Cloning and Characterization of the P Subunit of Glycine Decarboxylase from Pea (Pisum sativum)*

(Received for publication, May 6, 1991)

Simon R. Turner, Robert Ireland‡, and Stephen Rawsthorne§

From the Cambridge Laboratory, Agricultural and Food Research Council, Institute of Plant Science Research, John Innes Centre, Colney Lane, Norwich NR4 7UJ, United Kingdom and the §Department of Biology, Mount Allison University, Sackville, New Brunswick EBO 3C0, Canada

A pea leaf cDNA library constructed in λgt11 was screened with an antibody raised to the P subunit of glycine decarboxylase. One of the positive clones isolated was sequenced and shown to contain an open reading frame, which encoded the entire P subunit polypeptide. Aligning the deduced amino acid sequence with the amino acid sequence determined directly from the NH₂ terminus of the mature P subunit shows the presence of a putative 86 amino acid leader sequence, presumably required for import into the mitochondria, and gives a M, of the mature protein of 105,000. Comparison of this deduced amino acid sequence with the sequence of a pyridoxal phosphate-containing peptide isolated from the P subunit of chicken liver glycine decarboxylase shows remarkable conservation. The P subunit, however, shows little sequence homology with other published amino acid decarboxylases. Expression of the P subunit mRNA shows a pattern very similar to that of the corresponding polypeptide: it is strongly light induced and is expressed at a much higher level in leaves than in other tissues. Southern blot analysis suggests that the P subunit is encoded by a small multigene family.

In both animal and plant mitochondria glycine decarboxylase (GDC) catalyzes the oxidation of glycine to form CO₂, NH₃, NADH, and 5,10-methylene tetrahydropteroyl-L-glutamic acid (Motokawa and Kikuchi, 1974; Clandinin and Cossins, 1975). The enzyme is composed of four subunits: P (containing pyridoxal phosphate), H (a carrier protein containing lipoamide), T (a transferase responsible for producing 5,10-methylenetetrahydropteroyl-L-glutamic acid), and L (lipoamide dehydrogenase required to complete the cycle). The P subunit from chicken and bovine liver (Fujiwara et al., 1987), from chicken and bovine liver (Kim and Oliver, 1990; Macherel et al., 1991). In addition, the sequence of the T subunit from bovine liver has recently been published (Oka-ura-Ikeda et al., 1991). Studies on the sequence of the P subunit and its interaction with the other subunits of the enzyme have been hampered by the availability of only the amino acid sequence of a small pyridoxal phosphate-binding peptide from the protein purified from chicken liver (Fujiwara et al., 1987). However, Kume et al. (1991) recently reported the cloning of cDNAs encoding the P subunit from chicken and human, and here we report for the first time the cloning and characterization of a full-length cDNA clone for the P subunit of GDC from plants and the deduced amino acid sequence of the complete polypeptide.

EXPERIMENTAL PROCEDURES

Plant Material—Pisum sativum cultivar "Birle" was grown under natural day length in a greenhouse with a 20 °C day and 16 °C night temperature in seedtrays containing vermiculite after surface sterilization of the seeds in 10% (v/v) commercial bleach solution. Plants were harvested within 2 weeks except for those required for mRNA extraction from mature green leaves and embryos, which were grown in pots containing a mixture of John Innes No. 1 compost and grit.

---

* This work was funded by the Cambridge Laboratory. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper have been submitted to the GenBank™/EMBL Data Bank with accession number(s) X59773.

† To whom correspondence and requests for reprints should be addressed.

The abbreviations used are: GDC, glycine decarboxylase complex; SDS, sodium dodecyl sulfate; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid.

Vol. 267, No. 8, Issue of March 15, pp. 5353-5360, 1992
Printed in U.S.A.
Etiolated seedlings were grown for 6 days in the dark at 20 °C before exposure to continuous light at 20 °C for various periods of time (100 μmol quanta of photosynthetically active radiation m⁻² s⁻¹). Plants for dark treatment were grown in the glasshouse for 3 weeks prior to the dark period.

