Downloaded from rstb.royalsocietypublishing.org

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Future Prospects for Genetic Manipulation of Rubisco [and Discussion]

C. R. Somerville and G. Voordouw

Phil. Trans. R. Soc. Lond. B 1986 **313**, 459-469
doi: 10.1098/rstb.1986.0052

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

To subscribe to *Phil. Trans. R. Soc. Lond. B* go to: **http://rstb.royalsocietypublishing.org/subscriptions**

This journal is © 1986 The Royal Society

Downloaded from rstb.royalsocietypublishing.org

Phil. Trans. R. Soc. Lond. B **313**, 459–469 (1986)
Printed in Great Britain

Future prospects for genetic manipulation of Rubisco

By C. R. Somerville

MSU-DOE Plant Research Laboratory and Department of Botany and Plant Pathology, Michigan State University, East Lansing, Michigan 48824, U.S.A.

Recent advances in the development of techniques for the manipulation of gene structure *in vitro* and genetic transformation of plants have brought the goal of directed genetic modification of RuBP carboxylase–oxygenase (Rubisco) within grasp. Genes from both prokaryotic and eukaryotic species have been cloned, sequenced and expressed in *Escherichia coli*, and in several instances this has resulted in the production of large quantities of fully functional enzyme. Several specifically-modified enzymes have been produced by site-directed mutagenesis of a cloned gene and the effects of the mutations evaluated following expression of the modified genes in *E. coli*. Thus, there are no major technical barriers to the creation and analysis of modified enzymes. A number of new opportunities now exist to explore the structural basis of naturally occurring differences in kinetic constants of the enzymes from diverse taxonomic sources. The recent report of chloroplast transformation mediated by the Ti plasmid has also raised the possibility that, if useful natural variation can be identified, genes for both the large and small subunits of the enzyme may eventually be transferred between species. However, the opportunities for rational application of mutagenesis *in vitro* in the creation of useful or informative variants of the enzyme is currently limited by lack of information about tertiary structure and the role of specific residues in catalysis.

Introduction

There are, in principle, two distinct ways to approach the genetic manipulation of RuBP carboxylase–oxygenase (Rubisco). The simplest of these involves an attempt to identify natural variation by surveying the properties of the enzyme from diverse taxonomic sources with a view to transferring genetic information between species by genetic transformation. This approach is predicated on the recognition that homologous enzymes from organisms adapted to different niches may diverge substantially with respect to kinetic constants and structural properties. Thus, it is possible that some organisms may have evolved enzymes that are better suited to our needs than those already present in a particular crop species. The other approach seeks to create useful or informative variation by the appropriate application of mutagenesis *in vivo* or *in vitro*. Both of these approaches have been substantially invigorated by recent improvements in enzyme assay procedures, which permit a convincing determination of the various enzyme constants, and the rapid proliferation of information on primary amino acid sequences, deduced from DNA sequence analysis of cloned genes. Also, because of recent advances in techniques for transferring genes between species, the possibility now exists that if useful variation were identified or created it could be directly exploited by genetic engineering of crop species. Therefore, it seems appropriate to enquire at this juncture as to whether useful or informative genetic manipulation of Rubisco is imminent.

The Royal Society is collaborating with JSTOR to digitize, preserve, and extend access to
Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences.
www.jstor.org

Downloaded from rstb.royalsocietypublishing.org

Rubisco expression vectors

The essential prerequisite for protein engineering is a plasmid that permits production of the protein of interest in some easily manipulated host. Several such systems have recently been established for Rubisco genes from prokaryotic species. The most extensively investigated is the gene for Rubisco from *Rhodospirillum rubrum*, which has been cloned and sequenced and placed behind the *lac* promoter from *E. coli* in such a way that it is expressed in *E. coli* and amounts to as much as 15% of the total protein (Somerville & Somerville 1984). The protein produced in *E. coli* is a fusion polypeptide, in which the 20 amino-terminal amino acids are not naturally found in the enzyme (Nargang *et al.* 1984). However, this appears to have no significant effect on the enzyme's kinetic properties or stability during purification (Gutteridge *et al.* 1984). Conditions for growing very large cultures of the plasmid have been established, so that it is possible to produce 100 g of enzyme from a culture of about 300 l (Pierce & Gutteridge 1985). Recently the form-II *rbcL* gene from *Rhodopseudomonas sphaeroides* has also been cloned and expressed in *E. coli*, resulting in the production of an active hexameric (L_6) form of the enzyme (Quivey & Tabita 1984). This should greatly facilitate the analysis of this unusual form of the enzyme without the complications normally associated with separating it from the form I enzyme.

