Purification, molecular cloning, and sequence analysis of sucrose-6^F-phosphate phosphohydrolase from plants

John E. Lunn*†, Anthony R. Ashton*, Marshall D. Hatch*, and Hans W. Heldt‡

*Commonwealth Scientific and Industrial Research Organization Plant Industry, GPO Box 1600, Canberra, ACT 2601, Australia; and ‡Albrecht-von-Haller-Institut für Pflanzenwissenschaften der Universität Göttingen, Abteilung für Biochemie der Pflanze, Untere Karspüle 2, 37073 Göttingen, Germany

Sucrose-6^F-phosphate phosphohydrolase (SPP; EC 3.1.3.24) catalyzes the final step in the pathway of sucrose biosynthesis and is the only enzyme of photosynthetic carbon assimilation for which the gene has not been identified. The enzyme was purified to homogeneity from rice (*Oryza sativa* L.) leaves and partially sequenced. The rice leaf enzyme is a dimer with a native molecular mass of 100 kDa and a subunit molecular mass of 50 kDa. The enzyme is highly specific for sucrose 6^F-phosphate with a K_m of 65 μM and a specific activity of 1250 μmol min^{-1} mg^{-1} protein. The activity is dependent on Mg^{2+} with a remarkably low K_a of 8–9 μM and is weakly inhibited by sucrose. Three peptides from cleavage of the purified rice SPP with endoproteinase Lys-C showed similarity to the deduced amino acid sequences of three predicted open reading frames (ORF) in the *Arabidopsis thaliana* genome and one in the genome of the cyanobacterium *Synechocystis* sp. PCC6803, as well as cDNA clones from *Arabidopsis*, maize, and other species in the GenBank database of expressed sequence tags. The putative maize SPP cDNA clone contained an ORF encoding a 420-amino acid polypeptide. Heterologous expression in *Escherichia coli* showed that this cDNA clone encoded a functional SPP enzyme. The 260-amino acid N-terminal catalytic domain of the maize SPP is homologous to the C-terminal region of sucrose-phosphate synthase. A PSI-BLAST search of the GenBank database indicated that the maize SPP is a member of the haloacid dehalogenase hydrolase/phosphatase superfamily.

A phosphatase specific for sucrose 6^F-phosphate (Suc6P) is widely distributed in plant tissues (1, 2) and catalyzes the last step in sucrose synthesis following the formation of Suc6P via sucrose-phosphate synthase (SPS). The reaction catalyzed by sucrose 6^F-phosphate phosphohydrolase (SPP; EC 3.1.3.24) is essentially irreversible and pulls the reaction catalyzed by SPS in the direction of net sucrose synthesis (3). Earlier evidence that the potential activity of SPP *in vivo* is substantially higher than that of SPS (2, 4) suggested that SPP is unlikely to have a regulatory role in sucrose synthesis. However, more recent estimates of SPS activity in at least some of these species range from 1.0 to 3.5 μmol min^{-1} mg^{-1} Chl (5, 6), which are comparable to those recorded previously for SPP. This suggests that SPP could contribute to control of the flux through the pathway of sucrose synthesis. Recent evidence that SPP and SPS could exist as a complex in rice leaves (7, 8) has added further interest to the possible role of SPP in the regulation of sucrose synthesis.

There is only one report of the purification to homogeneity of SPP, from pea shoots (9). This enzyme was reported to be a homodimer with a molecular weight of 120,000 and a specific activity of about 8 μmol min^{-1} mg^{-1} protein. SPP has been partially purified from rice leaves and reported to have a similar molecular weight for the native enzyme (10).