**Protein Purification and Antibody Production**—The P subunit was purified from an acetone precipitate (Walker and Oliver, 1986a) of pea leaf mitochondria prepared by the method of Day et al. (1985) by running a 50–60% saturated ammonium sulfate fraction of the resuspended matrix proteins on a 12% SDS-polyacrylamide gel (Laemmli, 1970) and electrophoretically separating the very prominent band of M₉₇,₀₀₀ which has previously been identified as the P subunit (Walker and Oliver, 1986a). H protein was gel-purified in the same way from the 75–100% ammonium sulfate fraction (Walker and Oliver, 1986a). Electrophoresed proteins were freeze-dried and 300 μg of P and 100 μg of H were injected subcutaneously into separate rabbits after taking samples of preimmune serum. The proteins were equally divided between three injections made at 10-day intervals, the first using Freund's complete adjuvant and the others incomplete adjuvant. Serum was collected at 10-day intervals after the third injection, and IgG fractions were purified using a protein A-Sepharose column (Ey et al., 1978).

**Protein Sequence Analysis**—The NH₂-terminal protein sequence was determined using an Applied Biosystems (Warrington, Cheshire, United Kingdom (U. K.) 470A gas-phase protein sequencer equipped with a 120A on-line phenylthiohydantoin analyzer. The P protein bands cut from Immobilon blots of low pH SDS-polyacrylamide gels (Morris et al., 1988) containing the 50–60% ammonium sulfate fraction (see above) and stained with Brilliant Blue R were placed directly onto the sequencer reaction cartridge and held in place with trifluoroacetic acid-etched glass fiber discs (Hunkapiller et al., 1983).

**Antibody Characterization and GDC Assays**—SDS-polyacrylamide gels were electrophoresed onto nitrocellulose and development of Western blots was essentially as described by Blake (1984) but including a 1 M salt wash after incubation with the primary and secondary antibodies. The ability of the anti-P IgG to inhibit the activity of glycine decarboxylase was assessed by immunoprecipitation of the P subunit and assaying the supernatant for enzyme activity (Walker and Oliver, 1986a) containing the 50–60% ammonium sulfate fraction (see above) and stained with Brilliant Blue R were placed directly onto the sequencer reaction cartridge and held in place with trifluoroacetic acid-etched glass fiber discs (Hunkapiller et al., 1983).

**Identification of a P Subunit cDNA Clone**—The P protein antibody was used to screen 10⁵ plaques from a pea leaf cDNA library constructed in λgt11. Positive clones (15) were isolated and one of these clones, which contained an insert of 3.4 kilobases, was subcloned into a plasmid vector for sequencing and named pGDP1. The sequence (Fig. 3) contained an open reading frame of 3171 nucleotides, and the amino acid sequence deduced from this included a sequence matching exactly the amino acid sequence that we had determined directly from the NH₂-terminus of the mature protein. The only

**RESULTS**

**Characterization of the P Subunit Antibody**—The antibody was judged to be monospecific on the basis that it gave only one band of M₉₇,₀₀₀ when used to probe Western blots of SDS-polyacrylamide of total pea leaf protein extract (see Fig. 4). The antibody was tested for its ability to immunoprecipitate the P subunit by measuring the inhibition of the GDC activity of a mitochondrial extract following immunoprecipitation of the P subunit. The results (Fig. 2) show that over 90% of GDC activity could be inhibited by the antibody, consistent with the antibody recognizing the P subunit of GDC.

**Fig. 1. Restriction map of the clone pGDC1.** Positions of the restriction sites are shown in brackets: BamHI (261), HindIII (1272, 1640), EcoRI (2662, 2706), PstI (3542). Arrows underneath represent fragments sequenced from independent deletions.

