antiserum (described in chapter four) specific for the Anabaena small subunit.

insoluble aggregates was greater at higher temperatures (17). A third difference was that the amino terminus of the products expressed from the engineered rbcS genes in the experiments of Bradley et al. and Gatenby et al. was different from the amino terminus of the product of p77 described here. The structure of the protein expressed in their systems was missing the first four amino acids of the mature small subunit and had twelve residues from the E. coli β-galactosidase fused to the amino terminus of the small subunit sequence. The protein expressed by p77 contained all of the residues of the mature small subunit and had only seven extra residues fused to the amino terminus. The levels of expression of both subunits directed by pL56 and p77 were significantly higher than in either of the other two systems.

As an alternative to co-expression of the mature Z. mays small subunit with the Z. mays large subunit, I also co-transformed cells with the large subunit expression plasmid in combination with a plasmid directing the expression of the precursor form of the Z. mays small subunit. The rationale behind this experiment was that the transit peptide, which is cleaved off sometime during or after transport into the chloroplast, might be required for recognition of, or initial association with, the large subunits. The results were essentially the same as in the experiment with the mature small subunit. The subsequent demonstration of successful in vitro reconstitution of spinach mature small subunits with large subunit cores from RuBisCO from the cyanobacterium Synechococcus 6301, by Andrews and Lorimer (18), established that the transit peptide is not required for holoenzyme assembly.

The experiment in which Anabaena large and small subunits were expressed simultaneously differed only slightly from the experiments of Gurevitz et al. (4) and others (5,6,7). The principal difference was that the two subunits were expressed from two promotors and on two plasmids, rather than being in an operon. My results clearly showed that the operon structure, found in all cyanobacteria examined to date (19,20), and in Chromatium vinosum (21) and Alkaligenes eutrophus (22), was not a requirement for expression or assembly in E. coli. Levels of carboxylase activity in extracts from the two plasmid system were about two-fold higher than levels reported by Gurevitz et al. (4) for the one plasmid, one operon system, but the arrangement of the promotors relative to the small subunit coding sequence was different in the two experiments, and differences in the levels of expression of small subunit protein in the two systems could have accounted for the differences in levels of assembly and catalytic activity.

The results of the experiment in which Z. mays large subunits were co-expressed with cyanobacterial small subunits were not very informative when considered in isolation , as no activity could be detected and most or all of both subunits were found in the insoluble fraction. When considered in light of the results from the experiments expressing both Anabaena subunits and both Z. mays subunits, however, these results were more meaningful. They suggested that the problem preventing holoenzyme assembly in the case of co-expression of both Z. mays subunits was, at least in part, due to problems related to assembly of the large subunits. (This hypothesis was supported by the results of the other heterologous

combination and by results from other laboratories and will be discussed further.)

The experiment in which Anabaena large subunits were co-synthesized with Z. mays small subunits was similar to an experiment subsequently reported by van der Vies et al.(23), with similar results. In both experiments, subunits assembled an active holoenzyme with sedimentation velocity similar to that of authentic RuBisCO. Comparison of the levels of activity measured by van der Vies et al. (23) in soluble extracts from E. coli expressing Synechococcus PCC 6301 large subunits and wheat small subunits suggested that the hybrid enzyme had only 10% of the activity of the homologous Synechococcus enzyme. In my experiments, less activity was also observed in crude extracts from cultures expressing heterologous subunits, but because the hybrid enzyme was not purified (the enzyme was also not purified in the experiments of van der Vies et al.) the activities of the natural and hybrid enzymes could not be menaningfully compared. Careful kinetic characterization of a hybrid enzyme with cyanobacterial large subunits and spinach small subunits, made by in vitro reconstitution by Andrews and Lorimer (18), showed that this related hybrid enzyme had a specific activity about half that of the non-hybrid cyanobacterial enzyme.

Additionally the results from this combination served as a positive control for the experiment in which both Z. mays subunits were co-expressed. The activity in extracts from cells bearing pL56 and p77 demonstrated that small subunits synthesized from p77 could support catalytic activity, and thus that the extra residues at the amino terminus did not interfere with assembly or catalysis.

CONCLUSIONS

Comparing the data from the subunits co-synthesized in all of the combinations, I initially drew the conclusion that only cyanobacterial large subunits could assemble into holoenzyme in E. coli, and that both higher plant and cyanobacterial small subunits could assemble with the cyanobacterial large subunits. I was unable to determine why the Z. mays large subunits could not assemble and was left with several possible explanations for the results. Among the possible problems with expression of the Z. mays large subunits in E. coli could have been the formation of inappropriate disulfide bonds. The Z. mays large subunit contains 11 cysteine residues (24) and disulfide bridges not present in authentic RuBisCO could have formed as a result of a difference in reductive status between chloroplasts and bacteria. That inappropriate bonds did not form in the cyanobacterial large subunit, or that they did not interfere with assembly or activity, could have indicated either that the intracellular environment of E. coli is more similar to cyanobacteria than to chloroplasts, or that the chloroplast RuBisCO had evolved to a point where it was no longer able to fold correctly under as broad a range of conditions. A second possibility was that the Z. mays large subunits were aggregating so rapidly after synthesis that none were available for holoenzyme assembly. Supporting this possibility was the observation that when higher plant RuBisCO is stripped of its small subunits the large subunits aggregate and cannot be redissolved, whereas cyanobacterial large subunits can be redissolved quite readily (18). A third possibility was that there was some

additional factor needed for assembly of higher plant RuBisCO. Indeed, a protein had been identified that associated with newly synthesized large subunits *in vivo*, and antibodies to the protein had been shown to inhibit RuBisCO assembly in chloroplast extracts (25,26).

More recent results from other laboratories strongly support the third possibility listed above. The protein found associated with newly synthesized large subunits has been purified and shown to be comprised of heterogeneous subunits and the genes for the subunits have been cloned (27,28 and R. J. Ellis personal communication) A veritable breakthrough occurred when sequence analysis of one of the genes revealed substantial similarity to the *E. coli groEL* gene, also implicated in macromolecular assembly (28). This led to the description of a new class of proteins from bacteria, chloroplasts, and mitochondria, which are all involved in assembly of macromolecular complexes (for review, see 29). These proteins have been designated "chaperonins".

The role of chaperonin in RuBisCO assembly was further demonstrated, and a more straightforward explanation for the results of my experiments was provided, by the results of a recent experiment by Goloubinoff *et al*. (30). Using *E. coli* mutants deficient in the *E. coli* chaperonin subunits (*GroEL* and *GroES*) and a plasmid directing the synthesis of both *Synechococcus* 6301 RuBisCO subunits, they were able to show that endogenous chaperonin was required for assembly of the cyanobacterial RuBisCO in *E. coli*. Further, they showed that if they co-transformed *E. coli* with a plasmid directing the overexpression of the chaperonin subunits, in addition to the plasmid

directing the expression of the cyanobacterial RuBisCO subunits, more of the RuBisCO subunits were found to assemble into holoenzyme. In the absence of small subunits, assembly of RuBisCO large subunits into an octameric core was found to be dependent on the E. coli chaperonin (30).

These more recent findings strongly suggest that the specificity of the endogenous E. coli chaperonin was responsible for the differences in the results of the different subunit combinations. The most reasonable hypothesis is that the E. coli chaperonin was able to function with the cyanobacterial large subunits, forming a large subunit core with which the small subunits then associated, but, for some reason, possibly because of greater evolutionary divergence from some primordial chaperonin recognition signal, the E. coli chaperonin could not function with higher plant large subunits. As a result, octameric large subunit cores were not formed in the combinations where the large subunits were synthesized from Z. mays rbcL, and there was no structure available with which the small subunits could associate.

This explanation of my results is consistent with the data in that the specificity determining whether an active enzyme was assembled resided in the large subunit and chaperonin has been shown to bind the large subunit (25). The identity of the small subunits was not important for assembly, and chaperonin has not been shown to have a specific interaction with the small subunit (31). Additionally, if chaperonin activity was limiting in the assembly of RuBisCO subunits in E. coli in my experiments (as was the case in the experiments of Goloubinoff et al. (30)) then the result that most of

both the large and small subunit proteins was found in the insoluble fraction, in the combination of Anabaena large subunits and Z. mays small subunits in which assembly was possible, would also be explained.

Following from this hypothesis is a clear and testable prediction. If the higher plant subunits cannot assemble in E. coli because of the lack of a suitable chaperonin, then co-expression with a functional, higher plant chaperonin should result in assembly. This experiment is almost certainly in progress, although no results have been reported. If successful, it would conclusively disprove the original working hypothesis of this chapter: that co-expression of the two Z. mays RuBisCO subunits in E. coli would be necessary and sufficient for assembly of an active enzyme.

REFERENCES

1 Andrews, T. J. (1988) J. Biol. Chem. 263, 12213-12219.

2 Chapman, M. S., Suh, S. W., Cascio, D., Smith, W. W. & Eisenberg, D. (1987) Nature 329, 354-356.

3 Andersson, I., Knight, S., Schneider, G., Lindqvist, Y., Lundqvist, T., Branden, C. I. & Lorimer, G. H. (1989) Nature 337, 229-234.

4 Gurevitz, M., Somerville, C. R. & McIntosh, L. (1985) Proc. Natl. Acad. Sci. USA 82, 6546-6550.

5 Tabita, F. R. & Small, C. L. (1985) Proc. Natl. Acad. Sci. USA 82, 6100-6103.

6 Gatenby, A. A., van der Vies, S. M. & Bradley, D. (1985) Nature 314, 617-620.

7 Christeller, J. T., Terzaghi, B. E., Hill, D. F. & Laing, W. A. (1985) Plant Mol. Biol. 5, 257-263.

8 Bell, E. (1986) Ph.D. Thesis, Michigan State University.

9 Herrero, A., Elhai, J., Hohn, B. & Wolk, C. P. (1984) J. Bacteriol. 160, 781-784.

10 Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

11 Davison, J. (1984) Gene 28, 1-15.

12 Myers, R. M., Lerman, L. S. & Maniatis, T. (1985) Science 229, 242-247.

13 Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) Gene 33, 103-119.

14 Pierce, J. W., McCurry, S. D., Mulligan, R. M. & Tolbert, N. E. (1982) Methods Enzymol. 89, 47-55.

15 Gatenby, A. A., Van der Vies, S. M. & Rothstein, S. J. (1987) Eur. J. Biochem. 168, 227-231.

16 Bradley, D., Van der Vies, S. M. & Gatenby, A. A. (1986) Phil. Trans. R. Soc. Lond. B 313, 447-458.

17 Schein, C. H. & Noteborn, M. H. M. (1988) Bio/technology 6, 291-294.

18 Andrews, T. J. & Lorimer, G. H. (1985) J. Biol. Chem. 260, 4632-4636.

19 Shinozaki, K. & Sugiura, M. (1983) Nucleic Acids Res. 11, 6957-6964.

20 Nierzwicki-Bauer, S. A., Curtis, S. E. & Haselkorn, R. (1984) Proc. Natl. Acad. Sci. USA 81, 5961-5965.

21 Viale, A. M., Kabayashi, H., Takabe, T. & Akazawa, T. (1985) FEBS Lett. 192, 283-288.

22 Andersen, K. & Caton, J. (1987) J. Bacteriol. 169, 4547-4558.

23 Van der Vies, S. M., Bradley, D. & Gatenby, A. A. (1986) EMBO J. 5, 2439-2444.

24 Mcintosh, L., Poulsen, C. & Bogorad, L. (1980) Nature 288, 556-560.

25 Barraclough, R. & Ellis, R. J. (1980) Biochim. Biophys. Acta 608, 19-31.

26 Cannon, S., Wang, P. & Roy, H. (1986) J. Cell Biol. 103, 1327-1335.

27 Musgrove, J. E., Johnson, R. A. & Ellis, R. J. (1987) Eur. J. Biochem. 163, 529-534.

28 Hemmingsen, S. M., Woolford, C., van der Vies, S. M., Tilly, K., Dennis, D. T., Georgopoulos, C. P., Hendrix, R. W. & Ellis, R. J. (1988) Nature 333, 330-334.

29 Ellis, R. J. & Hemmingsen, S. M. (1989) Trends. Biochem. Sci., in press.

30 Goloubinoff, P., Gatenby, A. A. & Lorimer, G. H. (1989) Nature 337, 44-47.

31 Ellis, R. J. & Van der Vies, S. M. (1988) Photosynthesis Res. 16, 101-115.

Chapter 4

RESIDUES IN THREE CONSERVED REGIONS OF THE SMALL SUBUNIT OF RuBisCO ARE REQUIRED FOR QUATERNARY STRUCTURE*

ABSTRACT The function of the small subunit of Ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO, E.C. 4.1.1.39) is unknown. In order to explore the role of individual residues, small subunits with single amino acid substitutions in three regions of relative sequence conservation were produced by directed mutagenesis of the rbcS gene from Anabaena 7120. These altered small subunits were co-synthesized with large subunits (from an expressed Anabaena rbcL gene) in E. coli. Mutants were analyzed for effects on quaternary structure and catalytic activity. Changing Glu 13S (numbering used is that of spinach RuBisCO) to Val, Trp 67S to Arg, Pro 73S to His, or Tyr 98S to Asn all prevented accumulation of stable holoenzyme. Interpretation of these results using a model for the three-dimensional structure of spinach RuBisCO based on X-ray crystallographic data (Andersson et al. 1989 Nature 337, 229-234 and Knight et al. 1989 Science 244, 702-705) suggests that our small subunit mutants containing substitutions at positions 13S and

* manuscript submitted for publication

67S probably do not assemble because of mis-pairing or non-pairing of charged residues on the interfacing surfaces of the large and small subunits. The failure of small subunits substituted at positions 73S or 98S to assemble correctly may result from disruption of intersubunit or intrasubunit hydrophobic pockets, respectively.