As sucrose is the main product of photosynthesis in most plants, SPP can be regarded as the last enzyme in the pathway of photosynthetic carbon assimilation and is the only enzyme in the pathway for which no gene has been described. During the present studies, we purified SPP from rice leaves to homogeneity

and described some of the physical and kinetic properties of the enzyme. The specific activity was more than 100-fold higher than that reported previously for the pea enzyme (9). Amino acid sequencing of three peptides derived from the purified SPP allowed us to identify a cDNA clone coding for SPP from maize and putative SPP-encoding genes and cDNA clones from other species including *Arabidopsis* and *Synechocystis* sp. PCC6803.

Materials and Methods

Materials. Rice plants (*Oryza sativa* L. cv. Taipei 309) were grown in soil in a naturally illuminated glasshouse maintained between 20 and 30°C. Biochemicals and other special reagents and dyes were obtained from Sigma or Roche Molecular Biochemicals. The dye-based affinity columns were prepared by binding either Procion Red H-E3B (Reactive Red-120) or Cibacron Brilliant Red G-E to Sepharose CL-6B as described (11).

Purification of SPP. Young, fully expanded rice leaves were sliced into 2- to 5-mm segments. One hundred fifty grams of this tissue was blended for 2 min in a Waring blender in 600 ml of buffer A (25 mM Hepes-K$^+$, pH 7.1/8 mM MgCl$_2$/0.5 mM EDTA/8 mM DTT) containing the following protease inhibitors: 1 mM benzamidine, 5 mM ε-aminocaproic acid, 1 mM benzamide, 0.4 mM phenylmethylsulfonyl fluoride, 10 μM antipain, 2 μM leupeptin. For blending only, the medium included 1% (wt/vol) PolyclarAT. All procedures were conducted between 0 and 5°C. After filtering through two layers of Miracloth (Calbiochem), the extract was centrifuged at 10,000 $\times$ *g* for 15 min, and the supernatant was brought to 50% saturation with respect to (NH$_4$)$_2$SO$_4$ by adding an equal volume of a saturated solution of (NH$_4$)$_2$SO$_4$ (saturated at 3°C and adjusted to pH 6.8). After standing for 10 min at 0°C, this mixture was centrifuged at 10,000 $\times$ *g* for 10 min, and the precipitated protein was discarded. Solid (NH$_4$)$_2$SO$_4$ was added to the supernatant at the rate of 7 g per 100 ml to give the equivalent of a 62% saturated solution, and precipitated protein was dissolved in buffer A, including protease inhibitors, in a final volume of 70 ml. Finely powdered polyethylene glycol-8000 was added slowly to this solution with constant stirring to give a final concentration of 29% (wt/vol, 20.3 g to 70 ml). After standing for 10 min at 0°C,

the precipitated protein was pelleted ($25,000 \times g$, 10 min) and discarded. The clear supernatant was warmed to 20°C and the pH quickly adjusted to 5.0 with dilute acetic acid. The mixture was rapidly cooled again to 0°C and stood for 15 min before pelleting the precipitated protein ($25,000 \times g$, 10 min). This protein pellet was suspended in buffer A, including protease inhibitors, to give a final volume of about 6 ml. After adjusting the pH from about 6.0 to about 7.0, the fraction was stored overnight at -80°C.

On thawing, undissolved protein was removed by centrifugation, and the clear supernatant was applied to a 170-ml column of Sephacryl S-300-HR (Amersham Pharmacia) equilibrated with a solution containing 25 mM Hepes-K$^+$, pH 7.1, 8 mM MgCl$_2$, 0.5 mM EDTA, and 40 mM KCl, and 4-ml fractions were collected. The peak fractions emerging from the Sephacryl S-300 column were pooled and concentrated by centrifugation through a Centricon Plus 20 filter (Amicon). The concentrated enzyme (1.8 ml) was applied to an 18-ml Cibacron Brilliant Red G-E column previously equilibrated with 25 mM Hepes-K$^+$, pH 7.2, containing 8 mM MgCl$_2$ and 0.5 mM EDTA and then eluted with the same buffer mixture. Most of the activity appeared in fractions emerging just after the major peak of unbound protein. Similar behavior was observed on a Procion Red H-E3B column. Peak fractions were applied to a 1-ml Mono-Q (Amersham Pharmacia) anion exchange column. Bound protein was eluted with a linear gradient of KCl from 40 mM to 300 mM in 25 mM Hepes-K$^+$, pH 7.1, containing 8 mM MgCl$_2$, 0.5 mM EDTA and 0.03% (vol/vol) Tween-20, at a flow rate of 1 ml min^{-1}. Activity eluted at about 130 mM KCl, coinciding with a major A_{280} peak. Active fractions were concentrated by filter centrifugation (see above), and 200 μl were applied to a 24-ml Superdex 200 (Amersham Pharmacia) gel filtration column equilibrated with the above buffer medium containing 40 mM KCl at a flow rate of 0.5 ml min^{-1}, and 200-μl fractions were collected.