**Fig. 2. Inhibitions of GDC activity of mitochondrial extract using anti-P subunit IgG.** Inhibition is expressed relative to the bovine serum albumin (BSA) control and was measured by assaying the supernatant following immunoprecipitation of the P subunit. Immunoblots contained 15 μg of mitochondrial extract protein which gives an activity of 6 nmol min⁻¹. The amount of IgG was kept constant at 270 μg incubation⁻¹ by the addition of preimmune IgG. Each point represents the average of two independent measurements.
Cloning the P Subunit of Glycine Decarboxylase

FIG. 3. Nucleotide sequence of the clone pCDP1. The deduced amino acid sequence of the open reading frame is shown above in bold type, and the sequence matching the N-terminal amino acids determined directly from the mature protein (underlined).

The DNA sequence shows the presence of a 3'-untranslated region of 140 nucleotides downstream of two consecutive translation stop codons. Like many plant genes, there does not appear to be an AATAAA polyadenylation signal anywhere in this sequence. The most likely signal is the GATAAT motif located 13 nucleotides upstream from the poly(A) tail (Joshi, 1987).

Expression and Organization of the P Protein Genes—In order to examine not only the level of P protein mRNA expression but also that of other GDC subunits, all measurements of P subunit mRNA are accompanied by measurements of H protein mRNA levels on the same samples. In addition, to correlate mRNA levels with protein levels, Western blots of protein extracts made from the same plant material as used for the mRNA extractions were probed with the P and H subunit antibodies. Glycine oxidation occurs at a much faster rate in mitochondria isolated from leaves compared to those from other tissues (Gardestrom et al., 1980). We therefore examined the tissue-specific pattern of the GDC mRNA expression by probing Northern blots containing poly(A)+ mRNA from pea leaves, roots, and embryos with pGDP1 and pGDH1. The highest levels of mRNA expression are seen in the leaves, although for the P subunit in particular, a significant amount of mRNA can be found in other tissues (Fig. 4). In addition, for both P and H there is a strong correlation between the level of the mRNA and the level of the corresponding protein (Fig. 4).

Since it is known that the level of GDC in the mitochondria of etiolated seedlings increases dramatically on exposure to light (Day et al., 1985; Walker and Oliver, 1986b) and H subunit mRNA has previously been shown to be light inducible (Macherel et al., 1990; Kim and Oliver, 1990), the effect of light on the expression of the P subunit mRNA was examined. In order to standardize the results, mRNA levels were examined in total RNA and since the P subunit mRNA (3.7 kilobases) is masked on Northern blots by the very large upstream methionine residue, prior to an in frame stop codon, lies 86 amino acids upstream of the mature N-terminal amino acids determined directly from the mature protein and is assumed to be the translation start site. The 86 amino acid leader sequence shows many characteristics of a mitochondrial targeting sequence: rich in serine (20%) and arginine (15%) (von Heinje, 1985), with the regions of hydroxylated amino acids (including one run of 8 consecutive hydroxylated residues) interspersed with positively charged amino acids and an arginine residue at -2 relative to the cleavage site (Hendrick et al., 1989).
Cloning the P Subunit of Glycine Decarboxylase

FIG. 4. Northern blots (top) of pea poly(A)+ mRNA (2 μg/track) isolated from various tissues probed with the P clone pGDP1 or H clone pGDH1. Corresponding Western blots (below) (loaded to give 15 μg total protein/track) were probed with the P or H antibody.

FIG. 5. Accumulation of P (bottom) and H (top) subunit protein and mRNA in pea leaves following the exposure of etiolated seedlings to the light. Histograms represent the data obtained from probing dot-blots (see "Experimental Procedures") of total mRNA with either pGDP1 (P subunit) or pGDH1 (H subunit). Results are normalized by expressing the hybridization relative to the hybridization to the ribosomal RNA. The corresponding protein levels are shown as Western blots (loaded with 15 μg of total protein/track) probed with the P and H subunit antibodies.