INTRODUCTION Ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO, E.C.4.1.1.39) catalyzes the first committed step of both the reductive and oxidative photosynthetic carbon cycles (for review, see 1). The holoenzyme in cyanobacteria and plants is composed of eight large subunits (L) and eight small subunits (S). The large subunits form the catalytic sites of the enzyme and provide all catalytically essential residues (2). The contribution of the small subunits to the function of the holoenzyme is unresolved.

The importance of the small subunits to the catalytic activity of the enzyme is demonstrated by a 200-fold drop in the V_{max} when the small subunits are removed (2). Activity can be restored by reconstitution of the resulting L octamer with isolated small subunits (2), but the structural requirements for this restoration and the mechanism by which activity is effected are not understood.

Several systems have been described which allow the use of site-directed mutagenesis to study the role of individual residues in the small subunit. These systems are based on the observation that a normal holoenzyme is assembled when cyanobacterial *rbcL* and *rbcS* genes, encoding the large and small subunits, respectively, are expressed in *E. coli* (3,4, 5,6). Some of the mutant enzymes produced have shown reduced activity (7,8); however, the causes of

the reductions in activity have not been determined and specific functions have not been assigned to individual S residues.

We have used directed mutagenesis of rbcS from the cyanobacterium Anabaena 7120 and co-expression with Anabaena rbcL in E. coli to investigate the function of individual residues in three conserved regions of the small subunit. In this paper we present data from four single amino acid substitutions, representing the three regions of broadest sequence conservation, which were analyzed for their effect on catalytic function and holoenzyme formation. In order to better understand the function of these residues in the native holoenzyme, these amino acids have been examined in the context of a model for the three-dimensional structure of spinach RuBisCO (9). For each of the four mutations, we discuss the structural implications of the substitution made and propose structural explanations for the results observed.

MATERIALS AND METHODS

CONSTRUCTION OF PLASMIDS. DNA manipulations were according to standard protocols (10).

pL56, a plasmid to direct the expression of Anabaena RuBisCO L subunits in E. coli, was constructed by excising a PvuII-ScaI fragment including the lac promotor and Anabaena rbcL coding sequence from pANX105 (6) and subcloning it, with the aid of SalI linkers, into the SalI site of the cosmid vector pWH4 (11).

p205-3, a "phagemid" plasmid for production of single stranded template DNA for mutagenesis, was made by subcloning a 0.5kb ScaI-DraI fragment containing the Anabaena 7120 rbcS protein coding

region into pGC1 (12), which had been cleaved with BamH1 and treated with S1 nuclease to create appropriate ends for ligation. A plasmid with theorientation placing the 5' end of rbcS adjacent to the XbaI site of the vector was designated p205-3.

pSewt, pSe13, pSe67, pSe73 and pSe98, plasmids to direct the expression of authentic or substituted Anabaena RuBisCO S in E. coli, were made by subcloning XbaI-EcoRI fragments from p205-3 and its mutant derivatives into the XbaI and EcoRI sites of pUC19 (13). The rbcS coding region is not in the reading frame of the lacZ fragment of the vector and translation of rbcS message must begin at Met 1S.

MUTAGENESIS. Single-stranded DNA templates of p205-3 were generated by superinfection with M13K07 as described (14). Oligonucleotides for each conserved region were synthesized as mixtures of oligonucleotides without and with one, two, or more differences from the wild-type sequence. Mutagenesis was performed as described (15). Mutants were identified by sequencing the appropriate region of randomly selected progeny plasmids from the mutagenesis protocol using the dideoxy chain termination method (16). The entire rbcS cassette was sequenced in putative single mutants to exclude the possibility of additional, spurious, mutations.

GROWTH OF CELLS. Cells were grown in LB medium (10). Antibiotics were added when appropriate at the following concentrations: ampicillin, 50 mg/l; kanamycin, 30 mg/l. Three-hundred ml cultures for analysis of mutant or wild-type proteins were started with a 1 ml inoculum and grown for 24 hr at 37°C with vigorous aeration.

PREPARATION OF BACTERIAL EXTRACTS AND ASSAY OF CARBOXYLASE ACTIVITY Cells were harvested by centrifugation, resuspended in assay buffer

(50 mM Tricine, pH 8.0/10 mM $MgCl_2$/10 mM 2-mercaptoethanol/1 mM phenylmethylsulfonyl fluoride), and disrupted by ultrasonic treatment (6). The cell extract was centrifuged at 100,000 x g for 1 H. Unwashed pelleted material was resuspended in SDS-PAGE sample buffer (17) and boiled for 2 min. Saturated ammonium sulfate solution was added to a portion of the 100,000 x g supernatant fraction to a final concentration of 40% (w/v) ammonium sulfate. Aggregated proteins were collected by centrifugation at 10,000 x g for 15 min, resuspended in assay buffer and dialyzed against assay buffer. The carboxylase activity of this preparation was determined as $RbuP_2$-dependent incorporation of radioactivity from $^{14}CO_2$ into acid-stable forms (6). Protein was quantified by a modified Lowry procedure (18).

ASSAY FOR HOLOENZYME FORMATION. A 2.5 mg sample of the crude 100,000 x g supernatant fraction was loaded on a 5.2 ml sucrose gradient (0.2-0.8 M sucrose in 50 mM Tris, pH 8.0 /0.1 mM EDTA/1 mM 2-mercaptoethanol/50 mM $NaHCO_3$/10 mM $MgCl_2$)(19) and centrifuged at 50,000 rpm in a Beckman VTi65 rotor (239,000 x g) for 70 min at 4°C. Three 300 µl fractions centered around the position of 18S particles (as estimated by sedimentation of purified spinach RuBisCO, a gift from S. C. Somerville, in equivalent gradients) were pooled and proteins in an 80 ul sample were resolved by SDS-PAGE (20 as modified in 17) and electrophoretically transferred to nitrocellulose membrane (17). S was detected using S-specific polyclonal antibodies and protein-A-linked alkaline phosphatase according to standard immunodetection protocols (21).

STRUCTURAL INTERPRETATION OF MUTANTS. The model of the spinach L_8S_8 RuBisCO holoenzyme was derived from an isomorphous electron density map with phase angles refined by fourfold averaging. The structures of both the large and the small subunits have been briefly described (22, 9). The model was examined on a graphics display using FRODO (23) and the steric effects of mutations were examined by substituting appropriate side chains. Accessibility calculations were made using the program of Lee and Richards (24). Secondary structures in wild-type and substituted S were predicted with the algorithm of Chou and Fasman (25).

RESULTS AND DISCUSSION

SYNTHESIS AND ANALYSIS OF AUTHENTIC AND SUBSTITUTED PROTEINS

Experimental System. Transformation of E. coli with the rbcL and rbcS expression plasmids pL56 and pSewt results in the synthesis of functional heteromultimeric RuBisCO (Figure 1, row 1). This result is similar to that observed with RuBisCO genes from other cyanobacteria expressed in E. coli (3,4,5) and with both Anabaena 7120 genes expressed as an operon on a single plasmid (6). The RuBisCO produced from such systems is apparently identical to the native enzyme, although post-translational processing and modification of L, known to occur in higher plants (26), has not been investigated.

Mutagenesis. Oligonucleotide mixtures were used to generate a collection of mutations in each of three relatively conserved regions of the small subunit. The mutations presented in this paper result in the following substitutions: Glu 13S→Val, Trp 67S→Arg, Pro 73S→

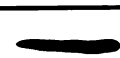
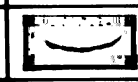
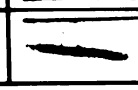
Substitution	Activity	18S	Pellet
None (wt)	51±17		
Glu 13S ⟶ Val	0.0		
Trp 67S ⟶ Arg	0.0		
Pro 73S ⟶ His	0.0		
Tyr 98S ⟶ Asn	0.0		

Figure 1. Loss of carboxylase activity and absence of holoenzyme resulting from substitutions in the small subunit. Residue numbering is that of spinach. Activity is expressed as nmol CO_2 fixed/min/mg protein and was assayed on a crude RuBisCO preparation from dually transformed (expressing L and native or substituted S) E. coli. The last two columns are sections of "western" blots of SDS-PAGE gels after immunodetection of S. The samples for the lanes in the 18S column were taken from sucrose gradients used to separate holoenzyme from unassembled subunits and included the region where assembled unsubstituted holoenzyme migrates. The samples for the pellet column were from unwashed 10,000 x g pellets of the ultrasonically disrupted cells and were prepared by partial solubilization in boiling 3% SDS PAGE sample buffer (20).

His, and Tyr 98S → Asn (Figure 1).

Analysis of Enzyme Activity. Cultures of E. coli containing pL56 and either pSewt or one of the mutant rbcS plasmids were grown and extracts were prepared for analysis. Relatively crude preparations were used for analysis to avoid the possible loss of an aberrant but still active or assembled enzyme. A very low level of activity (less than 18 pmol CO_2 fixed/min/mg total protein) was detected in control preparations from cells with pL56 but lacking an S-expressing plasmid. This activity is most likely the residual activity from L octamers (2) and sets the lower limit of detectable activity by mutant enzymes at about 0.2% of wild type activity. Oxygenase activity was not measured since partitioning between carboxylation and oxygenation appears to be a property determined solely by the large subunits (27). All four of the mutants were found to lack detectable carboxylase activity (Figure 1).

Analysis of Quaternary Structure. To explore a possible reason for the lack of activity in these mutants, extracts were analyzed by sedimentation through sucrose density gradients to separate assembled holoenzyme from unassembled subunits. After centrifugation, no S was detectable for any of the mutants at the position in the gradient where 18S holoenzyme is normally found (Figure 1). The sensitivity of our detection was such that levels of S greater than 2% of those in the non-substituted case would have been detected (data not shown). Low levels of S were sometimes detected at the top of the gradients (data not shown), and were reliably detected in crude pellets (Figure 1) indicating mutant S was being synthesized as expected.

CORRELATION OF MUTANTS WITH THE THREE DIMENSIONAL STRUCTURE.

To gain insight into the structural roles of the substituted residues and to try to understand how the substitutions made would interfere with formation or stability of the correct quaternary structure, we examined residues 13S, 67S, 73S, and 98S and residues in their vicinities in a model for the three dimensional structure of spinach RuBisCO. The amino acid sequences of the small subunits from Anabaena and spinach are 41% identical overall, but are more similar in the regions of our substitutions (Figure 2).

Glu 13S → Val. Glu 13S is in the interface region between the large and small subunits and is accessible to solvent in the free S but is buried in the L_8S_8 complex (Table 1). It is conserved in small subunits of known sequence (1,7,31) except those of the hydrogen bacterium Alcaligenes eutrophus (31) and the eukaryotic alga Olisthodiscus luteus (R. A. Cattolico, personal communication) which are quite different in this region of S. Glu 13S is close to Lys 164L (see Figure 3.c) and there is an intersubunit salt bridge between them. The surrounding region is highly charged (Figure 3.c). There are two nearby internal salt bridges in L, Glu 234L- Arg 421L and Asp 198L- Arg 167L. These four L residues are conserved in all L subunits including those of the Rhodospirillum rubrum enzyme (32), in which there are no small subunits. In R. rubrum the residue corresponding to 164L is an Arg.

Removal of the negative charge on Glu 13S in the mutant by replacement with Val could decrease the attraction between L and S, and, were they to assemble correctly, would result in an unpaired positive charge of Lys 164L buried in the interface between them. In

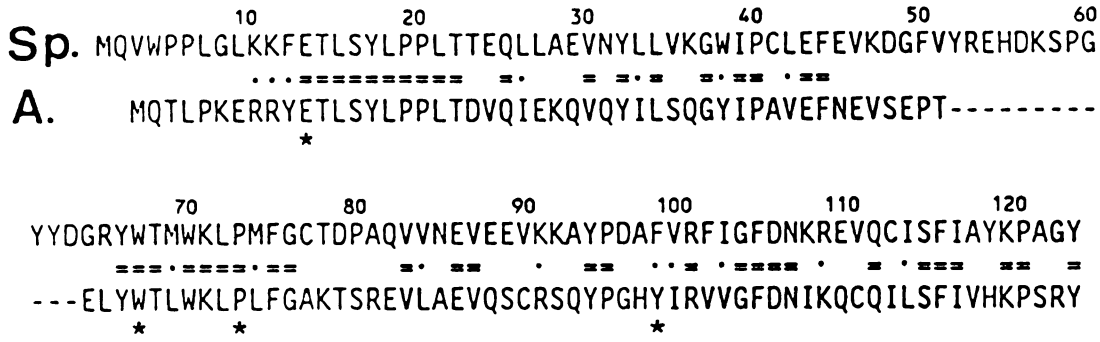


Figure 2. Comparison of spinach (Sp.)(28, corrected in 29 and 9) and Anabaena 7120 (A.)(30) RuBisCO S subunit sequences. The positioning of the gap in the Anabaena sequence is based upon the three dimensional structure of the spinach protein in which residues 52S to 63S form a loop. Similar (·) and identical (=) residues in the two sequences are indicated. Residues which were mutated are marked with an asterisk.