Assay of SPP. SPP activity was determined by following the release of orthophosphate from Suc6P. Reactions containing 1.25 mM Suc6P and 8 mM MgCl$_2$ in 25 mM Hepes-K$^+$, pH 7.0, in a total volume of 300 μl were incubated at 30°C and stopped by adding 30 μl of 2 M trichloroacetic acid. Orthophosphate was measured in samples using the ascorbic acid–ammonium molybdate reagent described in ref. 12. Fructose 6-phosphate (Fru6P) was used to test the specificity of the activity observed with Suc6P as substrate.

Determination of Protein. Protein was determined by the dye-binding assay described in ref. 13 with bovine γ-globulin as a standard. For the peak fractions off MonoQ and Superdex 200, protein was determined from the associated A_{280} peak using a value of $A_{280} = 1.47$ for 1 mg ml^{-1} determined from the amino acid composition of maize SPP ($\varepsilon_{280} = 69,340$ M^{-1} cm^{-1}). Protein in the peak fractions from Superdex 200 was also determined by densitometry of the Coomassie-stained band on SDS-polyacrylamide gels by comparison with the marker proteins run on the same gel.

SDS/PAGE. Proteins were separated on 12% polyacrylamide gels as described in ref. 14, and the bands were located by staining with Coomassie brilliant blue R-250.

Amino Acid Sequencing. The 50-kDa protein in samples of purified rice SPP was excised from a 12% SDS-polyacrylamide gel after staining with Coomassie blue. The protein (approximately 5 μg) was eluted from the gel and cleaved by incubation with endo-proteinase Lys-C from *Achromobacter lyticus* (Wako Pure Chemical, Osaka). Three hydrophobic peptides resulting from Lys-C cleavage were purified and sequenced at the Biomolecular Resource Facility of the Australian National University.

Table 1. Purification of SPP from rice leaves

Fraction	Activity, μmol min^{-1}	Specific activity, μmol min^{-1} mg^{-1}	Purification, -fold
Leaf crude extract	408	0.05	—
50–62% sat (NH$_4$)$_2$SO$_4$	270	0.08	1.6
29% PEG, pH 5	188	1.03	21
S300 (#12–15)	119	1.41	28
Brilliant Red G-E	12*	29	580
Mono-Q	8.5*	520	10,000
Superdex 200	0.96*	1250†	25,000

*Only part of the available activity from the preceding step was processed. Total recoveries of activity from the Brilliant Red G-E, MonoQ, and Superdex 200 columns were 90, 94, and 95%, respectively.

†Protein estimated from the intensity of Coomassie staining of bands on SDS/PAGE gels.

