Spinach Thylakoid Polyphenol Oxidase: Cloning, Characterization, and Relation to a Putative Protein Kinase[†,Δ]

Geoffrey Hind,[*,‡] Daniel R. Marshak,[§,||] and Sean J. Coughlan[‡,⊥]

Biology Department, Brookhaven National Laboratory, Upton, New York 11973, and Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724

Received December 19, 1994; Revised Manuscript Received March 21, 1995[⊗]

ABSTRACT: A 64-kDa protein was purified from an octyl glucoside/cholate extract of spinach thylakoids. N-Terminal analysis yielded 23 residues of sequence, of which the first 15 were identical to a sequence reported [Gal, A., Herrmann, R. G., Lottspeich, F., & Ohad, I. (1992) *FEBS Lett. 298*, 33−35] for a protein kinase with specificity toward the photosystem II light-harvesting complex (LHC-II). We report the complete sequence of this 64-kDa protein, deduced from cDNA clones. The transit peptide has a chloroplast import signal at the N-terminus and a C-terminal hydrophobic span bounded by basic amino acids that predicts localization of the protein to the thylakoid lumen. The mature protein sequence is about 50% identical to several polyphenol oxidases (PPOs). Canonical protein kinase motifs are absent, as are sequences characteristic of ATP-binding sites. The mature protein resembles arthropodan hemocyanin (Hc), possessing three major domains. The N-terminal domain is rich in cysteine residues and predicted α-helices. The central domain has a conserved motif, N-terminal to a presumptive Cu-A site, that is not found in tyrosinases or Hc and is proposed as the provider of a third imidazole ligand to Cu-A. An unusual 13-residue, glutamine-rich link begins a C-terminal domain containing 7 predicted β-strands which, by analogy with Hc, may form an antiparallel β-barrel. We conclude that this 64-kDa polypeptide is a lumenal PPO and the precursor of a 42.5-kDa PPO form described previously [Golbeck, J. H., & Cammarata, K. V. (1981) *Plant Physiol. 67*, 977−984]. In view of its lumenal location and primary sequence, it is unlikely to be a serine/threonine protein kinase.

Polyphenol oxidases (PPOs)[1] are copper metalloenzymes occurring in the chloroplasts of most angiosperms (Sherman et al., 1991). They catalyze the oxidation of various monophenols and *o*-diphenols to the corresponding quinones, which further react and polymerize to give brown products that are believed to protect damaged tissues from further destruction by herbivores and pathogens (Constabel et al., 1995). In undamaged tissues, PPO activity is latent; the isolated enzyme can be activated *in vivo* by partial proteolysis (Robinson & Dry, 1992) or by mild denaturants [reviewed by Mayer and Harel (1979)]. Since the catalytic activity of PPO survives during SDS−PAGE, and SDS is itself an excellent activator (Golbeck & Cammarata, 1981), PPO is conveniently detected by self-staining in the gel, using an infused chromogenic substrate (Sherman et al., 1991). PPOs so recognized and documented in the early literature have masses in the 40−50-kDa range, although these active forms are now thought to arise by proteolysis of 60−70-kDa precursors (Robinson & Dry, 1992). At least in the Solanaceae, PPO is expressed in tissue-specific isoforms (Newman et al., 1993; Hunt et al., 1993) having masses above 60 kDa that show high conservation of primary sequence, particularly in the binuclear Cu-A and Cu-B sites forming the presumptive catalytic center.

Activation *in vivo* is also achieved with free fatty acids, among which linolenic acid is particularly effective (Golbeck & Cammarata, 1981). The release of linolenic acid from membrane lipids *in vivo*, in response to wounding (Farmer & Ryan, 1992), is now viewed as the starting point for the jasmonic acid signaling pathway that promotes defensive processes such as the synthesis of proteinase inhibitors. Possibly the activation of PPO by linolenic acid occurs *in vivo* and provides an additional, prompt component of this defense; however, the actual *in vivo* mechanism and state of activation of PPO are unknown. In the functionally related tyrosinase (monophenol oxidase) from *Neurospora* (Kupper et al., 1989), activation involves the removal, by a chymotryptic type of protease, of a C-terminal portion of the 75-kDa proenzyme to yield active 46-kDa tyrosinase. Partial activation of PPO by proteolysis and loss of the C-terminal domain (Robinson & Dry, 1992) suggests that access to the catalytic center is comparably controlled in PPOs. The relation between activations of PPO by proteolysis and by fatty acids, as discussed earlier, is unclear. Unfortunately, no three-dimensional structure for a PPO or a tyrosinase is

[†] This research was carried out at Brookhaven National Laboratory under the auspices of the U.S. Department of Energy, with funding from its Office of Basic Energy Sciences, Division of Energy Biosciences, and from PHS Grant CA-13107 (to D.R.M.) and American Cancer Society Grant CB-72 (to D.R.M.).

[Δ] The sequence of tp64 has been entered in GenBank under accession number U19270.

* Author to whom correspondence should be addressed.

[‡] Brookhaven National Laboratory.

[§] Present address: Osiris Therapeutics, Inc., 11100 Euclid Ave., Cleveland, OH 44106.

[||] Cold Spring Harbor Laboratory.

[⊥] Present address: Pioneer Hi-Bred International, 7250 NW 62nd Avenue, Johnston, IA 50131.

[⊗] Abstract published in *Advance ACS Abstracts,* June 1, 1995.

[1] Abbreviations: Hc, arthropodan hemocyanin; LHC-II, light-harvesting complex of photosystem II; OEC, oxygen-evolving complex; OG, *n*-octyl β-D-glucopyranoside; PCR, polymerase chain reaction; PPO, polyphenol oxidase; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis.

0006-2960/95/0434-8157$09.00/0

available to aid in conceptualizing either mechanism of activation.

An alternative approach to acquiring structural insight is suggested by marked homologies in the motifs responsible for Cu binding in the primary sequences of PPOs, tyrosinases, and hemocyanins, which probably descended from a common ancestor with an archetypical binuclear copper center (Drexel et al., 1987). Commonalities is three-dimensional structure among disparate classes of protein are increasingly being recognized and explored by using improved analytical routines for secondary structure prediction, such as the PHD program (Rost & Sander, 1993, 1994). PHD performs neural network analysis on a database derived by merging two- and three-dimensional structure information with primary sequence homologies. With the spinach PPO primary sequence as input, PHD suggested secondary structures that could be matched to those of the *Panulirus* hemocyanin subunit (Hc), for which a crystal structure is known at 3.2-Å resolution (Gaykema et al., 1984; Volbeda & Hol, 1989). In this way, a tentative model for the gross three-dimensional structure of PPO was derived, and it is described here.