Figure 3. Drawings from model for three dimensional structure of spinach RuBisCO. A. S (bold lines) shown in relation to four L subunits. The intact holoenzyme is cube-shaped with the main body of the cube formed by four L dimers (two dimers are shown here). The S subunits are wedged between the ends of the dimers forming clusters of four each on the top and the bottom of the cube. The L subunit dimer shown on the left is formed by L subunits C and D and the dimer shown on the right is formed by L subunits A and B. B. S in roughly the same orientation as in A showing the location of residues 13S, 67S, 73S, and 98S. All non-hydrogen atoms are shown for these residues. Serial α-carbons are shown and connected for all other residues. C-F. Stereo drawings of residues in the vicinities of residues 13S(C.), 67S(D.), 73S(E.), and 98S(F.).

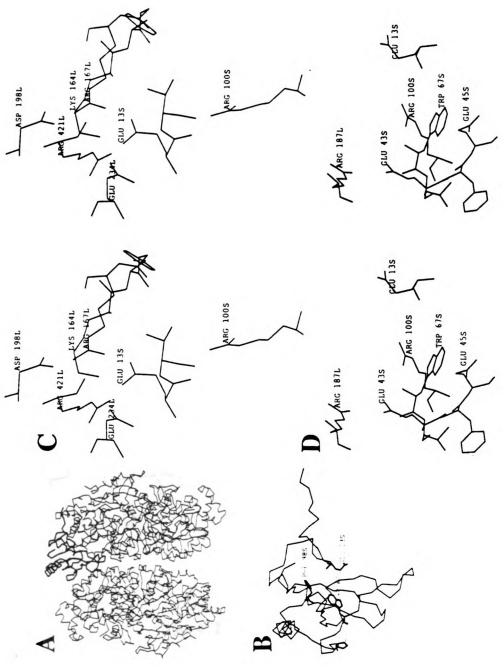

A
B
C
D
ASP 198L
ARG 421L
LYS 164L
ARG 167L
GLU 135
GLU 134L
ARG 100S
ARG 187L
GLU 43S
GLU 45S
TRP 67S

Figure 3

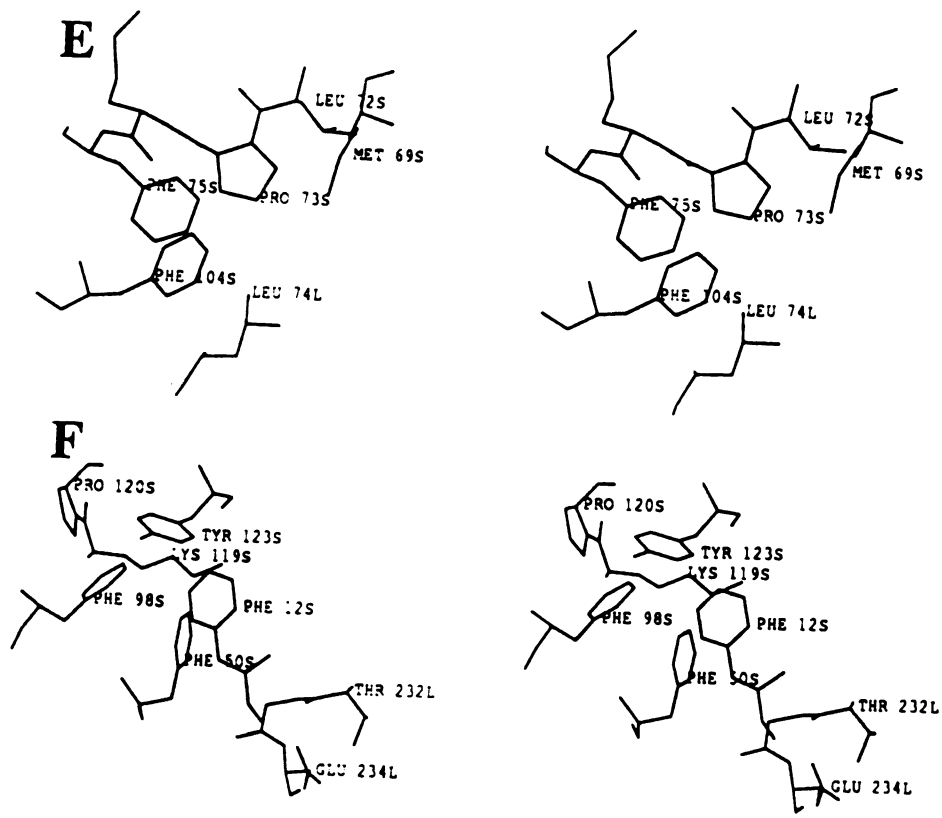


Figure 3 (cont'd.)

Table 1. Solvent accessibility of selected residues in model for three dimensional structure of spinach RuBisCO. Accessible surface area was calculated for individual residues in the model for the three dimensional structure of spinach RuBisCO and in the context of the appropriate S or L monomer alone, using the program of Lee and Richards (24) and is expressed as a percent of the accessible surface area of a tripeptide containing the amino acid found at that position.

residue	% accessible in monomer in isolation	% accessible in holoenzyme	decrease due to intersubunit contact
Phe 12S	16.1	4.6	11.5
Glu 13S	62.3	19.1	43.2
Glu 43S	14.8	3.3	11.5
Phe 50S	33.9	21.1	12.8
Trp 67S	11.2	4.6	6.6
Met 69S	40.2	10.8	29.4
Leu 72S	69.4	24.4	45.0
Pro 73S	20.3	4.2	16.1
Phe 75S	80.3	54.1	26.2
Phe 98S	1.8	1.8	0.0
Phe 104S	23.9	4.1	19.8
Lys 119S	33.6	33.6	0.0
Pro 120S	23.8	23.8	0.0
Tyr 123S	49.8	49.8	0.0
Leu 74L	58.3	8.9	49.4

addition to decreased stability of the assembled enzyme, this unpaired charge might interfere with proper alignment of the other charges in this region. Assembly and activity are probably both quite sensitive to small disruptions in this region since it is at the junction of the C and N domains of the L subunit (22).

Trp 67S → Arg. The conserved Trp 67S is in a quite polar environment near the interface between S and the two L subunits B and D (see Figure 3.a for designation of L subunits). In this region, Glu 13S interacts with subunit B as discussed for the previous mutant, and Glu 43S forms a salt bridge to Arg 187L of subunit D (Figure 3.d) which is conserved except in R. rubrum (32). This location of Glu 43S on the surface of S which is in contact with L in the holoenzyme is reflected in the residue being less accessible in the holoenzyme than in free S (table 1). Arg 100S is situated between Trp 67S and Glu 43S with the plane of the guanidinium group roughly parallel to the indole ring.

In the mutant, positive charge on the introduced Arg would repel Arg 100S, possibly forcing it closer to Glu 43S. This might result in a novel intrasubunit salt bridge between the two residues. At the very least, the negative charge in the immediate environment of Glu 43S on the interfacing surface of S would be diminished and the intersubunit attraction leading to a salt bridge between Glu 43S and Arg 187L in the wild-type enzyme would be reduced. Alternatively, the presence of two positively charged arginines close together could cause one of them to form a salt bridge with Glu 13S, preventing this residue from bridging to Lys 164L of large subunit B. In either case, assembly would be negatively affected. Additionally, the

substitution by an Arg side chain for the bulky Trp which is almost completely buried (Table 1) would leave an internal cavity in the small subunit which could destabilize local structure.

Pro 73S → His. The conserved residue Pro 73S is situated in a hydrophobic patch on the surface of S, which also includes conserved residues Leu 72S, Phe 75S and Phe 104S and Met 69S (Leu in Anabaena) (Figure 3.e). This patch forms an area of hydrophobic interaction with L subunit C involving Leu 74L. This Leu residue is conserved in all L subunits except L from R. rubrum (32). Consistent with their forming an intersubunit hydrophobic pocket, the accessibility of all of these residues to solvent decreases upon association of S and L (Table 1).

Replacement of the Pro side group by a polar His side chain would alter this pocket and make a more polar environment unsuitable for interaction with Leu 74L from the large subunit. It is also possible that the Pro residue is important for determining the local conformation of this region of the polypeptide chain. A search through a database of refined structures (23) showed that the majority of protein segments with similar conformation had a Pro residue either at, or adjacent to, this position.

Tyr 98S → Asn. Residue 98S is a Phe in spinach and is in the middle of an internal hydrophobic core in S and does not interact directly with any L residue (Figure 3.f). In addition to residue 98S this core is comprised of Phe 12S (Tyr in Anabaena), Pro 120S, Tyr 123S and the β and γ carbons of Lys 119S.

Phe 12S and the nearby residue Phe 50S contribute to interactions with L subunit B. The main chain atoms of Phe 12S and

the Cβ and Cγ atoms of Phe 50S are near (within 5 Å) the main chain atoms of residues 232L to 234L (Figure 3.f). Accordingly, the accessibility of residues 12S and 50S decreases upon association of subunits (Table 1). Introduction of an Asn at position 98S would increase the polarity of the region and could alter the core such that the subunit interactions involving residues 12S and 50S are adversely affected.

In S from Silene pratensis, residues 98S and 50S are histidines (33) and therefore have polar side chains in this core. Compared to other organisms, however, the polypeptide chain of S from this organism is shorter, and residue 123S is absent. The effect of this evolutionary change, as modeled by removal of Tyr 123S in the spinach structure, would be to open up the core and make it accessible to solvent. It is therefore logical that, in concert with the shorter chain, some of the hydrophobic residues in this region would have evolved into more polar residues in this organism.

CONFORMATION AND ASSEMBLY

The preceding discussions have as a premise the near-normal folding of the mutant small subunits. In the experimental system used, it could not be determined whether the substituted proteins differed in overall structure or stability or, if there were a difference, whether it was a cause or a consequence of the observed failure to assemble into holoenzyme. To address the possibility that one or more of the substitutions might cause mis-folding of the protein resulting in a structure grossly different from the wild-type small subunit, the algorithm of Chou and Fasman (25) was used to predict secondary structures in the wild-type protein and in each of the

substituted proteins. The predictions for the substituted proteins were not significantly different from the predictions for the wild-type protein, and the structures predicted corresponded well to the structures observed in the spinach protein (results not shown). This analysis suggests that the observed lack of assembly was due to disruption of specific interactions rather than gross conformational changes.

It has recently been shown that assembly of cyanobacterial RuBisCO in E. coli requires the E. coli groEL and groES gene products (34). GroEL protein shows extensive sequence similarity to the RuBisCO binding protein which is found in plant chloroplasts and which is probably involved in assembly of RuBisCO in green plants (35). Although it is possible that the substitutions described in this work prevent holoenzyme assembly by interfering with a necessary (but undemonstrated) interaction of the E. coli groEL or groES with the S subunit, we consider this interpretation to be unlikely in light of the ability of isolated S to assemble in vitro with pre-assembled L octamers.

Recent experiments have identified a region of higher plant S which is important for assembly in higher plant RuBisCO (36). This region, which includes residues 49S to 54S of the spinach protein, is missing in S from cyanobacteria.

OTHER MUTATIONS

Other previously reported S mutants have all been shown to have catalytic activity, and, by implication, to form a stable holoenzyme. A Trp 67S→Phe substitution in Anacystis nidulans rbcS (7, referred to as Trp 54) was shown to reduce V_{max} 2.5-fold without altering

either the K_m or the partition coefficient (relative substrate specificity for CO_2 and O_2) of the enzyme. This result is in contrast to our result for an Arg substitution at position 67S, but fits well with our explanation that it is the additional positive charge introduced by the Arg substitution, rather than the removal of the Trp, per se, which prevented assembly. Altered interaction with the large subunit, despite ability to assemble on a gross level, could lead to the reduction in V_{max} seen in the Phe substituted enzyme. Tryptophan 70S was also changed to Phe (7, referred to as Trp 57) with results very similar to those found for Trp 67S→Phe. Mutations resulting in pairs of substitutions in amino acids 10S, 11S, 14S, 16S, 19S, and 20S were shown to lead to reduced activity in crude preparations (8) but the amount of holoenzyme was not quantified. It is possible then, that these mutations result in an enzyme of reduced activity which, as with Trp70S→Phe and Trp67S→Phe, could be due to deleteriously altered L-S interactions.