FIG. 6. Changes in P and H subunit protein and mRNA in pea leaves during the exposure of mature plants to a period of 60 h of continual darkness followed by a period of continual illumination. Steady-state levels are shown for mature leaves before darkening (BD). Histograms represent the data obtained from probing dot-blots of total mRNA (see "Experimental Procedures") with either pGDP1 (P subunit) or pGDH1 (H subunit). Results are normalized by expressing the hybridization relative to the hybridization to the ribosomal RNA. The corresponding protein levels are shown as Western blots (loaded with 15 μg of total protein/track) probed with the P and H subunit antibodies.

FIG. 7. Copy number determination for the P subunit genes using a Southern blot probed with pGDP1. The three right lanes contained 10 μg of pea DNA digested with various restriction enzymes. The four left lanes contained loadings of pGDP1 digested to give 0.5, 1, 2, or 4 copies/haploid genome.

development, which may also contribute to the changes in gene expression observed. To try and separate these two factors, mature green plants were put in the dark for 60 h then re-exposed to light. After 60 h in the dark, the mRNA for both the P and H subunits had declined to a fraction (<2%) of their before-darkening levels (Fig. 6). After returning the plants to the light for 24 h, however, the mRNA level had returned to, or in the case of P subunit mRNA far exceeded, those of the level prior to dark treatment. Over the time course of the experiment, the Western blots indicated that there was very little change in the P or H subunit levels (Fig. 6).

In order to determine the copy number of the genes encoding the P subunit mRNA, Southern blots of pea genomic DNA as well as plasmid DNA markers were probed with pGDP1. The P subunit appears to be encoded by a small multigene family (Fig. 7) with approximately two genes for the P subunit/haploid genome in pea.

DISCUSSION

The clone (pGDP1) isolated contains an open reading frame which encodes the polypeptide sequence of the entire P sub-
Cloning the P Subunit of Glycine Decarboxylase

Fig. 8. Sequence comparison of the active site polypeptide isolated from the P subunit of chicken liver GDC (PLP peptide) with the deduced amino acid sequence of pGDP1. Arrows, lysine residue identified in the chicken liver sequence as the site of attachment of the pyridoxal phosphate group.

Unit of GDC. The deduced amino acid sequence together with our NH2-terminal amino acid data suggests that the mature protein contains 971 amino acids with a M, of 105,000. This is larger than previous estimates of 97,000-98,000 (Bourguignon et al., 1988; Walker and Oliver, 1986a) made from gel electrophoresis. The difference may be accounted for by the overall negative charge of the P subunit causing it to run anomalously on SDS-polyacrylamide gels. In addition, there is an unusually long 88 amino acid leader sequence, which is presumably for targeting the protein to the mitochondria. Long mitochondrial targeting sequences are generally associated with proteins which require further intramitochondrial sorting (see Schatz, 1987, for review). Available evidence, however, suggests that the P subunit is located in the mitochondrial matrix (Neuburger et al., 1986). The significance of the large targeting sequence and any special function it may perform awaits further functional analysis.

During the induction of the P subunit following the exposure of isolated seedlings to light, the level of the protein closely follows that of the corresponding mRNA: the same result is seen for the subunit (see also Macherel et al., 1990; Kim and Oliver, 1990), suggesting that under these conditions the protein level is controlled primarily by the level of corresponding mRNA. Following a period of 60 h of continuous darkness, the level of P subunit mRNA in the leaves of mature plants dropped sharply. The corresponding protein, however, showed relatively little change, implying that the P subunit is a stable protein, and little of it is turned over during this period. In comparison to the expression of the P subunit during light induction, the expression of the H subunit appears to show a lag and, although reasonable quantities of the P subunit mRNA are present in tissues other than leaves (roots or embryos), H subunit mRNA only appears at comparatively low levels. In the mitochondria of mature leaves, GDC exists as a stoichiometric complex and this stoichiometry presumably reflects that of the active enzyme. The reason why in other tissues, or at other stages of development/greening, GDC subunits appear to accumulate with different stoichiometry is unclear. A relationship between the levels of P and H subunit is implied from the work of Blackwell et al. (1990), who have characterized a GDC mutant from barley which shows reduced levels of the P and H subunits but apparently normal levels of the T and L subunits. Given that the P subunit is apparently stable in the absence of normal levels of the H protein, it appears that either the level of the H subunit is somehow governed by the level of the P subunit or that though the P and H genes show subtly different patterns of expression, they share a common regulatory factor.