The cellular location of an enzyme and its substrates is also important to evaluating possible function(s) and mode of activation. Cytochemical staining (Vaughn & Duke, 1984) and immunoelectron microscopy (Lax & Vaughn, 1991) show that PPO is associated with PS-II in the thylakoid grana stacks. Moreover, the precursor of tomato leaf PPO can be imported into the lumen of pea thylakoids and processed to a soluble, 59-kDa protein (Sommer et al., 1994). In contrast to these convincing findings on the location of the enzyme, the identity and subcellular compartmentation of PPO substrates are little understood [see reviews by Halliwell (1978) and Mayer (1987)]. PPO, in addition to being latent, may not have access to substrate *in situ* as a further restriction on its activity in undamaged tissues.

None of the leaf tissues (tomato, potato, broad bean) from which PPO clones are described is favorable to subcellular fractionation toward the isolation of latent enzyme and characterization of substrates and their distribution. The early study of spinach chloroplasts by Golbeck and Cammarata (1981) uncovered important details concerning PPO activation, and an oligomeric form of the enzyme was described; however, there was no subsequent attempt to exploit the extensively studied spinach chloroplast system or to characterize a cDNA clone for spinach PPO.

Additional interest in cloning spinach PPO stemmed from the discovery that a purification protocol for a thylakoid protein kinase also yielded fractions enriched for PPO activity (F. Lajko and G. Hind, unpublished data). Since the kinase is redox-regulated *in membrano*, and since PPO possesses both a redox center and a thylakoid location, it seemed possible that these attributes might be combined in a single molecule: possibly a PPO isoform or precursor that functions as a protein kinase.

Gal et al. (1990) fractionated a kinase-enriched preparation of spinach cytochrome b_6/f complex on SDS−PAGE and used a protein band, excised from the 64-kDa region of the gel, in order to raise a polyclonal antibody against a 64-kDa kinase. Immunobinding studies showed a physical association of the corresponding antigen with the cytochrome b_6/f complex, while immunoelectron microscopy further demonstrated its presence in marginal regions of the thylakoid grana stacks. N-Terminal sequence analysis was performed (Gal et al., 1992) on a 64-kDa band from an electroblot, corresponding to the material used as an antigen in the earlier work. This yielded the sequence APILPDVEKSTLSDA, which the researchers claimed to be the N-terminus of the 64-kDa protein kinase (Gal et al., 1992). In contrast to this literature suggestion (Gal et al., 1992), however, we were never satisfied with this proposition and sought confirmatory evidence from cDNA cloning. The protein defined by the preceding N-terminus is designated tp64 in this report.

EXPERIMENTAL PROCEDURES

Isolation of tp64. Thylakoids were isolated from 1 kg of spinach leaves, salt-washed, and extracted with OG/cholate; the supernatant was clarified and fractionated with ammonium sulfate essentially as described (Coughlan & Hind, 1986a). The pellet (100 mg of protein) was redissolved in 5 mL of 30 mM OG/12.5 mM cholate, 10 mM tricine/NaOH (pH 8.0), 0.1 mM EDTA, and 0.1 mM phenylmethanesulfonyl fluoride for 1−2 h on ice. The dissolved protein (10 mL) was desalted by passage over a Sephadex G-25 column (4 × 12 mL) equilibrated with 12.5 mM OG/5 mM cholate, 10 mM tricine/NaOH (pH 8.0), 0.1 mM EDTA, and 0.1 mM phenylmethanesulfonyl fluoride (desalting buffer). The eluate (20 mL) was loaded onto a hydroxylapatite column (5 × 8 cm) equilibrated in desalting buffer. Bound protein kinase activity was eluted from the column by desalting buffer containing 50 mM KH_2PO_4 (pH 8.0). The eluate (200 mL, 4 mg of protein) was concentrated by dialysis against desalting buffer minus detergent and phosphate and was rechromatographed on a 5-mL (1.3 × 4 cm) hydroxylapatite column. Bound protein (4 mg) was eluted as before. The solution was concentrated and desalted on Centricon-30 (Amicon Corp.) microconcentrators, and the product (1 mL) was loaded on a Superose-12 (Pharmacia) column (3 × 50 cm) preequilibrated in desalting buffer minus detergent. The tp64 eluted as an apparently homogeneous monomer in yields of 100−250 μg of protein. It lacked histone kinase activity and PPO activity as isolated.

Preparation of Probes and Antibody. N-Terminal sequencing of purified PPO was performed on an Applied Biosystems 475A sequencer using gas phase delivery of acid (R1) and base (R2), according to the manufacturer's specifications. The protein was not reduced and alkylated. The overall average repetitive yield was >90% and absolute initial yield was approximately 70%, giving 23 residues of sequence *N*-APILPDVEKSTLSDALWDGSVGD. The residue in cycle 10 was detected as phenylthiohydantoin−dehydroalanine, as derived from serine or cysteine by β-elimination. A synthetic peptide representing the 15 N-terminal residues was conjugated to ovalbumin, and antisera to the conjugate were raised in rabbit. On immune blots (not shown) of OG/cholate extract and purified tp64, the antiserum reacted with tp64 and also with a presumed proteolytic fragment (ca. 42 kDa). Synthetic oligonucleotide primers complementary to amino acid residues 3−10 and 14−20 were synthesized and used as primers for PCR on a template of first-strand cDNA, obtained by reverse transcription from 1 μg of spinach leaf mRNA (Ausubel et al., 1990) using an oligo(dT) primer. Twenty-five cycles (1 min, 94 °C; 2 min, 52 °C; 3 min, 72 °C) yielded the predicted 54-nt product, which was sequenced chemically in both directions. With this knowledge, a unique 29-mer, 5′-GAAAAATG-

TACCCTGAGTGACGCCCTATG-3′, was synthesized for use as a probe in screening expression libraries.