In the work presented here, we have identified specific residues widely separated in the primary sequence which play structurally important roles in the small subunit and its interaction with the large subunits. Disruption of the intrasubunit and intersubunit interactions of these and neighboring residues by amino acid substitutions can completely prevent formation of stable holoenzyme. In order to further understand assembly as well as stability of the holoenzyme, similar analyses of a larger number of mutations are underway.

REFERENCES

1. Miziorko, H. M. & Lorimer, G. H. (1983) Ann. Rev. Biochem. 52, 507-535.

2. Andrews, T. J. (1988) J. Biol. Chem. 263, 12213-12219.

3. Gatenby, A. A., van der Vies, S. M. & Bradley, D. (1985) Nature 314, 617-620.

4. Christeller, J. T., Terzaghi, B. E., Hill, D. F. & Laing, W. A. (1985) Plant Mol. Biol. 5, 257-263.

5. Tabita, F. R. & Small, C. L. (1985) Proc. Natl. Acad. Sci. USA 82, 6100-6103.

6. Gurevitz, M., Somerville, C. R. & McIntosh, L. (1985) Proc. Natl. Acad. Sci. USA 82, 6546-6550.

7. Voordouw, G., de Vries, P. A., van den Berg, W. A. M. & de Clerck, E. P. J. (1987) Eur. J. Biochem. 163, 591-598.

8. McFadden, B. A. & Small, C. L. (1988) Photosynthesis Res. 18, 245-260.

9. Knight, S., Andersson, I. & Branden, C. I. (1989) Science 244, 702-705.

10. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning (Cold Spring Harbor Laboratory, NY)

11. Herrero, A., Elhai, J., Hohn, B. & Wolk, C. P. (1984) J. Bacteriol. 160, 781-784.

12. Myers, R. M., Lerman, L. S. & Maniatis, T. (1985) Science 229, 242-247.

13. Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) Gene 33, 103-119.

14. Vieira, J. & Messing, J. (1987) Methods Enzymol. 153, 3-11.

15. Nakamaye, K. L. & Eckstein, F. (1986) Nucleic Acids Res. 14, 9679-9698.

16. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.

17. Elthon, T. E. & McIntosh, L. (1987) Proc. Natl. Acad. Sci. USA 84, 8399-8403.

18. Larson, E., Howlett, B. & Jagendorf, A. (1986) Anal. Biochem. 155, 243-248.

19. Berhow, M. A., Saluja, A. & McFadden, B. A. (1982) Plant Sci. Lett. 27, 51-57.

20. Laemmli, U. K. (1970) Nature 227, 680-685.

21. Harlow, E. & Lane, D. (1988) Antibodies (Cold Spring Harbor Laboratory, NY)

22. Andersson, I., Knight, S., Schneider, G., Lindqvist, Y., Lundqvist, T., Branden, C. I. & Lorimer, G. H. (1989) Nature 337, 229-234.

23. Jones, T. A. & Thirup, S. (1986) EMBO J. 5, 819-822.

24. Lee, B. & Richards, F. M. (1971) J. Mol. Biol. 55, 379-400.

25. Chou, P. Y. & Fasman, G. D. (1978) Adv. Enzymol. 47, 45-148.

26. Mulligan, R. M., Houtz, R. L. & Tolbert, N. E. (1988) Proc. Natl. Acad. Sci. USA 85, 1513-1517.

27. Andrews, T. J. & Lorimer, G. H. (1985) J. Biol. Chem. 260, 4632-4636.

28. Martin, P. G. (1979) Aust. J. Plant Physiol. 6, 401-408.

29. Takruri, I. A. H., Boulter, D. & Ellis, R. J. (1981) Phytochemistry 20, 413-415.

30. Nierzwicki-Bauer, S. A., Curtis, S. E. & Haselkorn, R. (1984) Proc. Natl. Acad. Sci. USA 81, 5961-5965.

31. Andersen, K. & Caton, J. (1987) J. Bacteriol. 169, 4547-4558.

32. Nargang, F., McIntosh, L. & Somerville, C. (1984) Mol. Gen. Genet. 193, 220-224.

33. Smeekens, S., van Oosten, J., de Groot, M. & Weisbeck, P. (1986) Plant Mol. Biol. 7, 433-440.

34. Goloubinoff, P., Gatenby, A. A. & Lorimer, G. H. (1989) Nature 337, 44-47.

35. Hemmingsen, S. M., Woolford, C., van der Vies, S. M., Tilly, K., Dennis, D. T., Georgopoulos, C. P., Hendrix, R. W. & Ellis, R. J. (1988) Nature 333, 330-334.

36. Wasmann, C. C., Ramage, R. T., Bohnert, H. J. & Ostrem, J. A. (1989) Proc. Natl. Acad. Sci. USA 86, 1198-1202.

Chapter 5

A CLASS OF MUTANTS IN THE SMALL SUBUNIT OF RIBULOSE BISPHOSPHATE CARBOXYLASE/OXYGENASE WHICH CAUSE PARTIAL REDUCTION IN HOLOENZYME ASSEMBLY*

ABSTRACT A collection of directed mutations was made in the rbcS gene encoding the small subunit of RbuP2 carboxylase/oxygenase (E.C. 4.1.1.39) from the cyanobacterium Anabaena 7120. The mutant rbcS genes were expressed concommitantly with wild type Anabaena 7120 rbcL (encoding the large subunits) in Escherichia coli, and the altered small subunits produced were examined for their ability to assemble to form stable holoenzyme and for their ability to support catalysis. In addition to mutants in which assembly of stable holoenzyme was completely blocked, and mutants which exhibited normal assembly and activity, the collection also included many mutants in which activity was reduced but not abolished. In these mutants with reduced activity, the amount of small subunit assembled into holoenzyme was also reduced and there was an apparent correlation between the two parameters. The residues substituted in this intermediate class are not localized to a single region, either in the primary sequence, or in the three dimensional structure of the protein. Examination of

* manuscript to be submitted

these and other nearby residues in a model for the crystal structure of spinach $RbuP_2$ carboxylase/oxygenase suggests that the molecular interactions impaired by these substitutions are similar to those disrupted in the mutants in which assembly is completely blocked.

INTRODUCTION

Ribulose 1,5-bisphosphate carboxylase/oxygenase (E.C. 4.1.1.39) catalyzes both the carboxylation and the oxygenation of $RbuP_2$ in the competing reactions of the first committed steps of the reductive and oxidative, respectively, photosynthetic carbon cycles (for review, see Andrews and Lorimer, 1987). The native enzyme in most photosynthetic organisms, including all plants and cyanobacteria, is composed of eight large subunits (L) of molecular weight approximately 55,000 each, and eight small subunits (S) each of molecular weight approximately 15,000. The large subunits provide all of the residues which participate directly in substrate binding and catalysis (Andrews, 1988). The small subunits are believed to force a conformational shift in the large subunits which in some way alters the positioning of active site residues of the large subunits to effect catalytic competence (Andrews, 1988). The residues and the inter-subunit interactions which are required to bring about this conformational shift, however, and their relationship to the residues and interactions required for subunit assembly, have not been identified.

When the gene encoding the small subunits of cyanobacterial $RbuP_2$ carboxylase/oxygenase (rbcS) is expressed concurrently with the

gene encoding the large subunits (rbcL) in E. coli, a functional holoenzyme is assembled in the bacteria (Gatenby et al., 1985; Tabita and Small, 1985; Gurevitz et al., 1985; Christeller et al., 1985). This heterologous expression provides a convenient system for site-directed mutagenesis to allow identification and analysis of small subunit residues essential for association with large subunits or catalysis. In a demonstration of such a system using the rbcS and rbcL genes from the cyanobacterium Anacystis nidulans, two evolutionarily conserved residues were replaced, resulting in enzymes with reduced V_{max} (Voordouw et al., 1987). In a similar system preliminary results from a collection of double mutants substituting residues near the amino terminus of the A. nidulans small subunit suggest that this region of the protein is important for activity (McFadden and Small, 1988). Structural requirements of higher-plant small subunits for assembly into holoenzyme have been studied in vitro with isolated chloroplasts. A putative assembly domain has been identified (Wasmann et al., 1989) but the importance of individual residues has not been assessed.

We have used site-directed mutagenesis of rbcS from the cyanobacterium Anabaena 7120 and expression in E. coli to search for small subunit residues which are important for assembly or catalytic function. Four of the mutants were found to prevent the assembly of stable holoenzyme and have been described previously (Fitchen et al., submitted). In the experiments reported in this paper, we investigated assembly and catalysis in eighteen additional mutants, and describe a new class in which the ability of mutant S to assemble into holoenzyme is intermediate between the four initially described

mutants and wild-type S. Residues homologous to those changed in the mutants were studied in a model for the three dimensional structure of spinach $RbuP_2$ carboxylase/oxygenase (Andersson et al., 1989; Knight et al., 1989) to identify structural interactions of these residues and neighboring residues which would have been disrupted by the substitutions made.

MATERIALS AND METHODS

Materials.

DNA polymerases, restriction and modifying enzymes were from New England Biolabs (Beverly, MA) or Bethesda Research Laboratories (Gaithersburg, MD). $RbuP_2$ and protein A-alkaline phosphatase were from Sigma (St. Louis, MO). 14C-Sodium bicarbonate (58 Ci/mole) was from ICN (Irvine, CA). Protein molecular weight standards were from BioRad (Richmond, CA) and prestained protein molecular weight standards, used in conjunction with western transfers, were from Diversified Biotech (Newton Centre, MA) Nitrocellulose membrane was from Schleicher and Schuell (Keene, NH). Purified spinach $RbuP_2$ carboxylase/oxygenase was a gift from Dr. S. C. Somerville.

Antiserum.

A recombinant protein including residues 33S-99S[2] of the small subunit of $RbuP_2$ carboxylase/oxygenase Anabaena 7120 subunit was used as antigen for raising polyclonal antiserum in rabbits as described (Harlow and Lane, 1988).

Bacterial strains.

E. coli strain TG1 [K12, Δ(*lac-pro*) *supE thi hsdD5* F' (*traD36 proA*$^{+}$*B*$^{+}$ *lacI*^{q} *lacZ* ΔM15)] (from Amersham) was used for preparation of single stranded DNA template for mutagenesis and sequencing. HB101 (*pro leu thi lacY hsdR endA recA rpsL ara galK xyl mtl supE44*) was used in all other experiments. Transformation of bacteria with multiple plasmids was performed serially.

Growth of bacteria.

Bacteria were grown in L-broth (Lennox, 1955) without glucose. Antibiotics were added at the following concentrations when appropriate: ampicillin, 50 mg/l; kanamycin, 30 mg/l. For analysis of the mutants, 300 ml cultures of HB101 carrying the large subunit expression plasmid (pL56) and a mutant-small subunit expression plasmid were inoculated with 1 ml of a fresh overnight culture and grown for 24 hr at 37°C with vigorous aeration.

DNA manipulations and DNA sequencing.

Subcloning of DNA fragments and analysis of DNA was according to standard protocols (Maniatis *et al.*, 1982). DNA restriction and modifying enzymes were used according to the manufacturers instructions. DNA sequence was determined by the dideoxy chain termination method (Sanger *et al.*, 1977).

Site-directed mutagenesis.

Oligonucleotides. Four oligonucleotide mixtures were synthesized by the solid-state phosphorimidate method with an Applied Biosystems

model 380A DNA synthesizer. At several positions in each oligomer, nucleotide-reagent mixtures were used for the synthesis to provide a fraction of the oligonucleotides which would differ from the wild-type sequence at that position. The sequences of the oligonucleotides are shown in Figure 1. The proportion of mutagenic base reagent included in the synthesis at each position was 11% for oligonucleotide mixtures SSm1 and SSm2 and 15% for mixtures SSm3 and SSm4, in order to maximize the yield of oligonucleotides with one and only one difference from the wildtype sequence. Mixtures were designed so that all mutants would result either in the loss or gain of an Adenine in the sequence, allowing screening by sequencing with only the dideoxy-A terminated reaction. The C at position 18 of SSm4 is a T in the wild-type sequence. This silent mutation allowed us to monitor the efficiency of mutagenesis.

DNA template for mutagenesis. Plasmid p205-3 includes the Anabaena 7120 rbcS coding region and the ori region of bacteriophage M13 and has been described (Fitchen et al., submitted). E. coli (strain TG1) containing this plasmid were superinfected with bacteriophage M13K07 and the single stranded form of the plasmid was purified as described (Vieira and Messing, 1987).

In-vitro mutagenesis. In-vitro mutagenesis of single stranded p205-3 was performed by the phosphorothioate-enrichment method as described (Nakamaye and Eckstein, 1986). The four mutagenic oligonucleotide mixtures were used separately.