The discovery that the Rubisco genes of cyanobacteria are cotranscribed stimulated the construction of several expression vectors for these genes. Thus, the *rbcL* and *rbcS* genes from *Anacystis nidulans*, *Anabaena variabilis* and *Synechococcus* PCC6301 have been coexpressed in *E. coli* and shown to form hexadecamers of apparently normal activity (Gatenby *et al.* 1985; Gurevitz *et al.* 1985; Christeller *et al.* 1985; Tabita & Small 1985). The constructions employing the *A. nidulans* gene appear to exhibit the highest level of expression in *E. coli* and are, therefore, likely to be the system of choice for future studies of the hexadecameric enzyme. The availability of these plasmids implies that there should be no technical barrier to the subsequent analysis of the hexadecameric enzymes from cyanobacteria by mutagenesis techniques *in vitro*. This should permit the eventual identification of the structural features required for subunit interaction and may permit delineation of the role of the small subunit. For instance, if the only role of the small subunit is to catalyze hexadecamer formation, it may be predicted that the *only* small subunit mutations that inhibit Rubisco activity will be those that also prevent assembly of the hexadecamer. The analysis of this particular problem is timely since it does not depend upon knowledge of the tertiary structure.

It has not yet been possible to obtain expression of a functional hexadecameric enzyme from higher plants in *E. coli*. When expressed in the absence of small subunit, the *rbcL* genes from maize or *Anabaena* produce an insoluble and nonfunctional polypeptide (Gatenby *et al.* 1985; Gurevitz *et al.* 1985; Somerville *et al.* 1985). Thus, one function of the small subunit appears to be the solubilization of the large subunit polypeptide. In a situation where the expression of the large subunit polypeptide greatly exceeds that of the small subunit, all of the small subunit polypeptide was found in hexadecamers (Gurevitz *et al.* 1985). Thus, it was argued that the hexadecameric form of the enzyme arises by the formation of heterodimers which then assemble into hexadecamers. Whatever the mechanism of assembly, it seems apparent that no factors other than the large and small subunits are required to obtain assembly of the hexadecameric form of the cyanobacterial enzyme.

Because the two subunits of Rubisco are translated in separate compartments in eukaryotes,

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES
PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Downloaded from rstb.royalsocietypublishing.org

it is possible that there is something fundamentally different about the mechanism of assembly as compared with the pathway in prokaryotes. It is conceivable that a companion protein, which has been observed to be associated with newly synthesized large subunit polypeptide (Barraclough & Ellis 1980; Bloom *et al.* 1983) is required in eukaryotes to keep the large subunit soluble prior to assembly with the small subunit. This could have important implications for future attempts to obtain expression of eukaryotic Rubisco in *E. coli*. However, the inability to obtain expression of a functional enzyme cannot be used as evidence for the involvement of another factor. Since the small subunit is normally translated as a precursor, it is possible that expression of a restructured small subunit that lacks the transit peptide in *E. coli* will not result in a protein of correct tertiary structure because the pathway of folding may be quite different.

An insight into the future prospects for obtaining expression of a functional higher plant Rubisco in *E. coli* may be inferred from reconstruction experiments *in vitro* in which the small subunits were first stripped from a cyanobacterial enzyme and a functional hexadecameric form was reconstituted with small subunits from cyanobacterial or spinach Rubisco (Andrews *et al.* 1984; Andrews & Lorimer 1985). Although this pathway of assembly may not be directly comparable to the normal pathway of assembly, it does indicate that the binding sites for large subunit–small subunit interaction are conserved and that the properties of the small subunit which are essential to normal function of the enzyme are conserved. Thus it may be possible to solubilize the large subunit from a eukaryote with the small subunit from a prokaryote.

Natural variation in Rubisco

Michaelis constants of Rubisco from a wide variety of higher plant, algal and bacterial species have recently been convincingly measured by a number of groups (Bird *et al.* 1982; Jordan & Ogren 1981, 1983; Seemann *et al.* 1984; Yeoh *et al.* 1981). These studies have uniformly revealed that variability exists with respect to all kinetic constants, including the ratio of carboxylase to oxygenase activity. C_3 species generally have the lowest $K_m[CO_2]$ values (table 1), suggesting that these species have experienced the greatest pressure to adapt the enzyme to the absolute concentration of atmospheric CO_2. Unfortunately it has been difficult to establish convincing quantitative estimates of V_{max}. However, the values in table 1 suggest that there may be a direct correlation between $K_m(CO_2)$ and V_{max}. The major discrepancy in this respect is a low V_{max} value for the *Anabaena* enzyme. Since previous estimates of the V_{max} for *Synechococcus* were also comparably low (Andrews & Abel 1981), it seems likely that