Expression of Maize SPP in *Escherichia coli*. Standard cloning procedures were carried out as described in ref. 15. The putative maize SPP coding region (GenBank accession no. AF283564) was modified by the PCR to introduce an *Nco*I site at the ATG translation initiation codon and an *Xho*I site 19 bp downstream of the TAG termination codon and was ligated between the corresponding sites of the bacterial expression plasmid pTYB4 (New England Biolabs). The recombinant plasmid was introduced into *E. coli* strain ER2566 (New England Biolabs). A stationary-phase culture of *E. coli* ER2566 (pTYB4/*ZmSPP1*) grown in LB medium was diluted 250-fold into 50 ml of LB containing 100 μg ml^{-1} ampicillin in a 250-ml flask and incubated with shaking (250 rpm) at 37°C until the OD$_{600}$ reached 0.5. Protein expression was induced by the addition of isopropyl β-D-thiogalactopyranoside to a final concentration of 0.3 mM. After incubation at 30°C for 3 h, the cells were harvested by centrifugation at $4000 \times g$ for 10 min (4°C). The cells were washed by resuspension in 50 ml of ice-cold water and recentrifugation. The cells were resuspended in 1 ml of buffer A containing 1 mM phenylmethylsulfonyl fluoride, 10 μM antipain and 2 μM leupeptin, with 0.5 ml of 0.1-mm-diameter glass beads and agitated for 1 min at maximum power with a Mini Bead-beater-8 cell disrupter (Biospec Products, Bartlesville, OK). The cell extract was centrifuged for 1 min and assayed for SPP activity as described above.

Results

Purification and Specific Activity. The rice leaf SPP was purified about 25,000-fold by the procedure described in *Materials and Methods* and summarized in Table 1. The 20-fold purification obtained by treatment on the Brilliant Red G-E dye column was due in part to the binding of most of the nonactive protein to the column under the conditions used and in part to the fact that SPP activity emerged just after the major peak of unbound protein. SPP activity appeared with one of the major protein peaks emerging from the MonoQ column and was coincident with the major protein peak (100 kDa) emerging from the Superdex 200 column. The peak fractions from Superdex 200 showed a single protein band on SDS-polyacrylamide gels with a molecular weight of about 50,000 (Fig. 1). The specific activity of SPP in the peak fractions from the Superdex 200 column was about 1,250 μmol min^{-1} mg^{-1} protein, based on estimates of protein concentration from the color intensity of the 50-kDa band on a Coomassie blue-stained SDS-polyacrylamide gel. Using the very low A_{280} values for the peak fractions from the Superdex 200 column, the specific activity was about one-half of this value.

Properties. The purified enzyme was totally inactive with Fru6P as substrate. A Hanes plot (s/*v* vs. s) for varying Suc6P showed that the enzyme has Michaelis–Menten kinetics with a K_m for
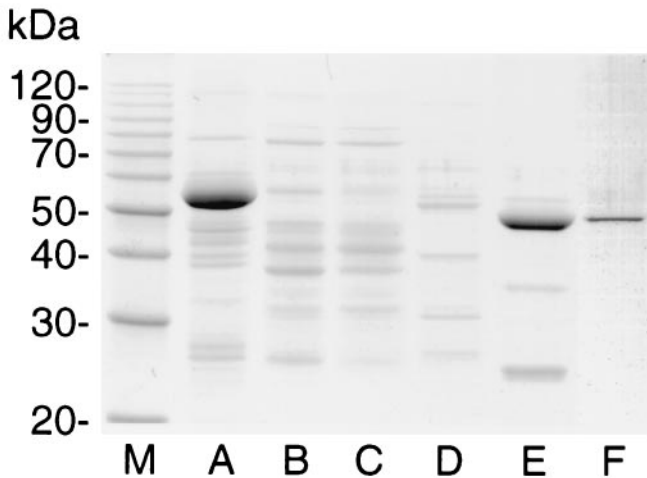
Fig. 1. SDS/PAGE of samples from each stage in the purification of SPP from rice leaves. Amounts of protein loaded were as follows: lane A, crude extract (15 μg); lane B, 29% polyethylene glycol-8000 precipitate (10 μg); lane C, Sephacryl S300 pooled fractions 12–15 (10 μg); lane D, Brilliant Red-Sepharose 6B fraction 12 (5 μg); lane E, Mono Q fraction 14 (10 μg); lane F, Superdex 200 fraction 30 (0.5 μg).