Preparation and Isolation of Clones. Spinach primary leaf mRNA was used as the starting material to construct a cDNA library in the bacteriophage λgt11 (Sambrook et al., 1989; Ausubel et al., 1990). Independent recombinants (4×10^5) were screened with the 5′-^{32}P-labeled oligonucleotide probe by conventional procedures. Positive clones were sequenced directly on the phage DNA (both strands) by the dideoxynucleotide method, using synthetic oligonucleotides as primers and T7 DNA polymerase (USB Sequenase version 2.0). Seven partial clones were obtained, which were all truncated at the 3′ end. To obtain the missing 3′ sequence, another oligonucleotide was synthesized complementary to the 3′ end of the longest insert. This was used to screen another 4×10^5 independent recombinant clones, and three partial clones possessing a poly(A) tail were obtained. The longest inserts from both rounds of screening proved to be almost full length: sc8-1 (nt 1-1804) and sc12-2 (nt 453-2065). The entire cDNA sequence, including a 78-nt 5′ extension, the preprotein, and a poly(A) tail, was deduced by DNA sequencing of these 10 clones through regions of overlap. This sequence has been entered in GenBank as accession number U19270. cDNAs corresponding to the preprotein and its mature form, with flanking *Eco*RI, *Nde*I(5′), and *Bgl*II(3′) restriction sites, were obtained by using PCR reactions (four cycles of 1 min 94 °C, 30 s 53 °C, 30 s 22 °C and then 32 cycles 20 s 94 °C, 30 s 53 °C, 5 min 72 °C) catalyzed by Vent DNA polymerase (New England Biolabs), with total cDNA as a template and the following primers: (i) for preprotein, 5′-TTGAATTCATATGGCAACTCTCTCT-TCTCCGACC (sense) and 5′-GCCAGATCTTCACTCGAT-GTCGATACCAGTGAT (antisense); (ii) for mature protein this same antisense primer and a forward primer that silently mutates the second codon, GCC, to the GCT preferred for expression in *Escherichia coli* (Looman et al., 1987), 5′-TTGAATTCATATGGCTCCCATTCTCCCTGACGTGG. After restriction with *Nde*I and *Bgl*II, the cDNAs were ligated between the *Nde*I and *Bam*HI cloning sites of the expression vector pET 13a [pET-11a (Studier et al., 1990), in which the selective marker confers kanamycin rather than ampicillin resistance], taking advantage of the compatability of *Bgl*II and *Bam*HI cohesive ends. This was necessary to avoid restriction of the insert at a unique *Bam*HI site (nt 695). Plasmids were isolated from transformed *E. coli* DH5α by conventional means (Sambrook et al., 1989), and the inserts were sequenced in both directions. *E. coli* host strains BL21-(DE3) and BL26(DE3) were transformed with these clones (RT11, preprotein; RZ22, mature protein), and expression of the foreign protein was induced by IPTG (Studier et al., 1990).

Chemicals and Supplies. OG was obtained from Jersey Lab & Glove Supply (Livingston, NJ). Cholate and other fine chemicals were from Sigma (St. Louis, MO). Protein chemistry chemicals were obtained from Applied Biosystems (Foster City, CA) and from Aldrich Chemical (St. Louis, MO). Solvents were from Baxter Diagnostics (McGraw Park, IL). Restriction and DNA-modifying enzymes were from New England Biolabs (Beverly, MA).

RESULTS

Description of tp64 cDNA and Gene. The deduced cDNA of ca. 2.2 kb encodes 78 nt of 5′ untranslated DNA having an in-frame TAA stop codon at nt −30, a canonical plant ATG initiation codon (at nt + 1) at the start of an open reading frame of 1917 nt, a TGA in-frame stop codon at nt 1918, and 167 nt of untranslated 3′ DNA followed by a poly-(A) tail (70 nt). Northern blot analysis of RNA from spinach leaf revealed a single strongly hybridizing band at 2.5 kb (Figure 1A). The deduced cDNA thus was essentially full length, having complete 3′ information, and missing ca. 300 nt of sequence comprising the 5′ cap. tp64 mRNA is present in the leaf; identical amounts of mRNA from the root yield no apparent positive signal. Furthermore, the mRNA appears to be light-regulated, since no positive signal was elicited when etiolated cotyledon or shoot tissue mRNA was probed (prior to floral induction, the shoot (stem) of mature spinach plants is compact, heavily shaded, and almost free of chlorophyll). Hybridization of the probe with total RNA was weak, giving a signal 1−2 orders of magnitude lower (not shown) than a probe to FNR, a major thylakoid protein whose mRNA is present at about 1% abundance in young greening tissue (Newman & Gray, 1988; Jansen et al., 1988); thus, tp64 mRNA is present in low copy number (low multiple copy/cell).

Genomic spinach DNA, prepared as in Ausubel et al. (1990), was digested with six restriction enzymes, and the corresponding blot was challenged with a 1.76-kb probe corresponding to the central part of tp64 cDNA. Only one strongly hybridizing band was observed under high-stringency conditions (Figure 1B), confirming that the gene encoding tp64 is present at low or single copy number.

Primary Sequence of tp64. The single open reading frame encodes a protein of 639 residues (MW = 73 255). The N-terminus of the mature protein is alanine 102; thus, the first 101 residues are the presumptive transit peptide, controlling import into the chloroplast and targeting to the thylakoid. The mature protein, shown in Figure 2, consists of 538 residues (MW = 62 243), in reasonable agreement with the electrophoretic mobility of the native enzyme on SDS−PAGE of ca. 64 kDa. The calculated amino acid compositional analysis and the p*I* of 5.4 also agree with those obtained experimentally from the native protein. The sequence from residue 1 to 23 was that derived by direct sequencing of the mature protein with the exception of C^{10}, which was detected as dehydroalanine (a derivative of either cysteine or serine) and originally denoted serine.

A database search (Altschul et al., 1990) against the entire mature protein revealed extensive homology to PPOs. Complete sequences of four representative mature PPOs are compared with tp64 in Figure 2. After optimal alignment, about one-half of the amino acids in tp64 are identical with those in other PPO sequences. Identity scores (not shown) for tp64 are only slightly less than those of a *Vicia faba* PPO when comparisons are drawn with closely interrelated PPOs from members of the Solanaceae (tomato and potato). The lower score is attributable to the presence in tp64 of a stretch of 11 glutamine residues interposed with two arginines beginning at Q^{359}, QQQQQQQRQQQRQ, which is the predicted point of maximum chain flexibility (Karplus & Schultz, 1985). In plants, glutamine runs are found in seed storage proteins (Ambler, 1991), yeast Cks1 protein (Altschul et al., 1990), and *Arabidopsis thaliana* TSL protein kinase (Roe et al., 1993), but not so far in other PPO sequences. The C-terminal region is rich in basic residues (theoretical p*I* = 7.7).