Table 1. Kinetic constants for Rubisco from several species

(The values were compiled from Andrews & Abel (1981), Andrews & Lorimer (1985), Badger (1980), Gutteridge *et al.* (1984), Jordan & Ogren (1981), Seemann *et al.* (1984) and Servaites (1985).)

species	$K_m[CO_2]$ / μM	$K_m[O_2]$ / μM	specific activity: μmol CO_2 min^{-1} mg $protein^{-1}$	specificity factor
Rhodospirillum rubrum	89	406	6.4	15
Anabaena variabilis	293	1008	2.4	31
Synechococcus sp.	212	997	13.4	40
Chlamydomonas reinhardtii	29	480	6.7	61
Zea mays	34	810	5.6	78
Spinacia oleracea	15	480	3.7	80
Nicotiana tabacum	11	650	1.9	77

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Downloaded from rstb.royalsocietypublishing.org

the value for *Anabaena* may simply be underestimated because of enzyme inactivation during purification.

The apparent relation between enzyme rate and the binding constant for CO_2 is consistent with the general proposal (Fersht 1977) that K_m and k_{cat} are inversely related for mechanistic reasons related to the affinity of the enzyme for the transition state intermediate. This relationship, if substantiated by subsequent analysis, could have important implications for future attempts to manipulate the enzyme.

Neglecting for the moment the technical difficulties associated with transferring genes from one species to another, it may be concluded at present that, even if this approach were technically feasible, there does not appear to be a variant form of the enzyme which would confer a significant benefit if transferred to a C_3 species, such as spinach. This is due to the fact that the higher V_{max} values of C_4 plants' or algal enzymes are associated with increases in K_m (CO_2), which would offset the gain. This may be seen by assuming a cellular CO_2 concentration (p_{CO_2}) of about 5 μM, setting $p_{O_2} = 240$ μM and substituting the values from table 1 into the following equation for the rate of carboxylation:

$$V_c = V_{max}(CO_2)/\{1+[K_m(CO_2)/p_{CO_2}]\}\{1+[p_{O_2}/K_m(O_2)]\}.$$

There may be exceptions to this situation. For instance, if the V_{max} estimate for tobacco is accurate (table 1), it might be possible to improve photosynthesis in tobacco by, for instance, introducing the genes for the spinach enzyme. It is, therefore, very important that careful and convincing estimates of the kinetic constants should be carried out on the enzymes from a diverse collection of C_3 species.

In addition to the possible opportunities with respect to plant breeding, natural variation may be useful in attempting to understand the structural basis of catalysis. The sequences for many Rubisco genes are now known (figure 1). Thus, in principle, one could attempt to correlate changes in amino acid residues with differences in enzyme properties (table 1). Unfortunately there is no way to do this quantitatively, and there are too many amino acid substitutions to permit a correlative approach solely on the basis of sequence comparisons. Indeed, many of the observed differences are likely to be neutral. The best available estimate for the proportion of differences that are neutral is probably derived from the study of the *lac* repressor, in which 58% of 302 amino acid replacements did not result in a detectable change in phenotype (Miller 1978). To exploit this growing base of structural information, a substantial reduction in complexity is required. This might, in principle, be accomplished by splicing fragments of genes together to produce hybrid enzymes. Because of remarkable conservation of primary sequence, some of the restriction sites remain in common between genes as diverse as the *Zea mays* and *Anabaena variabilis rbcL* genes. Thus, by constructing hybrid genes *in vitro* and placing them in a suitable expression vector, it might be possible to investigate the functional significance of these changes.

This approach has previously been explored by genetic techniques *in vivo* to create hybrid forms of the tryptophan synthetase enzymes from *E. coli* and *Salmonella typhimurium* (Schneider *et al.* 1981). These enzymes differ at 40 of their 286 amino acid residues and have 199 sequence differences in a coding region of 804 nucleotides. Five hybrid enzymes which had from 6 to 34 substitutions were examined in some detail. All of these hybrid enzymes had only slightly different catalytic characteristics but were substantially more sensitive to thermal denaturation than the progenitors. The major conclusion from this study was that, although it was apparent