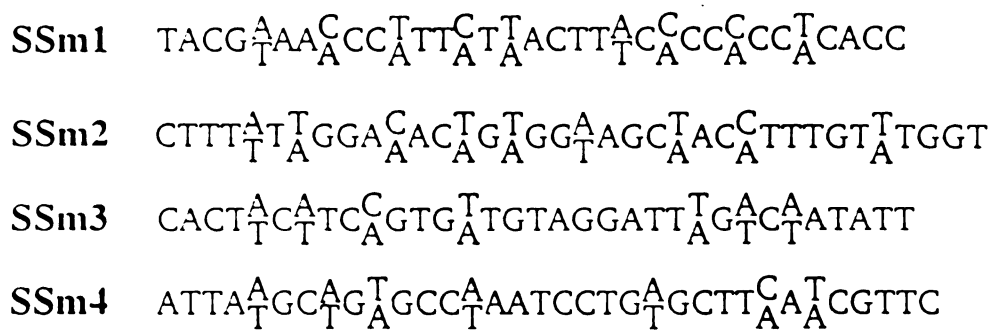


Fig. 1. Sequences of the oligonucleotide mixtures synthesized for mutagenesis of the small subunit. At positions where two bases are shown, the upper base is the base found in the wild type sequence and was included at that position in the majority of the oligonucleotides.

Screening. Mutants were identified by sequencing of randomly selected progeny plasmids from the mutagenesis protocol. Initial screening was by sequencing the appropriate region using only the dideoxy-A-terminated reaction. The entire rbcS cassette was sequenced (using all four termination reactions) in putative single mutants to exclude the possibility of additional, spurious mutations.

Prediction of secondary structures.

Secondary structures in wild-type and mutant small subunits were predicted using the algorithm of Chou and Fasman (Chou and Fasman, 1978).

Plasmids directing expression of $RbuP_2$ carboxylase/oxygenase subunits.

The construction of the plasmids expressing non-mutant subunits has been described previously (Fitchen et al., submitted). XbaI-EcoRI restriction fragments from p205-3 and its mutant derivatives were subcloned into pUC19 (Yanisch-Perron et al., 1985) such that expression of rbcS was under the transcriptional control of the E. coli lac promotor. In the resulting plasmids , the rbcS coding region is not in the reading frame of the lacZ fragment of the vector and translation of rbcS message begins at the initial methionine of authentic small subunit. Plasmid pL56, from which expression of Anabaena 7120 rbcL is under lac control, uses the bacteriophage lambda origin of replication and can be stably co-maintained with the pUC19-derived small subunit expression plasmids.

Analysis of activity.

The measurement of carboxylase activity in extracts of E. coli expressing large subunits and mutant small subunits was as described (Fitchen et al., submitted; Pierce et al., 1982). Activity was assayed in crude fractions consisting of protein precipitating at 40% (w/v) ammonium sulfate in a 100,000 x g supernatant of ultrasonically disrupted cells. Protein was quantified as described (Lowry et al., 1951 as modified in Larson et al., 1986).

Analysis of assembly.

The extent of assembly of mutant small subunits into holoenzyme was determined as described for the initial four mutants in this collection (Fitchen et al., submitted). Unassembled small subunits were separated from holoenzyme by sedimentation in sucrose gradients (Fitchen et al., submitted, modified from Berhow et al., 1982). Proteins sedimenting at or near the position in the gradient where wild type holoenzyme sedimented (as estimated by sedimentation of spinach $RbuP_2$ carboxylase/oxygenase in equivalent gradients) were resolved by SDS-PAGE (buffer system of Laemmli, 1970 using a 5% (wt/vol) stacking gel and a 10-17.5% (wt/vol) polyacrylamide gradient resolving gel) and transferred to nitrocellulose filters (Towbin et al., 1979). These western blots were incubated with polyclonal antibodies specific to the small subunit. Visualization of bound antibody after decoration with protein A-alkaline phosphatase was as described (Blake et al., 1984).

Structural interpretation of mutants.

Possible structural effects of the mutations were investigated by inspection of a model for the three- dimensional structure of spinach $RbuP_2$ carboxylase/oxygenase. The model was derived from an isomorphous electron density map with phase angles revised by four-fold averaging and has been briefly described (Andersson et al., 1989; Knight et al., 1989). Residues subjected to mutagenesis, other residues in their vicinities, and the location of the substituted residues in the context of the holoenzyme were examined on a graphics display using FRODO (Jones, 1985). Mutants were modeled by substitution of appropriate side chains.

RESULTS

Comparison of sequences of the small subunit of $RbuP_2$ carboxylase/oxygenase from many organisms reveals several regions of evolutionary sequence conservation (Andrews and Lorimer, 1987). Residues in each of three of the largest regions were targeted for mutagenesis. The mutations were designed so that the resulting collection of mutant proteins would include substitutions of dissimilar amino acids as well as conservative changes. Twenty-two mutants were recovered, resulting in substitutions at 23 of the 109 residues in the Anabaena small subunit. Four of the mutations have been described in a previous report (Fitchen et al., submitted); the locations of the eighteen mutants described in this paper, and the sequence of S, are shown in Figure 2.

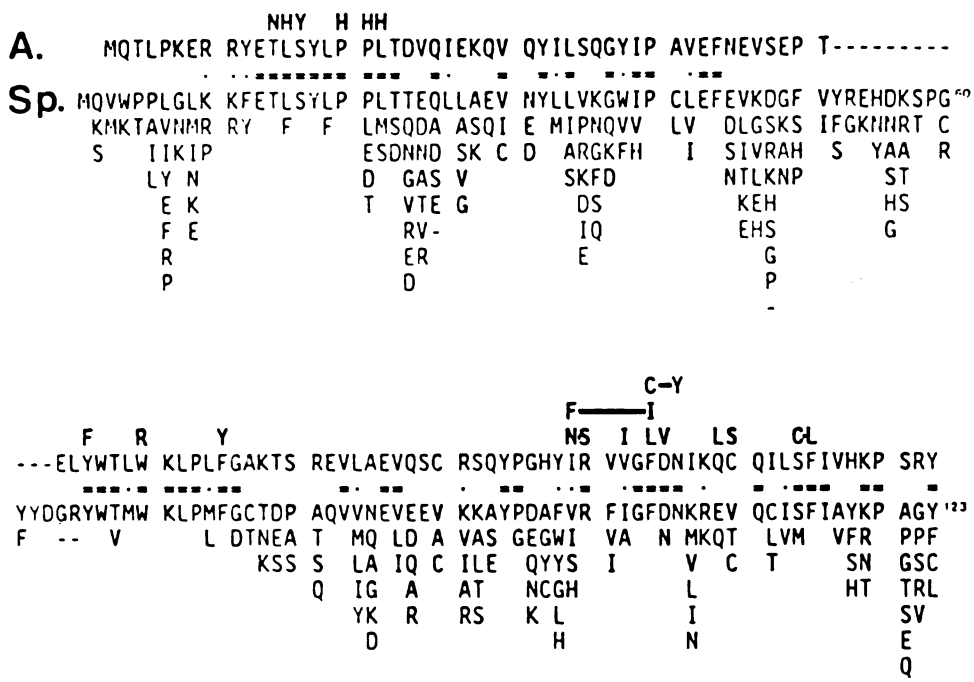


Fig. 2. Location of substituted residues in the primary sequence. The substitutions resulting from the mutants described in this paper are shown above the deduced amino acid sequence of the wild-type small subunit of RbuP$_2$ carboxylase/oxygenase from Anabaena 7120 (Nierzwicki-Bauer et al., 1984). Pairs of residues substituted in double mutants are connected by solid lines. The amino acid sequence of the small subunit of the enzyme from spinach (Martin, 1979, corrected in Takruri et al., 1981 and Knight et al., 1989) is shown below the Anabaena sequence. Identical residues in the two sequences are indicated by -; similar residues are indicated by·. Other residues occurring at each position in a collection of 16 other higher plants and one other cyanobacterium for which deduced small subunit sequences have been reported are shown below the spinach sequence. Sequences are from Adams et al., 1987; Baszczynski et al., 1988; Bedbrook et al., 1980; Berry-Lowe et al., 1982; Broglie et al., 1983; Greenland et al., 1987; Lebrun et al., 1987; Mazur and Chui, 1985; McKnight et al., 1986; Shinozaki and Sugiura, 1983; Smeekens et al., 1986; Smith et al., 1983; Steikema et al., 1983;Tumer et al., 1986; Waksman and Freyssinet, 1987; Xie and Wu, 1988; and Yamamoto et al., 1988.

Possible effects of the substitutions on the folding of the mutant proteins were predicted using the algorithm of Chou and Fasman (Chou and Fasman, 1978) to predict secondary structures in the wild-type small subunit and in each mutant small subunit (results not shown). The predictions for the wild-type Anabaena small subunit agreed well with the structures observed in the spinach small subunit, predicting all four beta-strands and one out of two alpha-helices. Proteins with substitutions of Pro 19S^2→His, Trp 70S →Arg, Phe 75S →Tyr, or Phe 104S →Cys and Asn 106S →Tyr were predicted to have structures differing from wild-type in the vicinities of the substitutions. Proteins with substitutions of Thr 14S →Asn, Leu 15S →His, Ile 99S→Phe and Phe 104S →Ile, or Ile 99S →Asn and Arg 100S →Ser, were predicted to have slight changes in the length of one or two structures in each case.

The effects of the substitutions on activity and assembly were studied using expression of L and mutant S in E. coli. The plasmids directing expression of wild-type Anabaena 7120 $RbuP_2$ carboxylase/oxygenase L and S subunits have been previously described, as has the demonstration of functional enzyme in extracts of E. coli transformed with these plasmids (Fitchen et al., submitted). DNA fragments containing the mutant rbcS coding regions were subcloned into pUC19 to create plasmids analogous to the wild-type S-expressing plasmid and the effect of the mutations on activity and assembly, when expressed in E. coli in conjunction with the large subunit, was determined.

Catalytic activity was measured as $RbuP_2$ dependent incorporation of radioactivity from ^{14}C-bicarbonate into acid-stable forms (see

Methods). Oxygenase activity was not measured because the relative substrate specificity for O_2 and CO_2 is a property determined solely by the large subunits (Andrews and Lorimer, 1985). Because our earlier studies suggested that assembly was likely to be impaired in some mutants (Fitchen et al., submitted), activity was measured on relatively crude preparations to avoid the possible loss of an aberrant, but still partially assembled, enzyme. A very low level of activity (up to 40 pmol CO_2 fixed/min/mg protein) was detected in extracts from control E. coli not expressing small subunits. This was probably due to residual activity of L octamers (Andrews, 1988) and set the lower limit of detectable activity in our system at 0.2% of the activity in wild-type extracts. The levels of carboxylase activity in extracts with the mutant small subunits are shown and compared to the level in wild-type extract in Figure 3. A complete range of activities was found for the different mutants, extending from several which did not have any detectable activity, to three in which activity was higher than wild-type.

Assembly of the mutant small subunits into holoenzyme was investigated by separation of assembled holoenzyme from unassembled small subunit on sucrose density gradients. Small subunit protein sedimenting at or near the position of wild-type holoenzyme was resolved by SDS-PAGE and detected by western blotting. The sensitivity of our detection was such that S could be detected at levels less than 2% of those found with wild-type S (data not shown). The relevant regions of the western blots are shown in Figure 3. Many of the mutant extracts were found to contain reduced or non-existant levels of S assembled into holoenzyme. Inspection of

Fig. 3. Carboxylase activity and extent of assembly of holoenzyme in extracts containing mutant small subunits. Activity is expressed as nmol CO_2 fixed/min/mg protein and was assayed on a crude extract of dually transformed (expressing the large subunit and native or substituted small subunits of $RbuP_2$ carboxylase/oxygenase) E. coli. The activities are the mean (+/- S. D.) calculated from three separate experiments. The activities are also expressed as a percent of the activity measured in wild-type extracts. The last two columns are sections of western blots of SDS-PAGE gels after immunodetection of small subunits. The samples for the lanes in the 18S column were taken from sucrose gradients used to separate holoenzyme from unassembled small subunits and contained equal volumes of gradients which had been loaded with equal amounts of protein. The pellet samples were from unwashed 10,000 x g pellets of the ultrasonically disrupted cells and were prepared by partial solubilization in boiling SDS-PAGE sample buffer (Laemmli, 1970).