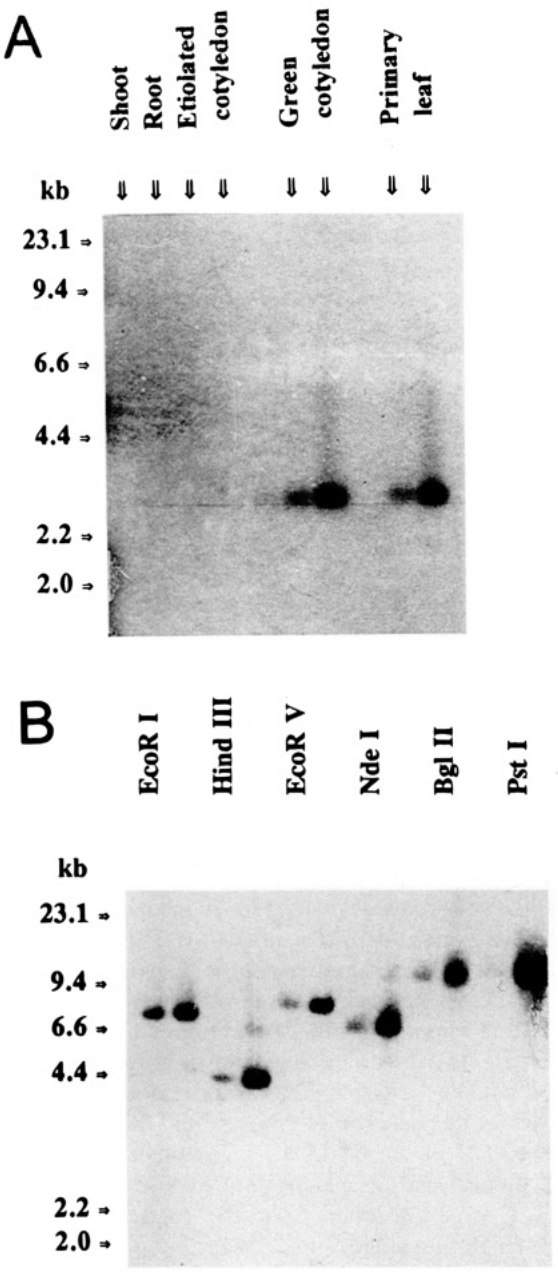

FIGURE 1: (A) Tissue-specific distribution of tp64 mRNA. Poly-(A)$^+$ RNA was isolated from plant tissues (Ausubel et al., 1990), fractionated on a 1% agarose−formaldehyde gel, and transferred to Nytran (Schleicher & Schuell). The blot was probed with a ^{32}P-labeled, random-primed probe derived from a PCR-amplified product (nt 132−1894) of the cDNA. The blot was rinsed at high stringency (65 °C), and radioactive areas were detected by autoradiography. Molecular weight markers denoted on the left side of the autoradiograph are from a *Hin*dIII digest of λ DNA. For shoot and root, 5 μg of poly(A)$^+$ RNA was loaded; in other samples, adjacent lanes contained 5 μg of total RNA (left) and 5 μg of poly(A)$^+$ RNA (right). (B) Southern blot analysis of genomic DNA. Aliquots of spinach leaf DNA [1 (left) or 4 μg (right)] were exhaustively digested with the denoted restriction enzyme, and the products were separated on a 0.75% agarose gel. The gel was denatured and neutralized, and the products were transferred to Nytran by conventional methods (Sambrook et al., 1989). The blot was probed and hybridizing regions were visualized as in (A). Markers are the same as in (A).

Conserved Motifs and Domains. In overall design, tp64 and PPOs resemble the ~75-kDa subunit of arthropodan hemocyanin (Hc); there is an N-terminal domain (tp64 residues 1−82), a C-terminal domain (tp64 residues 354−538), and between these a putative cataytic domain (Figure 2). Motifs conserved in PPO sequences and in tp64 are assigned bold numerals I−VII in Figure 2 and for convenience in discussion are shown as contiguous. The putative copper-binding motifs are given letter designations (A, B(I), B(II)). The C-terminal domain is discussed later in terms of its secondary, rather than primary motif structure.

Catalytic Center. In arthropodan Hc, the three-dimensional structure (Gaykema et al., 1984) in the vicinity of the Cu-A- and Cu-B-binding sites predicts that each Cu atom is liganded by two closely spaced histidine residues and a third that is remote and closer to the C-terminus. In tp64, there are two histidines in the putative Cu-A region, at positions 113 and 122 (Figure 2), and a third at position 151 that could serve as ligands according to the Hc model. However, the *Vicia faba* sequence is unusual in lacking a histidine residue equivalent in context to H^{151}. In tp64, H^{177} (Figure 2) is an alternative possible downstream ligand, but this histidine is lacking in tomato PPO (tom B). The conserved H^{92} in motif IV of tp64 and the PPOs shown in Figure 2 consequently is a better candidate for the third Cu-A ligand. Our assignment of motif IV within the catalytic domain reflects this proposed role and is thus a departure from the Cu center structure known for Hc (Volbeda & Hol, 1989). The conserved motif IV sequence QQA-2x-HC(A,T)YC-x-G-x-Y-3x-G is proposed here as a specific signature for the PPO class (threonine replaces alanine in *Vitis vinifera* PPO). The more N-terminal conserved histidine, H^{55}, in motif III is a viable alternative candidate for the third ligand to Cu-A. The Cu-B centers of tp64, PPOs, tyrosinases, and hemocyanins are highly conserved, as discussed elsewhere (Drexel et al., 1987; Shahar et al., 1992; Hunt et al., 1993).

Cysteine Residues. Larger proteins are often stabilized by internal disulfide bridging, and two such structures link polypeptide folds in the Hc C-terminal domain; thus, the existence of disulfide bridges in PPOs might be anticipated. The PPO sequences shown in Figure 2 contain from 6 to 9 cysteines/mol; in tp64, five of the six residues are N-terminal to the putative Cu-A site. Solanaceous PPOs are distinct in also having a C-terminal cysteine or cysteines in a C-terminal extension (Newman et al., 1993), which are matched by additional cysteines in motifs VI and VII (Figure 2). A cysteine also occurs within the Cu-B(II) site (Figure 2) of all but *V. faba* PPO. Although some of these cysteine residues may be involved in internal disulfide bonds, the overall picture shows no consistent pattern indicative of essential cross-bridges within the catalytic and C-terminal domains. The vicinal cysteines in motif II, CCPP, are, by contrast, well-conserved (as CC(P,A)(P,V,D)) among PPOs and animal (but not fungal) tyrosinases. Conserved cysteines also are found in the N-terminals motifs I and IV of tp64 and PPOs.