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES
PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Downloaded from rstb.royalsocietypublishing.org


```
    0       10        20        30        40        50        60         70        80       90
    MS-PQTETKASVGFKAGVKDYKLTYYTPDYTTKDTDILAAFRVTPQPGVPPEEAGAA-VAAESSTGTWTTVWTDGLTSLDRYKGRCYHI
An:  -  K QSA -- Y                 P     L    FS      AD       I               L  DM     K
Av:    YA K QTKS Y     Q  R         P                  F   A                   L  D         D
Ss:  -- --K QSAA Y                 P     L    P S     AD       I               L  DM     K
Cr:   V       GA          R       VVR          M   L      C                                 D
So:              E               E E L          S                                 N
Zm:                              E E               L                 A
Nt:                     E        E Q                                                        R
Rr: -------------MDQSSRYVN ALKEEDLIAGGEHV  C Y-IMKPKAGYGYV T  HF        NVE C --- DDFTRGVDALVY
```

```
    91       100       110       120       130       140       150       160       170       180
    EPVPGEDNQYIAY-VAYPLDLFEEGSVT-NMF-TSIVGNVFGFKALRALRLEDLRIPVAYVKTFQGPPHGIQVERDKLNK---YGRPLLGC
An:    Q  E S F F I               IL             I S     I F   L                L           M
Av:          F    I           I   VL                     I F     I
Ss:    Q  E S F F I               IL             I S     I F   L                L T         M
Cr:                   I                                       P      V                     G
So:    A  E    C
Zm:      DPD   C                                              P  S       R M        N
Nt:  R V  KD                                                  P
Rr:  VDEAR LTK   P  LFDRNITD KAMIAS L LTM  NQ MGDVEYAKMH FYV  E RAL  D SVN SALWKV GRPEVD GLVV  T
```

```
    181      190       200       210       220       230       240       250       260       270
    TIKPKLGLSAKNYGRAVYECLRGGLDFTKDDENVNSQPFMRWRDRFLFVAEAIYKAQAETGEIKGHYLNATAGTCEEMMKRAEFAKELGV
An:                                   I     Q          D  H S             V  P               M
Av:                                   I  A  Q          D  T                V  P     L    Y   KQ
Ss:                                   I     Q          D  H S             V  P               M
Cr:                                                                V                   VC
So:                                                  C   L                       D     V  R
Zm:                  C                             V C      S                   D  I G    RQ
Nt:                                                  C   L                         I   V  R
Rr: I       RP PFAE CHAFWL   -  I N PQGN   APL  TIALV D MRR   D    A LFSA I  DDPFEIIA GEYVL T-F
```

```
    271      280       290       300       310       320       330       340       350       360
    PIIM-H-DYLTGGFTANTTLAHYCRDNGLL--LHIHRAMH-AVIDRQ-KNHGIHFRVLAKALRLSGGDHIHSGTVV-GKLEGEREITLGF
An:         F   A        KW      V                   R           C        L            DKAS
Av:             A        RW                                                 T            G  M
Ss:         F   A        KW      V                   R           C        L            DKAS
Cr:                    S  I                          R              M     L               V
So:   V                  S     I                         M                               D
Zm:   V   D              S                               M  G       M
Nt:   V                S                                            M                    D
Rr: GENASHVAL VD YV GAAAITTA RRFPDNF   Y  G G  TSP S R- YTAF HC MA   Q ASG  T  MGF  M   SS-----
```

```
    361      370       380       390       400       410       420       430       440       450
    VDLLRDDYVEKDRSRGIYFTQDWASMPGVLPVASGGIHVWHMPALVEIFGDDSV-LQFGGGTLGHPWGNAPGATANRVALEACVQARNEG
An:    M E    A     VF                                                     L
Av:      EN   Q K            L   MA                                        R
Ss:    M E HI R     VF
Cr:    M                   C      M                     AC                    A         T
So:                      S V T                    T                           V
Zm:                        V      I               T  L       E             H  A
Nt:           Q            V L     E              T                           V          K
Rr: - RAIAYMLTQ EAQ PFYR S GG KACT II   MNALR  GFF NL NAN I TA    AF  ID PVA RSL Q WQ WRDGVPVL
```

```
    451      460       470       480       490       500
    RDLAREGNDII-REACKWS-PELAAACEL-W-KEIKFE-FEAMDTL
An:    Y   G  L    G          LD              T  K
Av:  N       V     A         V                    V
Ss:    Y   G  L    G          LD              T  K
Cr:        G V    S              V          DTI  K
So:         T      T             V           P    V
Zm:       VQ-   KA       A       I    D  DG  K    I
Nt:      Q  EV  I                V      V N  A V -VDK
Rr: - Y   HKELA- - FE- F GD DQIYPG  RALGV   DTRSALPA
```