Suc6P of 65 μM. Although leaf SPP has been shown to have an absolute requirement for Mg^{2+}, it has been reported that Mg^{2+} is difficult to remove (9, 10). The latter studies also reported that Mg^{2+} was required at concentrations approaching 1 mM or more for maximum activity. By contrast, we found that Mg^{2+} was readily removed by treatment of the highly purified enzyme on a Sephadex G-25 column (Table 2) and, based on the Hanes plot for varying Mg^{2+}, that SPP has an unusually low K_a for Mg^{2+} of about 8–9 μM. The activity with 1 mM $MgCl_2$ was almost completely abolished by a 13-fold excess of EDTA (Table 2). It was estimated that the low activity seen with the Sephadex G-25-treated enzyme assayed without added Mg^{2+} (Table 2) was equivalent to an Mg^{2+} concentration of less than 1 μM. It should be noted that under our conditions, the enzyme was unstable after processing on Sephadex G-25 to remove Mg^{2+}. This loss of activity was prevented by including EDTA in the equilibrating buffer at a concentration as low as 5 μM or by adding EDTA immediately after the enzyme emerged from the G-25 column. The enzyme was inactivated if Zn^{2+} was added in small excess over EDTA. The requirement for Mg^{2+} can be partially replaced by Mn^{2+}. For instance, at 0.1 mM the rate with Mn^{2+} was about 37% of that with Mg^{2+}. However, in the presence of Mn^{2+}, an equal concentration of Mg^{2+} did not activate SPP further.

Table 2. Mg^{2+} requirement for SPP activity

Additions	Activity, μmol min^{-1} ml^{-1}
8 mM $MgCl_2$*	2.95
None	0.34
13 mM EDTA	0.01
1 mM $MgCl_2$	2.96
1 mM $MgCl_2$ + 13 mM EDTA	0.05
0.1 mM $MgCl_2$	2.80
0.01 mM $MgCl_2$	1.61
0.002 mM $MgCl_2$	0.75

Purified SPP was desalted by passage through a Sephadex G-25 column equilibrated with 25 mM Hepes-K$^+$, pH 7.0, and assayed for activity with 1.2 mM Suc6P and various concentrations of $MgCl_2$ or EDTA as shown.
*Enzyme before desalting on Sephadex G-25.

Sucrose was a weak competitive inhibitor of SPP, with a K_i of 200–240 mM. For instance, at a concentration of 160 mM, a likely upper limit for the concentration of sucrose in photosynthetic cells (4), inhibition varied between about 15% at 550 μM Suc6P to 23% at 165 μM and 38% at 42 μM Suc6P.

Amino Acid Sequence. The N terminus of the purified SPP appeared to be blocked. Therefore, the protein was cleaved using endoproteinase Lys-C, which cleaves at the C-terminal side of Lys residues, and three of the resulting peptides were sequenced: peptide 1, WDRNIVVEETANVSELK; peptide 2, LGPN-VSPRDVDFPYVK; peptide 3, FYVLYEK.

A search of the GenBank database with the sequence of peptide 1 (including an initial Lys residue), using the TBLASTN algorithm, showed significant matches with three predicted open reading frames (ORF) in the genome of *Arabidopsis thaliana* in the following BAC clones: (*i*) accession no. AL132957 (Expect value [E] = 341, 55% identity, 10 of 18); (*ii*) AC007017 (E = 446, 55% identity, 10 of 18); and (*iii*) AL132972 (E = 260, 58% identity, 10 of 17), and a partial cDNA clone from *Brassica oleracea*: X97679 (E = 39, 70% identity, 12 of 17), all of unknown function. The deduced amino acid sequences of the *Arabidopsis* ORFs and *Brassica* cDNA clone also show regions with strong similarity to peptides 2 and 3 (Fig. 2). The deduced amino acid sequences of the putative *Arabidopsis* SPPs were used to search the database, and we identified a hypothetical protein coding region in the genome of the cyanobacterium *Synechocystis* sp. PCC6803 (D90914; E = 2 × 10^{-25}, 37% identity, 68 of 183) (Fig. 2). Interestingly, the next best match was the *Synechocystis* gene coding for SPS (D64006; E = 2 × 10^{-11}, 26% identity, 53 of 199) (16). The deduced amino acid sequences of the *Arabidopsis* ORFs were also used to search the GenBank database of expressed sequence tags (EST) and showed significant matches with EST clones from a number of species including maize, *Arabidopsis*, rice, sorghum, and *Medicago truncatula* (Table 3).