Secondary Structure Predictions. Profile-fed neural network analysis (Rost & Sander, 1993, 1994) suggests there are seven β-strands in the C-terminal domain, as indicated in Figure 2. The crystal structure of Hc (Gaykema et al., 1984; Volbeda & Hol, 1989) reveals that a comparable cluster of seven strands, involving different primary sequence, folds to form an antiparallel β-barrel. Comparable C-terminus architecture was predicted for fungal tyrosinase (Kupper et al., 1989). Thus, it seems reasonable to conclude that PPO C-termini share this configuration. An additional point of similarity with Hc is that the helix in motif VIII and an immediately preceding short helix interrupt between

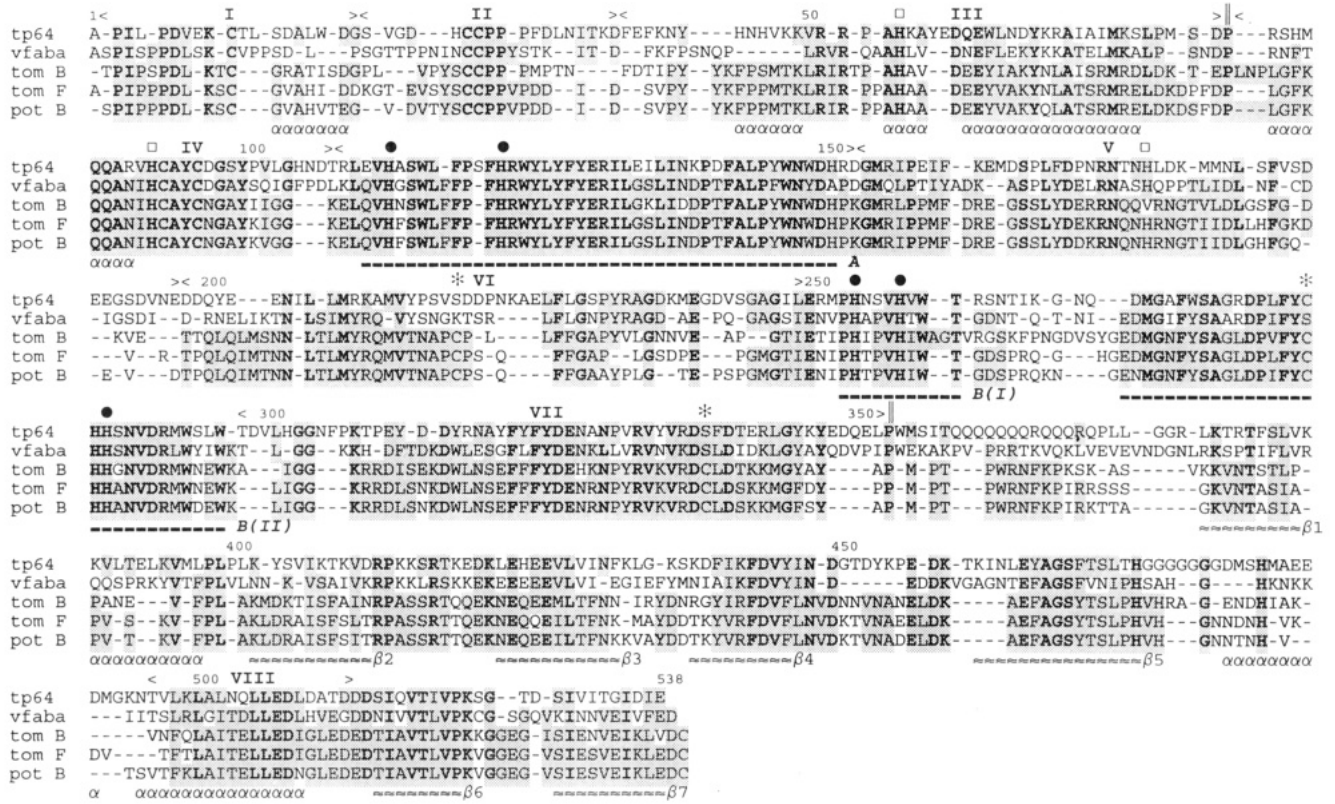

FIGURE 2: Homologies among deduced protein sequences of spinach tp64 and four PPO clones from other species, showing domain organization and secondary structure predictions. The PPO clones and Genbank accession numbers are as follows: vfaba, *Vicia faba* (Z11702); tom B and tom F, *Lycopersicon esculentum* (Z12834 and Z12838, respectively); pot B, *Solanum tuberosum* (M95197). Amino acid numbering is for tp64. Homologies were determined using the program ALIGN and manually optimized. Shaded areas show partial (60% or more) identity or homologous substitutions, while boldface letters indicate identity in all clones shown (though not necessarily in all known PPO sequences). Symbols: ||, proposed boundary between major domains; < and >, boundary between motifs (I−VIII); boldface dashed underlining, putative Cu-A and Cu-B sites; ●, presumed Cu-liganding histidine; □, potential ligand to Cu-A. Cysteine residues (*) are discussed in the text. Putative α-helices (α) and β-strands (≈β) are indicated only where relevant to the text.

strands 5 and 6 of the barrel in both proteins. In Hc, these helical and coiled structures project from the barrel and interact with the α-helices of the N-terminal domain (Volbeda & Hol, 1989); in PPOs, such interactions could occur with helices in motifs I, III, and IV (Figure 2). The motif VIII helix is amphipathic, with a hydrophobic moment of 0.33 (Eisenberg et al., 1984), as is the helix involving residues 388−397 (KVLTELKVML). Only the former, shown in Figure 3 as a helical wheel, is conserved among PPOs at the primary sequence level.

Transit Peptide Primary Sequence. Newman et al. (1993) found that all PPO isoforms from tomato possess a transit peptide of ca. 9 kDa, with a primary sequence characteristic of proteins targeted to the thylakoid lumen. General features of such peptides (Keegstra et al., 1989) are an N-terminal M(A,S)(S,T,Q) beginning a serine/threonine-rich hydrophilic stretch, followed by a hydrophobic region near the cleavage site. These are designated the chloroplast import domain and the thylakoid transfer domain, respectively (Smeekens et al., 1990). Homology is seen (Figure 4) in a comparison of tp64 with sequences of PPO precursors, especially at the conserved KVSC and at RR followed by a hydrophobic stretch. This stetch, in the thylakoid transfer domain, is predicted (Rao & Argos, 1986) to contain a membrane-spanning α-helix. Figure 4 shows that the lumenal 23-kDa protein of the oxygen-evolving complex (OEC) shares, in its shorter transit peptide, the key features of tp64 and PPO transit sequences. The transit peptide of the 16-kDa OEC protein (not shown) is also comparable, with a 16-amino acid