FIGURE 1. A comparison of the amino acid sequences of the large subunit polypeptides of Rubisco from eight species. The uppermost sequence is a consensus sequence, representing the most commonly used residue at each position. A dash indicates that no amino acid occurs in the corresponding sequence. The abbreviations for the sequences and the corresponding citations are: An, *Anacystis nidulans* (Shinozaki *et al.* 1983); Av, *Anabaena variabilis* (Curtis & Haselkorn 1983); Ss, *Synechococcus* sp. (Reichelt & Delaney 1983); Cr, *Chlamydomonas reinhardtii* (Dron *et al.* 1982); So, *Spinacia oleracea* (Zurawski *et al.* 1981); Zm, *Zea mays* (McIntosh *et al.* 1980); Nt, *Nicotiana tabacum* (Amiri *et al.* 1984); Rr, *Rhodospirillum rubrum* (Nargang *et al.* 1984).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Downloaded from rstb.royalsocietypublishing.org

that certain changes had been balanced with respect to overall protein conformation, there did not appear to be significant interdependence of amino acid substitutions with respect to residues required for catalysis. It would clearly be very interesting to explore this approach with hybrids between various forms of Rubisco. Such a study might reveal domains responsible for specific aspects of Rubisco function.

Site-directed mutagenesis

Oligonucleotide-directed mutagenesis promises to be a very powerful technique for the analysis of structure–function relationships in proteins. Although there are strong practical motivations to study the hexadecameric form of the enzyme, the *R. rubrum* enzyme has attracted interest because it is small enough to permit NMR spectroscopic analysis, and the absence of a second subunit reduces concern about mutations that disrupt activity by altering quaternary structure. Several mutations of the *R. rubrum* enzyme have been accomplished. In one instance the aspartate residue at position 188 (206 in figure 1) of the *R. rubrum* enzyme was changed to a glutamate in order to investigate the role of this residue in supporting CO_2-dependent activation (Gutteridge *et al.* 1984). Although subtle changes in kinetic constants were observed, they cannot currently be interpreted mechanistically. In another case, the Lys 191 residue (209 in figure 1), implicated as the residue that undergoes carbamylation during CO_2-dependent activation, was changed to a glutamate to mimic the effect of carbamate formation (Estelle *et al.* 1985). This change eliminated Mg^{2+}- and CO_2-dependent stabilization of carboxyarabinitol bisphosphate (CABP) binding. Since the γ-carboxyl of glutamate is functionally similar to the carboxyl of the carbamate, it seems likely that the precise position of the carboxyl is critical to formation of the quaternary complex (i.e. enzyme–CABP–CO_2–Mg^{2+}). In another instance, Met 330 of the *R. rubrum* Rubisco (residue 350 in figure 1), which has been identified as near the active site by affinity labelling, was changed to a leucine (Terzaghi *et al.* 1986). Since leucine is the consensus residue at this position in all other forms of the enzyme (figure 1), it might have been expected that the effect of the mutation would be neutral. However, this substitution caused numerous changes in the kinetic constants, including a 15-fold increase in the $K_m(CO_2)$ and a 36-fold increase in $K_i(O_2)$, but did not change the partitioning coefficient (Terzaghi *et al.* 1986). In this respect, the effects of this mutation provide a good example of the interplay between different substitutions which must have occurred during enzyme evolution, and provide additional evidence for the common structural basis for the two activities associated with the enzyme.

These studies demonstrate the feasibility of the approach. However, in the absence of a tertiary structure, they are not easily integrated into a rational model of enzyme structure and function. A potentially bothersome aspect of this approach is uncertainty as to whether any effects resulting from a particular mutation reflect the direct participation of the modified residue in catalysis, or whether the effects result from changes in the folding of the polypeptide. There are no simple answers to this problem, although techniques such as circular dichroism may be useful in determining whether mutant proteins have gross changes in tertiary or quaternary structure. It is hoped that the tertiary structure will soon be available for the *R. rubrum* enzyme so that we may undertake the analysis of Rubisco at the same level as has been so elegantly applied to tyrosyl tRNA synthetase (Wells & Fersht 1985).

Although it is difficult to phrase incisive questions at present, it seems worthwhile to assemble

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Downloaded from rstb.royalsocietypublishing.org

a collection of variant enzymes for subsequent interpretation when a tertiary structure becomes available, or when advances in protein-modelling techniques make it possible to infer structure from primary sequence. Indeed, a project designed to use site-directed mutagenesis to identify spectral features of the ^{19}F nuclear magnetic resonance spectrum of 3-fluorotyrosine-substituted Rubisco would be directly useful and might greatly facilitate the application of computer modelling techniques to Rubisco by providing spatial constraints. The basic techniques for such an approach are already available (Lu *et al.* 1976).

Genetic systems *in vivo*

The usefulness of genetic methods (at their best) for generating informative variants is most clearly exemplified by studies of the *E. coli lac*-repressor protein, encoded by *lacI*. The beauty of the system is that mutant phenotypes are readily evident by colony colour on indicator plates containing chromogenic substrates. Analysis of several thousand *lacI* mutants has revealed that the repressor is divided into two distinct domains (Miller 1978). One domain, of approximately 60 amino residues, is involved in DNA and operator binding but not in inducer binding or aggregation. The remainder of the protein carries out these functions. Mutationally altered repressors, defective in each of these functions, as well as those displaying increased affinity for the operator, have been characterized, and have provided the most detailed view available for any protein of the possibilities for structural variation.