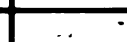
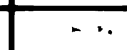
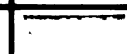
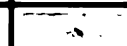
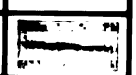
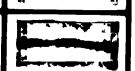
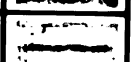
Substitution	Activity	%	18S	Pellet
None (wt)	51 ± 17	100		
Thr 14S → Asn	6.4 ± 4.8	13		
Leu 15S → His	0.4 ± 0.6	0.9		
Ser 16S → Tyr	80 ± 44	160		
Pro 19S → His	9.6 ± 13	19		
Pro 20S → His	16 ± 11	31		
Leu 21S → His	nd	-		
Tyr 66S → Phe	60 ± 48	120		
Trp 70S → Arg	nd	-		
Phe 75S → Tyr	11 ± 4.4	22		
Ile 99S → Phe Phe 104S → Ile	nd	-		
Ile 99S → Asn Arg 100S → Ser	nd	-		
Val 102S → Ile	55 ± 49	110		
Phe 104S → Leu	18 ± 1.0	36		
Phe 104S → Cys Asn 106S → Tyr	4.8 ± 2.0	9		
Asp 105S → Val	nd	-		
Gln 109S → Leu	1.0 ± 1.1	2		
Cys 110S → Ser	18 ± 3.9	36		
Ser 114S → Cys Phe 115S → Leu	0.8 ± 0.2	2		

Figure 3

the results shows an apparent correlation between activity and assembly. That the reduced levels of activity and assembled S are not due to failure of the expression system can be seen from the last column of Figure 3 which shows the presence of mutant S in unwashed pelleted material from crude extracts of the mutant cultures.

DISCUSSION

Several previously reported examples prove that the correlation, observed for the mutants made here, between extent of assembly and level of activity is not absolute. Mutations resulting in Trp to Phe substitutions at residues 67S[2] and 70S [Voordouw *et al.*, 1987 (designated therein as residues 54 and 57)] of the small subunit from the cyanobacterium *Anacystis nidulans* were made and the mutant protein was expressed with *A. nidulans* large subunits in *E. coli*. The V_{max} of $RbuP_2$ carboxylase/oxygenase containing these substituted small subunits were found to be 44% and 37% (respectively) of wild type. A hybrid enzyme, assembled *in vitro* from *Synechococcus* large subunits and spinach small subunits, which is similar to introducing about 65 mutations simultaneously into the cyanobacterial small subunit sequence[3], was found to have a V_{max} about half that of the non-hybrid reconstituted *Synechococcus* enzyme (Andrews and Lorimer, 1985). Preliminary data for additional mutations in *rbcS* of *A. nidulans* which result in decreased levels of activity in crude extracts have been reported (McFadden and Small, 1988). Each of the mutants was found to have at least a low level of catalytic activity,

but the impact of these mutations on assembly has not been assessed. That there are no demonstrated examples of small subunit mutants which completely block activity without concommittantly blocking assembly, however, suggests that many of the amino acid residues and intersubunit interactions required to bring about the conformational shift in the large subunits leading to full activity are the same as those required for subunit association.

It is possible that the mutant small subunits were incorrectly folded and thus unable to associate with the large subunits for gross structural reasons. Additionally, and possibly as a result of incorrect folding, some of the mutant small subunits may have been more succeptible to proteolytic degradation by intracellular E. coli proteases. We addressed the question of incorrect folding, indirectly, by comparing predicted secondary structures in wild-type and mutant small subunits. Markedly altered secondary structures, which could lead to aberrant folding, were predicted for only four of the mutant proteins, only one of which was a mutant lacking activity. Demonstration of intact small subunit protein in crude extracts showed that even if proteolysis was accelerated in some cases, it was not complete in any mutant.

To explore the possibility that the different classes of mutants could be due to the mutations being in different domains of the protein, we examined a model for the three dimensional structure of spinach $RbuP_2$ carboxylase/oxygenase. The model is based on x-ray crystallographic data from the spinach enzyme and has been described (Andersson et al., 1989; Knight et al., 1989). Although the sequences of small subunits from Anabaena and spinach are only 41

percent identical, the residues in the regions which we targeted for mutagenesis are more highly conserved (see Figure 2), and the secondary and tertiary structures of the small subunit are expected to be highly conserved. The three dimensional location of the residues mutated to produce each of the three classes of mutants is shown in the context of the structure of the small subunit in Figure 4. The different classes do not appear to be segregated to different regions of the protein. The lack of clustering of the mutants blocked or impaired in assembly to any one area of the small subunit suggests that determinants of association are not restricted to an "assembly domain" in the cyanobacterial small subunit, although some caution must be exercised in drawing conclusions as only limited regions of the protein were subjected to mutational analysis. This distribution contrasts with results recently reported (Wassman et al., 1989) suggesting that a region near the middle of the primary sequence of the higher plant small subunit constitutes an "assembly domain". Interestingly, the region identified in their experiments represents the most extensive sequence divergence between cyanobacterial and plant $RbuP_2$ carboxylase/oxygenases, and is not present in the cyanobacterial protein (see Figure 2).

In examining the three-dimensional structure in more detail to try to understand how the substitutions resulting in the intermediate class of mutants could interfere with assembly without entirely preventing it, we looked for two things: first, we wanted to identify what interactions with other residues would be disrupted by the substitution; second, we tried to roughly evaluate the importance of these disrupted interactions to L-S association in

Fig. 4. Location of mutants belonging to each of three classes in a model for the three-dimensional structure of the small subunit. The model was derived from x-ray crystallographic data for the spinach enzyme. A, The small subunit (bold lines) shown in relation to two dimers of large subunits. In the complete holoenzyme two more dimers are located directly behind the two shown here. The small subunits are situated in the crevices between large subunit dimers, with four clustered at the top and four clustered at the bottom. B, Stereo-drawing showing the position of residues which, when substituted, blocked carboxylase activity and prevented the assembly of holoenzyme (or reduced it to below 2% of wild-type which was the limit of detection in our system). Filled circles are positions of substituted residues in single mutants (13S, 15S, 21S, 67S, 70S, 73S, 98S, and 105S). Open circles are substituted residues in double mutants (99S, 100S, and 104S). C, Stereo drawing showing the position of residues which, when substituted, resulted in holoenzyme and activity levels between 2% and 70% of wild-type. Filled circles (residues substituted in single mutants) identify positions 14S, 19S, 20S, 75S, 104S, 109S, and 110S. Open circles (residues substituted only in double mutants) identify positions 106S, 115S, and 116S. D, Stereo drawing as in B and C showing location of those residues which, when substituted, did not have a deleterious effect on assembly or activity (16S, 66S and 102S).

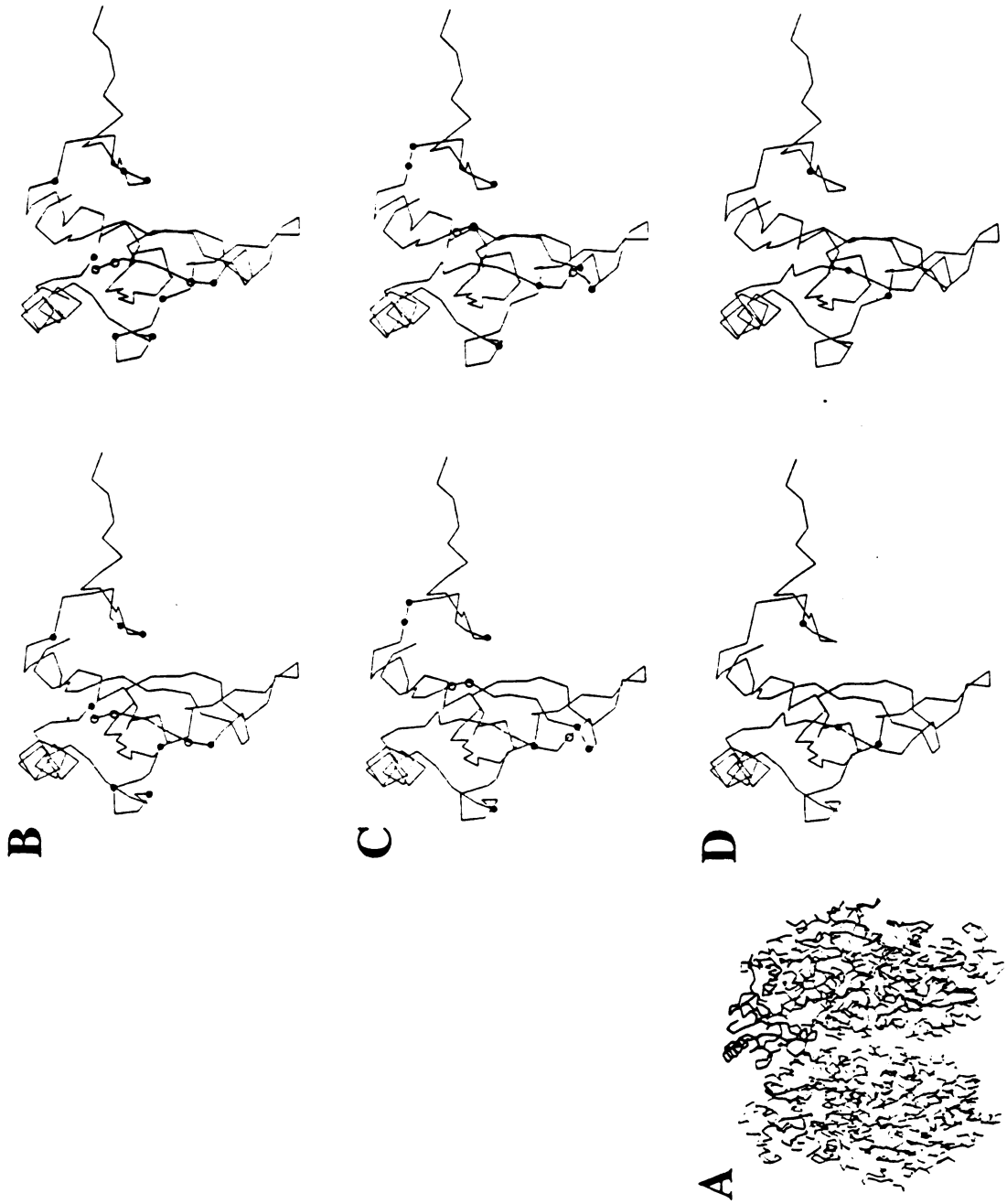


Figure 4

order to understand the observed difference between these mutants and mutants in which formation of stable holoenzyme was completely blocked. Two mutations for which these two criteria are particularly well met are the mutations resulting in Phe 75S to Tyr and Phe 104S to Leu which reduced assembly by approximately 80% and 40% respectively. In the wild type, both of these evolutionarily conserved hydrophobic residues contribute to the same hydrophobic patch on the surface of the small subunit as shown in Figure 5a. This patch makes an important hydrophobic interaction with the large subunit (especially leucine 74L) and its disruption by substitution of His for Pro 73S completely prevented assembly of stable holoenzyme (Fitchen et al., submitted). In the case of Phe 75S→Tyr the addition of a hydroxyl group at the C(zeta) position of the existing phenylalanine would increase the polarity of the region, making it less suitable for interaction with the hydrophobic Leu 74L. Alternatively, the location of the residue near the solvent-exposed surface of the protein in the holoenzyme would allow the side chain to rotate around the Ca-Cb bond, effectively removing the residue from the hydrophobic patch as represented in the model in Figure 5b. This rotation would, however, decrease the hydrophobic surface available for interaction with Leu 74L. In the case of Phe 104S Leu, the substitution does not require repositioning of the side chain, but the smaller size of the Leu side chain reduces the area of the "hydrophobic target" presented to Leu 74L and would be predicted to decrease the strength of the hydrophobic interaction. Neither of these two substitutions, however, would be predicted to have as much of an effect on the hydrophobicity of this microenvironment as would

Fig. 5. Drawings from model for the three-dimensional structure of RbuP$_2$ carboxylase/ oxygenase. A, stereo drawing showing the intersubunit hydrophobic pocket in which residues Phe 75S and Phe 104S are situated. All atoms shown in the upper-left belong to the large subunit; the residues in the lower-right are from the small subunit. B, Possible structural effect of a Phe (dotted lines) to Tyr (bold lines) substitution at position 75S. The orientation of the Tyr side-chain shown here is arbitrary. C, Structural representation of Phe (bold and dotted lines) to Leu (bold lines) substitution at position 104S. Local rearrangements of neighboring residues which might occur in response to the substitution are not shown.