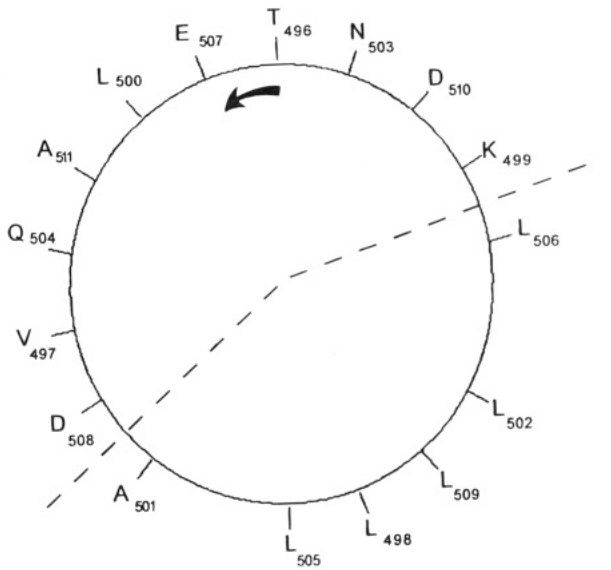

FIGURE 3: Helical wheel representation of the amphiphathic α-helix in tp64 motif VIII. The dashed line shows the effective boundary between hydrophobic and charged faces of the helix.

hydrophobic stretch bracketed by SRR at its N-terminal side and KAVLA at the C-terminal cleavage site.

Candidacy as a Protein Kinase. Eukaryotic protein kinases possess a number of characteristic primary sequence motifs; those in the catalytic site are highly conserved (Hanks et al., 1988). tp64 has no significant sequence homology with this enzyme class. Thus, motif I, G-x-G-2x-(G,S)-15x-

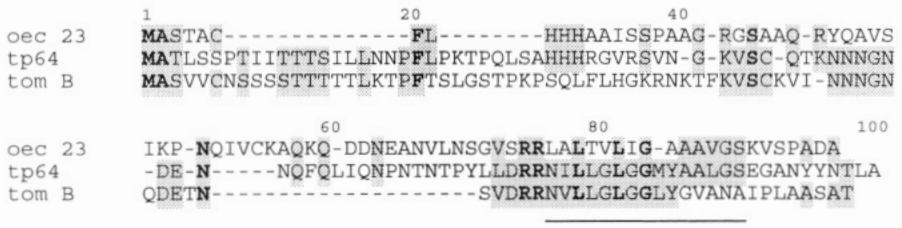

FIGURE 4: Comparison of transit peptide sequences among tp64, tomato PPO, and the 23-kDa protein of the OEC. Partial identities or homologous substitutions are shaded; identities are shown in boldface type. The hydrophobic region presumed to be involved in lumen targeting is underlined.

K, close to the N-terminus of protein kinases is absent, as is the less variable kinase motif VII, (I,L,V)-x-DFG, or in some plant protein kinases (Hayashida et al., 1993), (I,L,V)-x-DFD. Unusual serine/threonine protein kinases appear in the GenBank database that lack, for example a canonical motif I, but these proteins show sufficient homology in other conserved features to qualify for tentative membership in the class. Such is not true of tp64. We also found no sequence homology between tp64 and the phosphate-binding loop (Saraste et al., 1990) of ATP-binding proteins or homology with hexokinase (Adams et al., 1991). Parallel searches were made for each of the seven isoforms of thylakoid lumenal PPO described from tomato (Newman et al., 1993) without uncovering an instance of significant homology with serine/threonine protein kinases or ATP-binding proteins.

Expression of cDNA Clones. Inclusion bodies were produced from clones for the mature or precursor protein, regardless of the strategy adopted (individually or in combination) to restrict the rate of expression, such as suboptimal growth temperature, suboptimal inducer concentration with strain BL26(DE3), or osmotic stress. Similar results were obtained upon the expression of fusion constructs of clone RT11 in pMAL vectors (New England Biolabs).

DISCUSSION

tp64 and PPO Primary Sequences. tp64 is clearly identifiable as a PPO on the basis of the 50% sequence identity with PPOs from three other species. Its apparent mass is close to that predicted from the sequence. Historically, a confusing factor in the consideration of PPO size is the tendency of PPOs to run at abnormally low masses on SDS−PAGE unless adequate reduction of SS cross-links is achieved (Robinson & Dry, 1992; Cary et al., 1992). In the case of spinach PPO (Golbeck & Cammarata, 1981), the reported mass of 42.5 kDa is more probably a result of cleavage from tp64 by endogenous protease activity, since protease inhibitors were not included in earlier isolation protocols. In support of this view, we find that our tp64 preparations contain a minor 42-kDa protein that is recognized by antiserum to the tp64 N-terminus (data not shown), implying cleavage of tp64 near the junction of the catalytic and C-terminal domains. A thermolysin cleavage site in *V. faba* PPO has been identified (Robinson & Dry, 1992) after the β1-strand (Figure 2), yielding an active 42-kDa PPO with an intact N-terminus and an 18-kDa peptide commencing with VRQQSP. The significance of proteolytic cleavage of PPOs to their activation *in vivo* is uncertain. Detergent activators such as SDS give much higher V_{max} values (Robinson & Dry, 1992), which is possibly indicative of a predominant role for wound-inducible activators such as linolenic acid.

PPOs ranging between 160 and 220 kDa accompany lower mass forms in the leaves of many plants (Sherman et al., 1991) and are presumably oligomers. A forthcoming report (J. Sheng and G. Hind, in preparation) describes the recovery of tp64 from such an oligomer upon complete denaturation in the presence of a thiol. The N-terminal domain and motif IV, with their high cysteine content and high conservation in motifs involving cysteine, are the presumed sites of intermolecular interactions and cross-bridging.

Localization and Uptake of tp64 and PPOs. Recognition of PPOs as lumenal proteins receives support from insight into features of primary sequence that direct protein translocation across the thylakoid membrane (Keegstra et al., 1989; Cline et al., 1993; Robinson et al., 1994). A stretch of 16 hydrophobic amino acids flanked by positively charged or neutral hydrophilic residues in presequences of the 16- and 23-kDa proteins of the OEC directs these to the lumen. A homologous signal occurs in PPOs and tp64 (Figure 4), leaving no doubt that tp64 resides in the lumen. These OEC proteins constitute a subgroup that shares a specific transport pathway into the lumen (Cline et al., 1993; Robinson et al., 1994; Robinson & Klösgen, 1994), which is energetically dependent on the transthylakoid proton gradient and is distinct from the import pathway for plastocyanin and the 33-kDa OEC protein. It is noteworthy that competition among preproteins for the 16- and 23-kDa protein import pathway can be demonstrated *in organello* (Cline et al., 1993). Though not yet demonstrated, PPO transport into the lumen can be assumed to compete with the uptake of OEC 16- and 23-kDa proteins; thus, PPO and PS-II should interact at this level, if no other.