Unfortunately, it is not apparent how to adapt the elegant technology used in the study of *lacI* to the study of Rubisco. There are no chromogenic substrates, and bacterial cells are impermeable to sugar phosphates. It seems likely that the best that can be done is simply to isolate mutants of a transformable photosynthetic heterotroph by screening for colonies that do not grow phototrophically. Putative Rubisco mutants could be verified by transforming with the wild-type gene and the mutations subsequently mapped to specific regions of the gene by transformation with fragments of the gene and genomic sequencing. This approach is not particularly elegant but it could certainly provide a great deal of useful information.

Genetic engineering of higher plants

It is now routinely possible to introduce genes into the nuclear genomes of many dicotyledons by Ti-plasmid-mediated transformation, and the continued development of techniques for direct gene transfer will undoubtedly extend the possibility to monocotyledons. Thus, if useful variants of the *rbcS* gene could be identified it would be possible to introduce these. One interesting possibility in this respect is that, by adding additional copies of the *rbcS* gene, it might be possible to increase the amount of Rubisco. The rationale for this is the observation that the chloroplast has a mechanism for turning over unassembled Rubisco small subunit polypeptides (Schmidt & Mishkind 1983). The implication is that this is a general mechanism and that any 'abnormal' protein, such as unassembled large subunit, will be rapidly degraded. Thus, since it is not yet known what regulates the amount of Rubisco, it is possible that there is excess synthesis of large subunit translation and that the amount of Rubisco is limited by the rate of small subunit synthesis. This is directly testable by simply increasing the copy number of the *rbcS* gene.

Because the catalytic activity is associated with the large subunit polypeptide (Andrews &

Downloaded from rstb.royalsocietypublishing.org

Abel 1981; Andrews *et al.* 1984), it will be necessary to transform chloroplasts in order to manipulate genetically the properties of Rubisco in higher plants. Until recently this seemed a virtually insurmountable hurdle. The report of fortuitous genetic transformation of chloroplasts (DeBlock *et al.* 1985) raises the possibility that genes can be introduced into the chloroplast. The transformants obtained so far were unstable and ultimately lost the introduced gene. However, this seems a comparatively minor problem and will almost certainly be overcome. Presumably an introduced *rbcL* gene would either replace the endogenous *rbcL* gene or would integrate somewhere in the chloroplast, creating a duplication. If the orientation of the insert were inverted relative to the existing copy, the situation would mimic the duplication that characterizes the chloroplast genome of most non-leguminous higher plants. In the few instances where detailed chloroplast genetic studies have been possible, it has been observed that a phenomenon of gene conversion involving the inverted repeats seems to occur. This is apparent in mutants of *Chlamydomonas reinhardtii* in which the same mutation is found in both copies of genes in the inverted repeat (Erickson *et al.* 1984; Myers *et al.* 1982) and has been inferred from the observation of symmetrical restriction patterns in the inverted repeats of *Oenothera* (Herrmann *et al.* 1980). Thus, it is possible that an introduced copy will cause 'correction' of the endogenous copy. The alternative possibility, of introducing the *rbcL* gene into the nucleus and having it transported into the chloroplast by the addition of a transit sequence, as was done for a bacterial gene for antibiotic resistance (Van den Broeck *et al.* 1985), does not seem likely in the case of Rubisco because of the insolubility of the large subunit. Indeed, it may well be this characteristic of Rubisco that is the reason that it is encoded in the chloroplast DNA.

Since, as noted earlier, we have not yet identified any convincingly 'superior' variants of Rubisco among C_3 plants, the transformation technology is not currently a limitation.

Prospects

The application of recombinant DNA techniques to the study of Rubisco has led to rapid progress on several fronts during the past several years. We are now poised to ask specific questions pertaining to the role of the small subunit and, more generally, to the structural basis of catalysis. Unfortunately, the reactions catalysed by the enzyme are poorly suited for genetics *in vivo*. Thus, it seems likely that most manipulations of the primary sequence will be directed. Although this can be accomplished with available technology, the approach requires a level of effort comparable to that required for conventional enzymology. Thus, it is important that this effort be invested in addressing questions that are based on a firmly grounded appreciation of the probable role of specific amino acid residues in catalysis. In this respect, future progress is likely to be largely dependent on the interpretation of X-ray crystallographic information and the integration of this structural information with information pertaining to the sequence of chemical reactions catalysed by Rubisco.