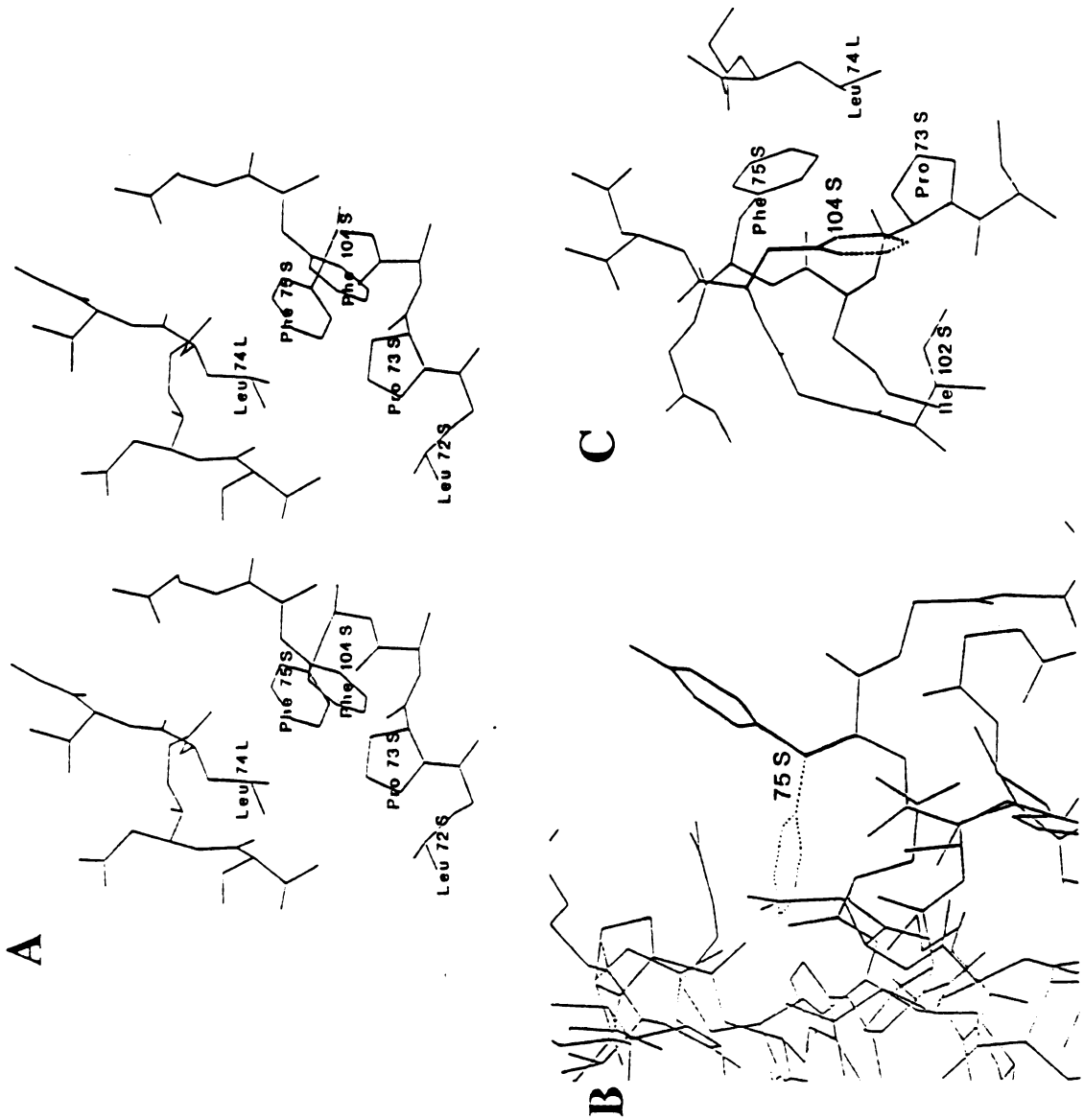

A
Leu 74 L
Phe 75 S
Phe 104 S
Pro 73 S
Leu 72 S
Leu 74 L
Phe 75S
Phe 104 S
Pro 73 S
Leu 72 S
B
75 S
C
Leu 74 L
Phe 75S
104 S
Pro 73 S
Ile 102 S

Figure 5

the Pro 73S →His substitution. Their less drastic impact on the suitability for interaction with Leu 74L could be responsible for the intermediate phenotype of these mutants. Of the two, the Phe 75S→ Tyr substitution would be predicted to have the greater effect on hydrophobicity, and the reduction in assembly caused by this mutation is observed to be greater than the reduction for Phe 104S Leu.

In addition to demonstrating that many of the residues of the small subunit are important for assembly of RbuP2 carboxylase/oxygenase, the results of the mutants in this collection are significant in that they reflect the pathway of holoenzyme assembly. In higher-plant chloroplasts, assembly of $RbuP_2$ carboxylase/oxygenase is mediated by the binding of a nucleus-encoded protein to the large subunits (Barraclough and Ellis, 1980; Cannon et al., 1986; Ellis and van der Vies, 1988) This protein is a member of a recently described class of proteins involved in macromolecular assembly called chaperonins (Hemmingsen et al., 1988). The presence of a chaperonin in cyanobacteria has not been demonstrated (Mcfadden and Torres-Ruiz, 1988), but a recent experiment showed that the E. coli chaperonin (comprised of the GroES and GroEL gene products and homologous to the chloroplastic chaperonin) is required for assembly of $RbuP_2$ carboxylase/oxygenase subunits from cyanobacterial rbcL and rbcS genes expressed in E coli (Goloubinoff et al., 1989). Our explanations for how the Phe 75S to Tyr and Phe 104S to Leu substitutions could interfere with assembly, and consequently catalysis, have as their basis the disruption of association of the small subunit with the large subunit. Inspection of several of the other mutations presented here suggested that they too may result in

reduced assembly through interfering with docking of the small subunit (structures not shown). Additionally, contact site analysis provided plausible explanations for the failure to assemble of the Glu 13S →Val, Trp 67S →Arg, Pro 73S →His, and Tyr 98S →Asn mutants from this collection which were analyzed previously (Fitchen et al., submitted). Further, we were not able to identify a group of mutations whose position on the small subunit would suggest a binding site or interface for interaction with another protein. Together these results suggest that the determinants of incorporation of the small subunits into holoenzyme are resticted to the small and large subunits, and the structural interactions between them. This conclusion is consistent with the ability of isolated small subunits to reassociate with large subunit octamers in vivo without involvement of other proteins (Andrews and Abel, 1981), and supports the involvement of chaperonin being limited to the assembly of large subunit octamers in chaperonin-assisted assembly of $RbuP_2$ carboxylase/oxygenase.

REFERENCES

Adams, C. A., Babcock, M., Leung, F., and Sun, S. M. (1987) Nucleic Acids Res. 15, 1875.

Andersson, I., Knight, S., Schneider, G., Lindqvist, Y., Lundqvist, T., Branden, C.-I., and Lorimer, G. H. (1989) Nature 337, 229-234.

Andrews, T. J. (1988) J. Biol. Chem. 263, 12213-12219.

Andrews, T. J., and Abel, K. M., (1981) J. Biol. Chem. 256, 8445-8451.

Andrews, T. J., and Lorimer, G. H. (1985) J. Biol. Chem. 260, 4632-4636.

Andrews, T. J., and Lorimer, G. H. (1987) in The Biochemistry of Plants, (Hatch, M. D. and Boardman, N. K., eds) vol. 10, pp. 131-218, Academic Press, San Diego.

Barraclough, R. and Ellis, R. J. (1980) Biochim. Biophys. Acta 608, 19-31.

Baszczynski, C. L., Fallis, L., and Bellemare, G. (1988) Nucleic Acids Res. 16, 4732.

Bedbrook, J. R., Smith, S. M., and Ellis, R. J. (1980) Nature 287, 692-697.

Berhow, M. A., Saluja, A., and McFadden, B. A. (1982) Plant Sci. Lett. 27, 51-57.

Berry-Lowe, S. L., McKnight, T. D., Shah, D. M., and Meagher, R. B. (1982) J. Mol. Appl. Genet. 1, 483-498.

Blake, M. S., Johnston, K. H., Russell-Jones, G. J., and Gotschlich, E. C. (1984) Anal. Biochem. 136, 175-179.

Broglie, R., Coruzzi, G., Lamppa, G., Keith, B., and Chua, N.-H. (1983) Bio/Technology 1, 55-61.

Cannon, S., Wang, P., and Roy, H. (1986) J. Cell Biol. 103, 1327-1335.

Chou, P. Y. and Fasman, G. D. (1978) Adv. Enzymol. 47, 45-148.

Christeller, J. T., Terzaghi, B. E., Hill, D. F., and Laing, W. A. (1985) Plant Mol. Biol. 5, 257-263.

Ellis, R. J., and van der Vies, S. M. (1988) Photosynthesis Res. 16, 101-115.

Fitchen, J. H., Knight, S., Andersson, I., Branden, C.-I., and McIntosh, L. (1989?) manuscript submitted.

Gatenby, A. A., van der Vies, S. M., and Bradley, D. (1985) Nature 314, 617-620.

Goloubinoff, P., Gatenby, A. A., and Lorimer, G. H. (1989) Nature 337, 44-47.

Greenland, A. J., Thomas, M. V., and Walden, R. M. (1987) Planta 170, 99-110.

Guidet, F., and Fourcroy, P. (1988) Nucleic Acids Res. 16, 2336.

Gurevitz, M., Somerville, C. R., and McIntosh, L. (1985) Proc. Natl. Acad. Sci. USA 82, 6546-6550.

Harlow, E., and Lane, D. (1988) Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Hemmingsen, S. M., Woolford, C., van der Vies, S. M., Tilly, K., Dennis, D. T., Georgopoulos, C. P., Hendrix, R. W., and Ellis, R. J. (1988) Nature 333, 330-334.

Jones, T. A. (1985) Methods Enzymol. 115, 157-171 .

Knight, S., Andersson, I., and Branden, C.-I. (1989) Science 244, 702-705.

Laemmli, U. K. (1970) Nature 227, 680-685.

Larson, E., Howlett, B., and Jagendorf, A. (1986) Anal. Biochem. 155, 243-248.

Lebrun, M., Waksman, G., and Freyssinet, G. (1987) Nucleic Acids Res. 15, 4360.

Lennox, E. S. (1955) Virology 1, 190-206.

Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J. (1951) J. Biol. Chem. 193, 265-275.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Martin, P. G. (1979) Aust. J. Plant Physiol. 6, 401-408.

Mazur, B. J., and Chui, C. F. (1985) Nucleic Acids Res. 13, 2373-2386.

McFadden, B. A., and Small, C. L. (1988) Photosynthesis Res. 18, 245-260.

McKnight, T. D., Alexander, D. C., Babcock, M. S., and Simpson, R. B., (1986) Gene 48, 23-32.

Nakamaye, K. L., and Eckstein, F. (1986) Nucleic Acids Res. 14, 9679-9698.

Nierzwicki-Bauer, S. A., Curtis, S. E., and Haselkorn, R. (1984) Proc. Natl. Acad. Sci. USA 81, 5961-5965.

Pierce, J. W., McCurry, S. D., Mulligan, R. M., and Tolbert, N. E. (1982) Methods Enzymol. 89, 47-55.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.

Shinozaki, K., and Sugiura, M. (1983) Nucleic Acids Res. 11, 6957-6964.

Smeekens, S., van Oosten, J., de Groot, M., and Weisbeck, P. (1986) Plant Mol. Biol. 7, 433-440.

Smith, S. M., Bedbrook, J., and Speirs, J. (1983) Nucleic Acids Res. 11, 8719-8734.

Steikema, W. J., Wimpee, C. F., and Tobin, E. M. (1983) Nucleic Acids Res. 11, 8051-8061.

Tabita, F. R., and Small, C. L. (1985) Proc. Natl. Acad. Sci. USA 82, 6100-6103.

Takruri, I. A. H., Boulter, D., and Ellis, R. J. (1981) Phytochemistry 20, 413-415.

Torres-Ruiz, J. A., and McFadden, B. A. (1988) Arch. Biochem. Biophys. 261, 196-204.

Towbin, H., Staehelin, T., and Gordon, J. (1979) Proc. Natl. Acad. Sci. USA 76, 4350-4354.

Tumer, N. E., Clark, W. G., Tabor, G. J., Hironaka, C. M., Fraley, R. T., and Shah, D. M. (1986) Nucleic Acids Res. 14, 3325-3342.

Vieira, J., and Messing, J. (1987) Methods Enzymol. 153, 3-11.

Voordouw, G., de Vries, P. A., van den Berg, W. A. M., and de Clerck, E. P. J. (1987) Eur. J. Biochem. 163, 591-598.

Waksman, G., and Freyssinet, G. (1987) Nucleic Acids Res. 15, 1328.

Wasmann, C. C., Ramage, R. T., Bohnert, H. J., and Ostrem, J. A. (1989) Proc. Natl. Acad. Sci. USA 86, 1198-1202.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Gene 33, 103-119.

Xie, Y., and Wu, R. (1988) Nucleic Acids Res. 16, 7749.

Yamamoto, N., Kano-Murakami, Y., Matsuoka, M., Ohashi, Y., and Tanaka, Y. (1988) Nucleic Acids Res. 16, 11830.

FOOTNOTES

footnote 1

The abbreviations used are: $RbuP_2$, ribulose 1,5-bisphosphate; S, small subunit of $RbuP_2$ carboxylase/oxygenase; rbcS, gene encoding S; L, large subunit of $RbuP_2$ carboxylase/oxygenase; rbcL, gene encoding L; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

footnote 2

Residue numbering used throughout this paper is that of the spinach enzyme. Large subunit residue numbers are followed by an L; small subunit residue numbers are followed by an S.

footnote 3

The sequence of the small subunit from Synechococcus ACMM323, used to make the hybrid enzyme is not known. The deduced amino acid sequence of the small subunit from Synechococcus PCC6301 (Shinozaki and Sugiura, 1983) was used for comparison with the spinach sequence (Martin, 1979 as corrected in Takruri et al., 1981 and Knight et al., 1989).