Relation of tp64 to Thylakoid Protein Kinase. The protein described here corresponds in N-terminal sequence to that of a putative LHC-II kinase (Gal et al., 1992) isolated by detergent extraction of spinach thylakoids, ammonium sulfate fractionation, and chromatography on histone−Sepharose. The primary sequence of tp64 clearly is not that of a protein kinase and has high homology with a PPO, as discussed earlier. Furthermore, its transit peptide locates it to the opposite side of the thylakoid membrane relative to the kinase substrates, ATP, and the LHC-II N-terminus (Bennett, 1991), in a compartment with a pH too low (Hind & McCarty, 1973) to support kinase activity in the light.

The inevitable conclusion is that these preparations of LHC-II kinase are contaminated with tp64. Presumably, they will elicit mixed antibody populations: it is noteworthy, for example, that the demonstrated localization of LHC-II kinase in grana margins by immunoelectron microscopy (Gal et al., 1990) is essentially identical to the PPO distribution observed (Lax & Vaughn, 1991) by using antibody to PPO. The LHC-II kinase antibody was further applied (Gal et al., 1990) to test for physical association between LHC-II kinase and the

cytochrome complex, using immunoaffinity column chromatography and immunoprecipitation from an OG extract purified as far as ammonium sulfate precipitation. An association between LHC-II kinase and cytochrome *f* was claimed, since immobilized antibody to the cytochrome also retained a 64-kDa polypeptide. There was comparatively little retention of kinase activity, however. The immobilized anti-kinase antibody of Gal et al. (1990) retained kinase activity and many polypeptides on the affinity column, including a doublet in the 62−66-kDa range that could represent tp64 and the LHC-II kinase. These interesting experiments clearly merit further attention.

Earlier work from our laboratory (Coughlan & Hind, 1986a, 1987a,b; Coughlan et al., 1988) utilized purified kinase fractions of high specific activity. Although some tp64 may have persisted through the purification, there is no reason to believe that this would have affected the characterization of histone kinase activity. The antibody preparation used to study kinase function in the intact membrane (Coughlan & Hind, 1987a) may have possessed anti-tp64 reactivity, but would not have had access to tp64 in the lumen. The observed inhibition of protein phosphorylation presumably resulted from the recognition of kinase, on the stromal aspect of the membrane, by anti-kinase immunoglobulin in the polyclonal antiserum.

We also described (Coughlan & Hind, 1986b, 1987b) the *in membrano* reversible phosphorylation of a 64-kDa protein, presumed to be LHC-II kinase. Its phosphorylation was kinetically comparable to that of LHC-II and the dephosphorylation was faster, in keeping with our interpretation that this 64-kDa protein was undergoing reversible autophosphorylation. Following extraction of membranes with OG and partial purification of an active histone kinase fraction, much slower phosphorylation of a serine residue (Coughlan & Hind, 1987b) in a 64-kDa constituent of the fraction was observed, and the preparation lost its ability to use added LHC-II as substrate. This phosphoprotein is known to be tp64 (Race et al., 1995). Mild trypsinolysis of tp64 produces an unlabeled 42-kDa polypeptide and a 22-kDa phosphoprotein (J. W. Davenport and G. Hind unpublished data), consistent with esterification in the C-terminal domain rather than close to the N-terminus, as previously suggested (Gal et al., 1992). The C-terminal domain contains elements resembling those of LHC-II N-termini (Bennett, 1991) in having hydroxyamino acids imbedded in a basic region; thus, strand β1 (Figure 2) contains KTRTFSLVKK, and at K^{414} there follows KKSRTKED in an exposed coil between β-strands 2 and 3.

Is There a 64-kDa Protein Kinase in the Thylakoid? The work of Gal et al. (1990), in which kinase activity appeared to be bound by an antiserum raised against protein cut from the 64-kDa region of an SDS−PAGE gel, suggests the affirmative, in agreement with our earlier studies (Coughlan & Hind, 1987a,b). The *in membrano* phosphorylation of a 64-kDa polypeptide can also be cited in favor of this notion (Coughlan & Hind, 1986b, 1987b). A possible interpretation that accommodates most of the available literature on the LHC-II kinase is that tp64 and the kinase are sufficiently similar in mass and properties as to cofractionate in a variety of procedures. N-Terminal analysis of a sample containing both proteins would preferentially yield tp64 sequence if the kinase N-terminus is blocked. The missing 30% of total protein recovery in our analysis, reported earlier, could owe its origin to a blocked component.

Conclusion. A 64-kDa protein in spinach thylakoids, tp64, is the mature, latent form of a lumenal PPO, closely resembling PPOs cloned from other dicotyledonous plants. It probably utilizes a lumenal import mechanism shared with the 16- and 23-kDa proteins of the OEC. The mature protein's C-terminal domain is predicted to consist of seven antiparallel β-strands forming a β-barrel comparable to that of Hc and proteins such as Cu/Zn superoxide dismutase and immunoglobulin (Gaykema et al., 1984). The primary sequences of spinach and other PPOs lack any recognizable binding site for ATP or homology with known protein kinases. A kinase with physicochemical properties similar to those of PPO is presumably enriched in detergent extracts of spinach thylakoids and can use spinach PPO (tp64) as an adventitious phosphorylation substrate.

ACKNOWLEDGMENT

We thank Dr. J. J. Dunn for advice on cloning and J. Wysocki for confirmatory protein sequence analysis. Dr. H. L. Race provided valuable comments on the manuscript.

SUPPLEMENTARY MATERIAL AVAILABLE

Figure showing the complete sequence of the tp64 cDNA, as well as its 5′- and 3′-terminal noncoding extensions (a translation using standard three-letter amino acid codes underlies the coding region), and table comparing percentage amino acid identity among the primary sequences represented in Figure 2 (2 pages). Ordering information is given on any current masthead page.