I am grateful to George Haughn, Bill Laing, John Pierce, John Andrews and George Lorimer for stimulating discussions and for communicating unpublished results.

Downloaded from rstb.royalsocietypublishing.org

References

Amiri, I., Salnikow, J. & Vater, J. 1984 Amino acid sequence of the large subunit of D-ribulose bisphosphate carboxylase/oxygenase from *Nicotiana tabacum*. *Biochim. biophys. Acta* **784**, 116–123.

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

Andrews, T. J., Greenwood, D. M. & Yellowlees, D. 1984 Catalytically active hybrids formed *in vitro* between large and small subunits of different procaryotic ribulose bisphosphate carboxylases. *Archs Biochem. Biophys.* **234**, 313–317.

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

Badger, M. R. 1980 Kinetic properties of ribulose 1,5-bisphosphate carboxylase/oxygenase from *Anabaena variabilis*. *Archs Biochem. Biophys.* **201**, 247–254.

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

Bird, I. F., Cornelius, M. J. & Keys, A. J. 1982 Affinity of RuBP carboxylases for carbon dioxide and inhibition of the enzymes by oxygen. *J. exp. Bot.* **33**, 1004–1013.

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

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

Curtis, S. E. & Haselkorn, R. 1983 Isolation and sequence of the gene for the large subunit of ribulose-1,5-bisphosphate carboxylase from the cyanobacterium *Anabaena* 7120. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1835–1839.

De Block, M. Schell, J. & Van Montagu, M. 1985 Chloroplast transformation by *Agrobacterium tumefaciens*. *EMBO J.* **4**, 1367–1372.

Dron, M., Rahire, M. & Rochaix, J. D. 1982 Sequence of the chloroplast DNA region of *Chlamydomonas reinhardtii* containing the gene of the large subunit of ribulose bisphosphate carboxylase and parts of its flanking gene. *J. molec. Biol.* **162**, 775–793.

Erickson, J., Rahire, M., Bennoun, P., Delepelaire, P., Diner, B. & Rochaix, J. D. 1984 Herbicide resistance in *Chlamydomonas reinhardtii* results from a mutation in the chloroplast gene for the 32-kilodalton protein of photosystem II. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3617–3621.

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

Fersht, A. 1977 *Enzyme structure and mechanism*. Reading: W. H. Freeman.

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

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

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

Herrmann, R. G., Seyer, P., Schedel, R., Gordon, K., Bisanz, C., Winter, P., Hilderbrandt, J. W., Wlaschek, M., Alt, J., Driesel, A. & Sears, B. B. 1980 The plastid chromosome of several dicotyledons. In *Biological chemistry of organelle formation* (ed. T. Bucher, W. Sebald & H. Weiss), pp. 97–112. Berlin: Springer-Verlag.

Jordan, D. B. & Ogren, W. L. 1981 Species variation in the specificity of ribulose bisphosphate carboxylase/oxygenase. *Nature, Lond* **219**, 513–515.

Jordan, D. B. & Ogren, W. L. 1983 Species variation in kinetic properties of ribulose 1,5-bisphosphate carboxylase/oxygenase. *Archs Biochem. Biophys.* **277**, 425–433.

Lu, P., Jarema, M., Mosser, K. & Daniel, W. 1976 Lac repressor: 3-fluorotyrosine substitution for nuclear magnetic resonance studies. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3471–3475.

McIntosh, L., Poulsen, C. & Bogorad, L. 1980 Chloroplast gene sequence for the large subunit of ribulose bisphosphate carboxylase of maize. *Nature, Lond.* **288**, 556–560.

Miller, J. H. 1978 The *lac*I gene: its role in lac operon control and its use as a genetic system. In *The operon* (ed. J. H. Miller & W. S. Reznikoff), pp. 31–88. Cold Spring Harbor Laboratory.

Myers, A. M., Grant, D. M., Robert, D. K., Harris, E. H., Boynton, J. E. & Gillham, N. W. 1982 Mutants of *Chlamydomonas reinhardtii* with physical alterations in their chloroplast DNA. *Plasmid* **7**, 133–151.

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

Pierce, J. & Gutteridge, S. 1985 Large-scale preparation of ribulosebisphosphate carboxylase from a recombinant system in *Escherichia coli* characterized by extreme plasmid instability. *Appl. Environ. Microbiol.* **49**, 1094–1100.

Quivey, R. G. & Tabita, F. R. 1984 Cloning and expression in *Escherichia coli* of the form II ribulose 1,5-bisphosphate carboxylase/oxygenase gene from *Rhodopseudomonas sphaeroides*. *Gene* **31**, 91–101.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Downloaded from rstb.royalsocietypublishing.org

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

Schmidt, G. W. & Mishkind, M. L. 1983 Rapid degradation of unassembled ribulose 1,5-bisphosphate carboxylase small subunits in chloroplasts. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2632–2636.