Expression of the Maize *SPP1* Gene in *E. coli*. Sequencing of the maize EST clone showed that it contained a complete ORF coding for a polypeptide with a molecular mass of 47,213 Da, which is close to the subunit molecular mass of the rice leaf SPP (50 kDa). The putative maize SPP coding region was cloned into the bacterial expression plasmid pTYB4, and the resulting recombinant plasmid pTYB4/ZmSPP1 was expressed in *E. coli*. With extracts from cells containing the pTYB4 plasmid with no insert, Suc6P was hydrolyzed at a rate of only about 3 nmol min^{-1} mg^{-1} protein. This rate was substantially less than that for Fru6P hydrolysis and only slightly inhibited by EDTA (Table 4). In contrast, extracts from cells containing pTYB4/*ZmSPP1* showed SPP activity of 370 nmol min^{-1} mg^{-1} protein, which was 85% inhibited by 20 mM EDTA. The 15% residual SPP activity was attributable to residual free Mg^{2+} (about 3 μM free Mg^{2+} with 20 mM EDTA) and the high affinity of SPP for Mg^{2+} (see above). The extracts from cells carrying either pTYB4 or pTYB4/*ZmSPP1* showed similarly low phosphatase activity with Fru6P as the substrate (Table 4).

Comparison of the Maize *SPP1* Gene with Other SPP-Like Sequences. The EST clones from *Arabidopsis* and *M. truncatula* were also sequenced. The *Arabidopsis* EST shows 100% identity with the ORF in one gene (AC007017) but is missing the first four bases of the coding region. Three other ESTs from *Arabidopsis* show perfect or near perfect sequence identity with AC007017. The *M. truncatula* cDNA clone contains a complete ORF. The maize *SPP1* cDNA shows between 62 and 96% identity with the other *SPP1*-like nucleotide sequences from higher plants, listed in Table 3. The maize SPP protein shows between 52 and 95%

```
                <Motif I>                                         Motif II
                *** *    *                                        * * * * * *
ZmSPP    MDKLSGSARLMIVSDLDHTMVD------HHDEENLSLLRFGALWESVYCEDSLLVFSTGRSPTLYKEL   62
AtSPP    MERLTSPPRLMIVSDLDHTMVD------HHDPENLLLRFNSLWEHAYRHDSLLVFSTGRSPTLYKEL   62
SynSPP   ......MRQLLLISDLDNIWVG--------DQQALEHLQ-EYLGDR--RGNFYLAYATGRSYHSAREL   51
SynSPS   KRRRTLYYNGALVTSLDQNLLGALQGGLPGDRQTLDEL-LEVLYQH--RKNVGFCIATGRRLDSVLKI  523
ZmSPS    KYPLLRRRRRLFVIAVDCYQDD--GRASKKMLQVIQEVFRAVRSDSQMFKISGFTLSTAMPLSETLQL  835

                                                <--Peptide 1---->