REFERENCES

Adams, V., Griffin, L. D., Gelb, B. D., & McCabe, E. R. B. (1991) *Biochem. Biophys. Res. Commun. 177*, 1101−1106.
Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990) *J. Mol. Biol. 215*, 403−410.
Ambler, R. P. (1991) *Biochem. Soc. Trans. 19*, 517−522.
Ausubel, F. A., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., & Struhl, K., Eds. (1990) *Current Protocols in Molecular Biology*, Greene Publishing & Wiley-Interscience, New York.
Bennett, J. (1991) *Annu. Rev. Plant Physiol. Plant Mol. Biol. 42*, 281−311.
Cary, J. W., Lax, A. R., & Flurkey, W. H. (1992) *Plant Mol. Biol. 20*, 245−253.
Cline, K., Henry, R., Li, C., & Yuan, J. (1993) *EMBO J. 12*, 4105−4114.
Constabel, C. P., Bergey, D. R., & Ryan, C. A. (1995) *Plant Physiol. 92*, 407−411.
Coughlan, S. J., & Hind, G. (1986a) *J. Biol. Chem. 261*, 11378−11385.
Coughlan, S. J., & Hind, G. (1986b) in *Hungarian-USA Binational Symposium on Photosynthesis, Salve Regina College, August 15−18, 1986* (Gibbs, M., Ed.), pp 141−152.
Coughlan, S. J., & Hind, G. (1987a) *J. Biol. Chem. 262*, 8402−8408.
Coughlan, S. J., & Hind, G. (1987b) *Biochemistry 26*, 6515−6521.
Coughlan, S. J., Kieleczawa, J., & Hind, G. (1988) *J. Biol. Chem. 263*, 16631−16636.
Drexel, R., Siegmund, S., Schneider, H.-J., Linzen, B., Gielens, C., Préaux, G., Lontie, R., Kellermann, J., & Lottspeich, F. (1987) *Biol. Chem. Hoppe-Seyler 368*, 617−635.
Eisenberg, D., Schwarz, E., Komaromy, M, & Wall, R. (1984) *J. Mol. Biol. 179*, 125−142.
Farmer, E. E., & Ryan, C. A. (1992) *Plant Cell 4*, 129−134.
Gal, A., Hauska, G., Herrmann, R., & Ohad, I. (1990) *J. Biol. Chem. 265*, 19742−19749.

Gal, A., Herrmann, R. G., Lottspeich, F., & Ohad, I. (1992) *FEBS Lett. 298*, 33−35.
Gaykema, W. P. J., Hol, W. G. J., Vereijken, J. M., Soeter, N. M., Bak, H. J., & Beintema, J. J. (1984) *Nature 309*, 23−29.
Golbeck, J. H., & Cammarata, K. V. (1981) *Plant Physiol. 67*, 977−984.
Halliwell, B. (1978) *Prog. Biophys. Mol. Biol. 33*, 1−54.
Hanks, S. K., Quinn, A. M., & Hunter, T. (1988) *Science 241*, 42−52.
Hayashida, N., Mizoguchi, T., & Shinozaki, K. (1993) *Gene*, 251−255.
Hind, G., & McCarty, R. E. (1973) in *Photophysiology* (Giese, A. C., Ed.), pp Vol. VIII, 113−156, Academic Press, New York.
Hunt, M. D., Eannetta, N. T., Yu, H., Newman, S. M., & Steffens, J. C. (1993) *Plant Mol. Biol. 21*, 59−68.
Jansen, T., Reilander, H., Steppuhn, J., & Herrmann, R. G. (1988) *Curr. Genet. 13*, 517−522.
Karplus, P. A., & Schultz, G. E. (1985) *Naturwissenschaften 72*, 212−213.
Keegstra, K., Olsen, L. J., & Theg, S. M. (1989) *Annu. Rev. Plant Physiol. Plant Mol. Biol. 40*, 471−501.
Kupper, U., Niedermann, D. M., Travaglini, G., & Lerch, K. (1989) *J. Biol. Chem. 264*, 17250−17258.
Lax, A. R., & Vaughn, K. C. (1991) *Plant Physiol. 96*, 26−31.
Looman, A. C., Bodlaender, J., Comstock, L. J., Eaton, D., Jhurani, P., de Boer, H. A., & van Knippenberg, P. H. (1987) *EMBO J. 6*, 2489−2492.
Mayer, A. M. (1987) *Phytochemistry 26*, 11−20.
Mayer, A. M., & Harel, E. (1979) *Phytochemistry 18*, 193−215.
Newman, B. J., & Gray, J. C. (1988) *Plant Mol. Biol. 10*, 511−520.
Newman, S. M., Eannetta, N. T., Yu, H., Prince, J. P., de Vicente, M. C., Tanksley, S. D., & Steffens, J. C. (1993) *Plant Mol. Biol. 21*, 1035−1051.
Race, H. L., Eaton-Rye, J., & Hind, G. (1995) *Phytosynth. Res.* (in press).
Rao, M. J. K., & Argos, P. (1986) *Biochim. Biophys. Acta 869*, 197−214.
Robinson, S. P., & Dry, I. B. (1992) *Plant Physiol. 99*, 317−323.
Robinson, C., & Klösgen, R. B. (1994) *Plant Mol. Biol. 26*, 15−24.
Robinson, C., Cai, D., Hulford, A., Brock, I. W., Michl, D., Hazell, L., & Schmidt, I. (1994) *EMBO J. 13*, 279−285.
Roe, J. L., Rivin, C. J., Sessions, R. A., Feldmann, K. A., & Zambryski, P. C. (1993) *Cell 75*, 939−950.
Rost, B., & Sander, C. (1993) *J. Mol. Biol. 232*, 584−599.
Rost, B., & Sander, C. (1994) *Proteins 19*, 55−72.
Sambrook, J., Fritsch, E., & Maniatis, T. (1989) in *Molecular Cloning: A Laboratory Manual*, 2nd ed., pp 8.60 and 14.20, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Saraste, M., Sibbald, P. R., & Wittinghofer, A. (1990) *Trends Biochem. Sci. 15*, 430−434.
Shahar, T., Hennig, N., Gutfinger, T., Hareven, D., & Lifschitz, E. (1992) *Plant Cell 4*, 135−147.
Sherman, T. D., Vaughn, K. C., & Duke, S. O. (1991) *Phytochemistry 30*, 2499−2506.
Smeekens, S., Weisbeek, P., & Robinson, C. (1990) *Trends Biochem Sci. 15*, 73−76.
Sommer, A., Ne'eman, E., Steffens, J. C., Mayer, A. M., & Harel, E. (1994) *Plant Physiol. 105*, 1301−1311.
Studier, F. W., Rosenberg, A. H., Dunn, J. J., & Dubendorff, J. W. (1990) *Methods Enzymol. 185*, 60−89.
Vaughn, K. C., & Duke, S. O. (1984) *Physiol. Plant. 53*, 421−428.
Volbeda, A., & Hol, W. G. J. (1989) *J. Mol. Biol. 209*, 249−279.

BI9429187