Schneider, W. P., Nichols, B. P. & Yanofsky, C. 1981 Procedure for production of hybrid genes and proteins and its use in assessing significance of amino acid differences in homologous tryptophan synthetase polypeptides. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2169–2173.

Seemann, J. R., Badger, M. R. & Berry, J. A. 1984 Variations in the specific activity of ribulose-1,5-bisphosphate carboxylase between species utilizing different photosynthetic pathways. *Pl. Physiol.* **74**, 791–794.

Servaites, J. C. 1985 Crystalline ribulose bisphosphate carboxylase/oxygenase of high integrity and catalytic activity from *Nicotiana tabacum*. *Archs Biochem. Biophys.* **238**, 154–160.

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

Somerville, C. R., McIntosh, L., Fitchen, J. & Gurevitz, M. 1985 The cloning and expression in *Escherichia coli* of RuBP carboxylase/oxygenase large subunit genes. *Meth. Enzymol.* **118**, 419–433.

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

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

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

Van den Broeck, G., Timko, P., Kausch, A. P., Cashmore, A. R., Van Montagu, M. & Herrera-Estrella, L. 1985 Targeting of a foreign protein to chloroplasts by fusion to the transit peptide from the small subunit of ribulose-1,5-bisphosphate carboxylase. *Nature, Lond.* **313**, 358–363.

Wells, T. N. C. & Fersht, A. R. 1985 Hydrogen bonding in enzymatic catalysis analysed by protein engineering. *Nature, Lond.* **316**, 656–657.

Yeoh, H. H., Badger, M. R. & Watson, L. 1981 Variations in kinetic properties of ribulose 1,5-bisphosphate carboxylase among plants. *Pl. Physiol.* **67**, 1151–1155.

Zurawski, G., Perrot, B., Bottomley, W. & Whitfield, P. R. 1981 The structure of the gene for the large subunit of ribulose 1,5-bisphosphate carboxylase from spinach chloroplast DNA. *Nucl. Acids Res.* **9**, 3251–3270.

Discussion

G. Voordouw *Department of Biochemistry, Agricultural University, Wageningen, The Netherlands*. There appears to be a discrepancy between the results obtained for the assembly of Rubisco from *Synechococcus*, presented by Dr D. Bradley, and Dr Somerville's results for *Anabaena* Rubisco, both enzymes being assembled in *E. coli*. Dr Bradley finds an L_8S_{2-3} enzyme and Dr Somerville, exclusively, an L_8S_8 enzyme plus an L-aggregate. In thinking about this in the days after the meeting, it occurred to me that this difference could well be due to a difference in solubility of the L_8 octamer formed. The octamer of *Synechococcus* is soluble and has little tendency to self-aggregate, whereas that of *Anabaena* has a stronger tendency towards self-aggregation. The L_8S_{2-3} enzyme of *Synechococcus* formed in *E. coli* is thus likely to be a mixture of L_8S_8 and (soluble) L_8 forms (whereas the aggregated octamers from *Anabaena* are removed by centrifugation) when a cell-free *E. coli* extract is prepared, leaving just the L_8S_8 enzyme.

C. R. Somerville. Andrews and collaborators have found no evidence for cooperative binding of small subunits to L_8 cores from *Synechococcus*. Furthermore, these L_8 cores are soluble. Thus, there seems no reason to believe that Dr Bradley's preparations are a mixture of L_8S_8 and L_8 forms. Since our interpretation of the observations with *Anabaena* Rubisco suggest an obligate L_1S_1 intermediate is required for assembly of the L_8S_8 form, I suggest that Dr Bradley's preparations arise by competition between L_8 cores and newly synthesized L_1 monomers for

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Downloaded from rstb.royalsocietypublishing.org

small subunit. Thus, it is possible that L_1S_1 forms rapidly assemble to L_8S_8 forms, which, in the case of *Synechococcus*, may degenerate to L_8S_{2-3}. The fact that this does not happen with the *Anabaena* enzyme may reflect tighter binding of the small subunit to the L_8 core from *Anabaena*. Indeed, the atypically loose binding of small subunits to *Synechococcus* is the reason this enzyme is used for reconstruction experiments.

BIOLOGICAL SCIENCES
THE ROYAL SOCIETY
PHILOSOPHICAL TRANSACTIONS OF

BIOLOGICAL SCIENCES
THE ROYAL SOCIETY
PHILOSOPHICAL TRANSACTIONS OF