Comparison of the deduced amino acid sequence shown here with the amino acid sequence of a pyridoxal phosphate-containing peptide from the GDC P subunit of chicken liver (Fujikawa et al., 1987) shows remarkable similarity (Fig. 8). Some 46 of the 54 amino acids are identical including a run of 25 consecutive amino acids, which includes a region rich in glucose just to the carboxyl side of the site of attachment of the pyridoxal phosphate. Given the high degree of sequence conservation it is probable that Lys392 in our sequence corresponds to the site of attachment of the pyridoxal phosphate. In common with all other characterized amino acid decarboxylases, this lysine is preceded by a histidine (Vaaler and Snell, 1989). The hydrophobic glycine-rich region is conserved in several amino acid decarboxylases and has been suggested to be involved in substrate binding (Moore and Boyle, 1990). In addition, Fujikawa et al. (1987) suggest that the large number of glycine residues are required to give the steric freedom required for access of the lipoyl moiety of the H subunit to the active site. Alignment of the complete amino acid sequences for the mature proteins of human and chicken (Kume et al., 1991) with that of pea shows that the similarity between them extends throughout the protein sequence. The amino acid sequence of the mature Pisum P subunit is 56% identical and 73% similar to the mature sequences of the human and chicken proteins which are more closely related (91% similar and 83% identical; Kume et al., 1991). Two notable regions of identity exist between all three proteins. The first is from Gin466 to Met479, and the second is from His530 to Pro549 in the Pisum protein. The second domain includes the lysine residue which binds the pyridoxal phosphate and the glycine-rich region which have been identified in the human and chicken proteins (Kume et al., 1991). Such strong conservation between unrelated species clearly suggests involvement of these domains in either the catalytic activity or in the interaction between subunits of the glycine decarboxylase complex. Apart from those mentioned above, the P subunit of GDC shows no obvious sequence similarity to any other characterized amino acid decarboxylases.

The P subunit of GDC shows several unique catalytic properties, such as its ability to catalyze the exchange of the glycine carboxyl carbon with CO2 (Hiraga and Kikuchi, 1989b). It is also characterized by its large size (M, 105,000) when compared to other amino acid decarboxylases (e.g. M, of 56,000 for tryptophan decarboxylase from Catharanthus roseus (De Luca et al., 1989)). Interestingly, bacterial GDC in which the polypeptide size and composition of the P subunit have been determined exists as a,68 tetramer containing polypeptides of M, 54,000-63,000 (Freudeaerg and Andreesen, 1989; Gariboldi and Drake, 1984). Comparison of the sequence of a bacterial enzyme with that of the P subunit from Pisum may shed some light on the evolution of the eukaryotic GDC P subunit.

Acknowledgments — We thank Dr. Pat Barker (AFRC Institute of Animal Production and Grasslands Research, Babraham, Cambridge) for performing the NH2-terminal amino acid sequencing, Dr. M. Trick and Dr. G. Murphy for technical advice, and Dr. A. Smith, S. Collin, and Dr. D. Murphy for useful comments on the manuscript.

REFERENCES

Blackwell, R. D., Murray, A. J. S., and Lea, P. J. (1990) Plant Physiol. 94, 1316-1322
Blake, M. S. (1984) Anal. Biochem. 136, 175-179
Bourguignon, J., Neuburger, M., and Douce, R. (1988) Biochem. J. 255, 169-178
Clandinin, M. T., and Cossins, A. E. (1975) Biochemistry 14, 387-391
Dey, A. D., Neuburger, M., and Douce, R. (1985) Aust. J. Plant Physiol. 12, 219-225
Domoney, C., and Casey, R. (1987) Nucleic Acids Res. 3, 687-699
De Luca, V., Marieau, C., and Brisson, N. (1987) Prog. Natl. Acad. Sci. U. S. A. 86, 2582-2586
De Vries, S., Hoge, H., and Biesseling, T. (1986) in Plant Molecular Biology Manual (Gelvin, S. B., and Schilperoot, R. A., eds) B47/1-13, Kluwer Academic Publishers, Dordrecht, Belgium
Eyz, P. L., Prowse, S. J., and Jenkins, C. R. (1978) Biochemistry 15, 429-436
Cloning the P Subunit of Glycine Decarboxylase