Chapter 6

CONCLUDING REMARKS

INTRODUCTION

In this dissertation I have described the development of plasmid systems for expression of RuBisCO subunits in E. coli, and have presented the results of experiments making use of one of these systems to study the function of individual residues in the small subunit. In this chapter, the major conclusions from my experiments are listed and the utility of these, and similar, systems is reviewed. Directions in which future investigation of the function of the small subunit and its individual residues could progress beyond the experiments reported in the previous two chapters are discussed.

RESULTS FROM EXPERIMENTS WITH RuBisCO EXPRESED IN E. coli.

System expressing Z. mays large subunit.

- Analysis of extracts of cells bearing a Z. mays RuBisCO large subunit expression plasmid showed that large subunit protein could be correctly synthesized and accumulate to high levels in E. coli.

- Results obtained with this expression system suggested that, if the large subunit were proteolytically processed *in vivo* the size of the removed peptide was less than previously hypothesized (1). Subsequent studies have confirmed this by demonstrating that the mature protein sequence begins at the third encoded residue (2).
- Analysis of the solubility of the large subunit produced, showed that the majority of the protein was in a rapidly sedimenting fraction of the bacterial extracts. This was true for several clones, expressing varying levels of large subunit protein. The observed insolubility was consistant with earlier biochemical results obtained by dissociation of the enzyme (3), and showed that the previous results had not been a consequence of the denaturing conditions used to dissociate the subunits.

System expressing *Z. mays* small subunit.

- A fusion protein consisting of seven amino acids of B-galactosidase fused to the amino terminus of the mature *Z. mays* small subunit was shown to accumulate to high levels in cultures bearing appropriate expression plasmids.
- The small subunit fusion protein was almost completely sequestered into a rapidly sedimenting fraction of the bacterial extracts.

System expressing *Z. mays* and *Anabaena* PCC 7120 large and small subunits in homologous and heterologous combinations.

- Co-expression of both *Z. mays* RuBisCO subunits in *E. coli* did not result in holoenzyme assembly or catalytic

activity.

- Co-expression of Z. mays large subunits with the precursor form of the Z. mays small subunit, also did not result in either holoenzyme assembly or catalytic activity.
- Co-expression of both Anabaena subunits resulted in assembly of normal RuBisCO holoenzyme, even if the genes for the two subunits were on separate plasmids.
- Co-expression of Anabaena large subunits with Z. mays small subunits resulted in an active enzyme with sedimentation properties similar to wild-type holoenzyme.
- Co-expression of Z. mays large subunits with Anabaena small subunits did not result in holoenzyme assembly or activity.
- The results of these combinations are consistent with there being a factor necessary for assembly, with specificity for Anabaena large subunits, and without specificity for one or the other small subunits. In light of subsequent results from other laboratories on the role of chaperonin in RuBisCO assembly (4), and the demonstration that the endogenous E. coli chaperonin participates in assembly of cyanobacterial RuBisCO in E. coli (5), it is likely that the E. coli chaperonin (GroES and GroEL gene products) confers this specificity.

Mutational analysis of the Anabaena small subunit

- Residues in all three evolutionarily-conserved regions of the small subunit are important for association of the small subunit with the large subunit.
- Both hydrophobic and electrostatic interactions are

important for association of small subunits with large subunits.

- Many mutations in rbcS (the gene for the small subunit) resulted in reduced levels of assembled holoenzyme. The residues substituted in these mutants were not localized to a single cluster, either in the primary sequence or in the three-dimensional structure of the enzyme. At least in some cases, the molecular interactions which inspection of the three-dimensional structure suggested were interrupted by the substitutions were of the same nature as those disrupted in mutations in which assembly was completely blocked.
- No mutations were recovered with a significant reduction in activity in bacterial extracts without a parallel reduction in the amount of assembled holoenzyme. With more than a fifth of the residues of the small subunit represented in the sample (including most of the conserved residues), this suggests that the small subunit residues essential for full catalytic activity are the same as, or substantially overlap, the residues required for association of the small subunits with the large subunit core.
- No cluster of residues identifiable by substitution and suggestive of a site on the small subunit for interaction with chaperonin could be discerned from our collection of mutants.

ADVANTAGES OF SYSTEMS EXPRESSING RuBisCO IN *E. coli*

Most of the results described above could not have been achieved with RuBisCO isolated from plants, and could only have been derived from a expression of RuBisCO subunits in an organism amenable to the tools of molecular genetics. In the case of the systems for expression of individual subunits described in chapter two, the systems developed provided a source for each of the subunits without possibility of contamination by the other and without the need for denaturing conditions. In the case of the combinations of subunits described in chapter three, the two that resulted in activity (cyanobacterial large subunits and either cyanobacterial or higher plant small subunits) could be assembled *in vitro*, but the other two combinations, although not able to be assembled in *E. coli* either, could not even be attempted *in vitro* because higher plant large subunit cores proved impossible to isolate (6). Most importantly, many of the mutants reported in chapters four and five could not have been studied in plants because of the background of wild-type small subunit, as discussed in chapter one. They would also have been impossible to study in, for example, cyanobacteria because there are no established procedures for the growth of the readily transformable cyanobacteria without RuBisCO activity. (This may change in the near future, and should be entertained as a possible alternative to *E. coli* expression systems as it would add the possibility of genetic selections on the basis of RuBisCO activity.) At least at the time of these experiments, the ease with which mutant genes could be introduced, and the altered proteins isolated, was unique to systems based on expression in *E. coli* and was essential for analysis of any

sizable collection of directed mutants.

The RuBisCO expression systems described in this thesis are not unlike others developed concurrently in other laboratories (reviewed in 7). The experimental uses to which the other systems have been put further demonstrate the utility of such systems in general. An important example is the demonstration by Andrews (8) of RuBisCO activity in cyanobacterial large subunits in the absence of small subunits. Because this activity is very low compared to holoenzyme activity, and because small subunits are difficult to detect with sufficient sensitivity, it was not possible, previously, to exclude the possibility that the levels of activity detected in small-subunit-depleted preparations of large subunits were the result contamination by traces of un-dissociated small subunits. The use of an E. coli expression system provided a source of large subunits in the complete absence of small subunits, and made the experiment conclusive. Similarly, the demonstration by Goloubinoff et al. (5) of the requirement for chaperonin for holoenzyme assembly was only feasible using RuBisCO expressed in E. coli. It depended on being able to regulate RuBisCO biosynthesis experimentally, and made use of characterized chaperonin mutants available in E. coli. Neither the easily manipulable RuBisCO genes, nor the necessary mutants were available in any photosynthetic organism.

REMAINING QUESTIONS AND FUTURE APPROACHES

The question "how does the small subunit effect catalytic competence in the large subunits?" remains unanswered. The demonstration by Andrews (8) that large subunits have all the

residues mechanistically required for catalysis, and the results from X-ray crystallography (9,10,11,12) showing that the small subunits are at the poles of the enzyme, distant fron the active sites, merely show that its mode of action must be indirect. The results of my experiments (and similar results from the experiments of others) have increased our understanding of the function of individual residues, but did not resolve the question of how the small subunit has such a substantial effect on catalysis.

Future experiments will have the pronounced advantage of having been planned using insight gleaned from the three-dimensional structures of form I and form II RuBisCOs. From detailed comparisons of the arrangement of homologous residues in the form II enzyme and in the large subunits of the form I enzyme it may be possible to discover a structural difference, such as the orientation or position of a beta-strand or alpha-helix, which could be attributed to the presence and absence of the small subunit. For example, if, in the spinach or tobacco structures an -helix were part of the -barrel which forms the active site at one end, and near the binding site for the small subunit at the other end, but were shifted in position at the small subunit end relative to the position of the same strand in the R. rubrum structure, it would be reasonable to propose that interaction of these residues with small subunit residues might be responsible for the shift in position and that this shift could be transmitted to the active site. Such a proposal might be experimentally tested in a system similar to the E. coli expression system used in chapters four and five of this thesis by substituting small subunit residues, selectively, to perturb but not completely

disrupt, the critical interactions.

Alternatively, rbcL could be modified and expressed in the absence of small subunits. Although restoration to full activity by mutation would probably be impossible, several concerted changes, designed to mimic the presence of small subunits in the crucial interaction area, might increase the residual activity of small-subunit-free large subunit octamers. Such a mutant could provide strong support for the hypotheses upderlying its design, and might provide an understanding of the inherent limitations of the dimeric architecture of the form II enzyme. In any case, a better understanding of how the small subunit functions would suggest reasons for why the more complex form I enzyme is so widely distributed.

A second major unfinished research question involves the role of chaperonin in RuBisCO assembly. Expression of RuBisCO subunits in E. coli is well suited to experiments in this area as was demonstrated by Goloubinoff et al. (5). Now that the higher plant chaperonin genes are being/ have been cloned (13; R. J. Ellis, personal communication), the obvious experiment is to determine if their expression in E. coli, concurrently with both higher plant RuBisCO subunits, results in assembly of higher plant holoenzyme in the bacteria. Current models suggest that it would.

One aspect of this question is the determination of which stage(s) of RuBisCO assembly require chaperonin. The antibody inhibition data of Cannon et al. (14) suggest that chaperonin is required, at least, for oligomerization of large subunits. In vitro reassembly of isolated small subunits with cyanobacterial large

subunit cores (6), without any requirement for chaperonin, suggests that, for cyanobacterial RuBisCO assembly, the chaperonin requirement is limited to assembly of large subunit cores, but the equivalent experiment using higher plant large subunit cores cannot be performed due to problems with insolubility (6). Indeed,there is evidence to suggest that in higher plant chloroplasts, there is an association between chaperonin and the small subunits (4). A system for expression of higher plant RuBisCO in E. coli would allow one to look for mutations in the small subunit gene that blocked assembly through interference with the association of small subunit with chaperonin, demonstrating the essential nature of the small subunit-chaperonin interactions. Further, using temperature sensitive alleles of chaperonin and temperature regulated promotors for expression of individual subunits, or by construction of hybrid enzymes (plant large subunits and cyanobacterial small subunits) not currently obtainable, and exchange of small subunits it may be possible to further clarify the stages of RuBisCO assembly in which chaperonin is involved.

Even if comparison of form I and form II crystal structures does not reveal differences suggestive of a possible mechanism for small subunit action, and even if the expression of higher plant chaperonin in E. coli concurrently with higher plant RuBisCO subunits does not result in assembled holoenzyme, much further research on RuBisCO is certainly justified. Gene expression studies on RuBisCO small subunit genes (rbcS) provided the first documented plant examples of enhancers and biochemically-identified trans-acting factors. Studies of transport of mutant RuBisCO small subunits and derivatives into

chloroplasts have provided most of our current knowledge of the requirements for transport and processing of nucleus-encoded chloroplast proteins, and putative identification of a chloroplast membrane import receptor (15). Studies of RuBisCO biosynthesis contributed significantly to the discovery and understanding of chaperonins and the recognition of their widespread importance. Analysis of a mutant unable to activate RuBisCO *in vivo* led to the surprising discovery of RuBisCO activase. There is no reason to believe that there are not more discoveries yet to be made.

REFERENCES

1 Langridge, P. (1981) FEBS Lett. 123, 85-89.

2 Mulligan, R. M., Houtz, R. L. & Tolbert, N. E. (1988) Proc. Natl. Acad. Sci. USA 85, 1513-1517.

3 Houtz, R. L., Stults, J. T., Mulligan, R. M. & Tolbert, N. E. (1989) Proc. Natl. Acad. Sci. USA, in press.

4 Ellis, R. J. & Van der Vies, S. M. (1988) Photosynthesis Res. 16, 101-115.

5 Goloubinoff, P., Gatenby, A. A. & Lorimer, G. H. (1989) Nature 337, 44-47.

6 Andrews, T. J. & Lorimer, G. H. (1985) J. Biol. Chem. 260, 4632-4636.

7 McFadden, B. A. & Small, C. L. (1988) Photosynthesis Res. 18, 245-260.

8 Andrews, T. J. (1988) J. Biol. Chem. 263, 12213-12219.

9 Chapman, M. S., Suh, S. W., Cascio, D., Smith, W. W. & Eisenberg, D. (1987) Nature 329, 354-356.

10 Chapman, M. S., Suh, S. W., Curmi, P. M. G., Cascio, D., Smith, W. W. & Eisenberg, D. S. (1988) Science 241, 71-74.

11 Andersson, I., Knight, S., Schneider, G., Lindqvist, Y., Lundqvist, T., Branden, C. I. & Lorimer, G. H. (1989) Nature 337, 229-234.

12 Knight, S., Andersson, I. & Branden, C. I. (1989) Science 244, 702-705.

13 Hemmingsen, S. M., Woolford, C., Van der Vies, S. M., Tilly, K., Dennis, D. T., Georgopoulos, C. P., Hendrix, R. W. & Ellis, R. J. (1988) Nature 333, 330-334.

14 Cannon, S., Wang, P. & Roy, H. (1986) J. Cell Biol. 103, 1327-1335.

15 Pain, D., Kanwar, Y. S. & Blobel, G. (1988) Nature 331, 232-237.