Freudenberg, W., and Andreesen, J. R. (1989) J. Bacteriol. 171, 2209-2215
Fujiwara, K., Okamura-Ikeda, K., and Motokawa, Y. (1986) J. Biol. Chem. 261, 8836-8841
Fujiwara, K., Okamura-Ikeda, K., and Motokawa, Y. (1987) Biochem. Biophys. Res. Commun. 149, 621-627
Fujiwara, K., Okamura-Ikeda, K., and Motokawa, Y. (1990) J. Biol. Chem. 265, 17463-17467
Gariboldi, R. T., and Drake, H. L. (1984) J. Biol. Chem. 259, 6085-6089
Gardestrom, P., Bergman, A., and Ericson, I. (1980) Plant Physiol. 65, 387-391
Hendrick, J. P., Hodge, P. E., and Rosenberg, L. E. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 4056-4060
Hiraga, K., and Kikuchi, G. (1980a) J. Biol. Chem. 255, 11671-11676
Hiraga, K., and Kikuchi, G. (1980b) J. Biol. Chem. 255, 11664-11670
Hunkapillar, M. W., Hewick, R. M., Dreyer, W. J., and Hood, L. E. (1983) Methods Enzymol. 91, 399-413
Joshi, C. P. (1987) Nucleic Acids Res. 23, 9627-9640
Kim, Y., and Oliver, D. J. (1990) J. Biol. Chem. 265, 848-853
Kim, Y., Shah, K., and Oliver, D. J. (1991) Physiol. Plant. 81, (1991) 501-506
Kume, A., Koyata, H., Sakakibara, T., Ishiguro, Y., Kure, S., and Hiraga, K. (1991) J. Biol. Chem. 266, 3323-3329
Laemmli, U. K. (1970) Nature 227, 680-685
Macherel, D., Lebrun, M., Gagnon, G., Neuburger, M., and Douce, R. (1990) Biochem. J. 268, 783-789
Moore, R. C., and Boyle, S. M. (1990) J. Bacteriol. 172, 4631-4640
Moos, M., Nguyen N. Y., and Liu T.-Y. (1988) J. Biol. Chem. 263, 6005-6008
Motokawa, Y., and Kikuchi, G. (1974) Arch. Biochem. Biophys. 164, 624-633
Murphy, G., and Kavanagh, T. A. (1988) Nucleic Acids Res. 16, 5198
Neuberger, M., Bourguignon, J., and Douce, R. (1986) FEBS Lett. 207, 18-22
Okumura-Ikeda, K., Fujiwara, K., Yamamoto, M., Hiraga, H., and Motokawa, Y (1991) J. Biol. Chem. 266, 4917-4921
Oliver, D. J., Neuburger, M., Bourguignon, J., and Douce, R. (1990) Plant Physiol. 94, 833-839
Rosie, D., Howarth, S. J., Macdonald, P., and Schatz, G. (1986) EMBO J. 5, 1327-1334
Schatz, G. (1987) Eur. J. Biochem. 165, 1-6
von Heijne, G. (1985) J. Mol. Biol. 184, 1327-1334
Vaaler, G. L., and Snell, E. E. (1989) Biochemistry 28, 7306-7313
Walker, J. L., and Oliver, D. J. (1986a) J. Biol. Chem. 261, 2214-2221
Walker, J. L., and Oliver, D. J. (1986b) Arch. Biochem. Biophys. 248, 626-638
White, B. A., and Bancroft, F. C. (1982) J. Biol. Chem. 257, 8569-8572