                                                WDRNIVVEETANVSELK
ZmSPP    RKEKPMLTPDI--TIMSVGTEITYGEA----------MVPDDGWEEYLNNKWDRNIVVAETVSFSELK-  118
AtSPP    RKEKPLLTPDI--TIMSVGTEITYGNS----------MVPDHGWVEALNNKWDLGIVKQEASNFPELK-  118
SynSPP   QKQVGLMEPDI--WLTAVGSEIYHPEG----------L-DQHWADYLSEHWQRDILQAIADGFEALK-  105
SynSPS   LREYRIPQPDM--LITSMGTEIYSSPD----------LIPDQSWRNHIDYLWNRNAIVRILGELPGLA-  579
ZmSPS    LQLGKIPATDFDALICGSGSEVYYPGTANCMDAEGKLRPDQDYLMHISHRWSHDGARQTIAKLMGAQD  903

                                                     <------------
                                                     * *      *
ZmSPP    ------LQPETEQRPHKVSFFVDKKNAQEVIKSVAERLDKCGLDAKIIYSGGQD-LDILPQGAGKGQA  179
AtSPP    ------LQAETEQRPHKVSFYVEKSKAQEVTKELSQRFLKRGLDVKIIYSGGMD-LDILPQGAGKGQA  179
SynSPP   ------PQSPLEQNPWKISYHLDPQACPTVIDQLTEMLKETGIPVQVIFSSGKD-VDLLPQRSNKGNA  166
SynSPS   ------LQPKEELSAYKISYFYDAAIAPN-LEEIRQLLHKGEQTVNTIISFGQF-LDILPIRASKCYA  639
ZmSPS    GSGDAVEQDVASSNAHCVAFLIKDPQKVKTVDEMRERLRMRGLRCHIMYCRNSTRLQVVPLLASRSQA  971

         ------Motif III------------------------->
         *          * * * * *      *    *      * * *
ZmSPP    IAYLLEKLSSCGKPPNNTLVCGDSGN-DAELFSIPGVHGVMVSNAQEELLQWYTENAKDNPK--IIHS  244
AtSPP    IAYLLKKLKTEGKLPVNTLACGDSGN-DAELFSIPDVYGVMVSNAQEELLKWHAENAKDNPK--VIHA  244
SynSPP   TQYLQQHLAME---PSQTLVCGDSGN-DIGLFE-TSARGVIVRNAQPELLHWYDQ--WGDSR--HYRA  225
SynSPS   VRWLSQQWNI---PLEHVFTAGGSGA-DEDMMR-GNTLSVVVANRHHEEL----SNLGEIEP--IYFS  696
ZmSPS    LRYLSVRWGVSVG--NMYLITEHGDTDLEEMLSGLHKTVIVRGVTEKGSEALVRSPGSYKRDDVVPS  1037

                    <---Peptide 2--->               Peptide 3
                    LGPNVSPRDV-DFPYVK               FYVLYEK
ZmSPP    NERCAA-----GIIQAIGHFKLGPNISPRDL-QFPYAKEASFKPTDAVVKFYVLYEKWRRAEVPKSDSVI  308
AtSPP    KERCAG-----GIIQAIGHFKLGPNLSPRDVSDFLEIKVENVNPGHEVVKFFLFYERWRRGEVENSEAYT  309
SynSPP   QSSHAG---AILEAIAHFDFLS*................................................  244
SynSPS   EKRYAA---GILDGLAHYRFFELLDPV*...........................................  720
ZmSPS    ETPLAAYTTGELKADEIMRALKQVSKTSSGM*.......................................1068


ZmSPP    KYFKNITHANGVTIHPAGLELSLHASIDALGSCYGDKQGRKYRAWVDRLFITQTGSDSWVGRFDLWES  376
AtSPP    ASLKASVHPGGVFVHPSGTEKSLRDTIDELRKYHGDKQGKKFRVWADQVLATDTTPGTWIVKLDKWEQ  377


ZmSPP    EGDVRVCSLSSLALILKAESPEGFVLTHIQKTWLNGYSSGVEQAFKL* 423
AtSPP    DGDERRCCTTTVKFTSK--EGEGLVWEHVQQTWSKETMVKDDSSWII* 422
```
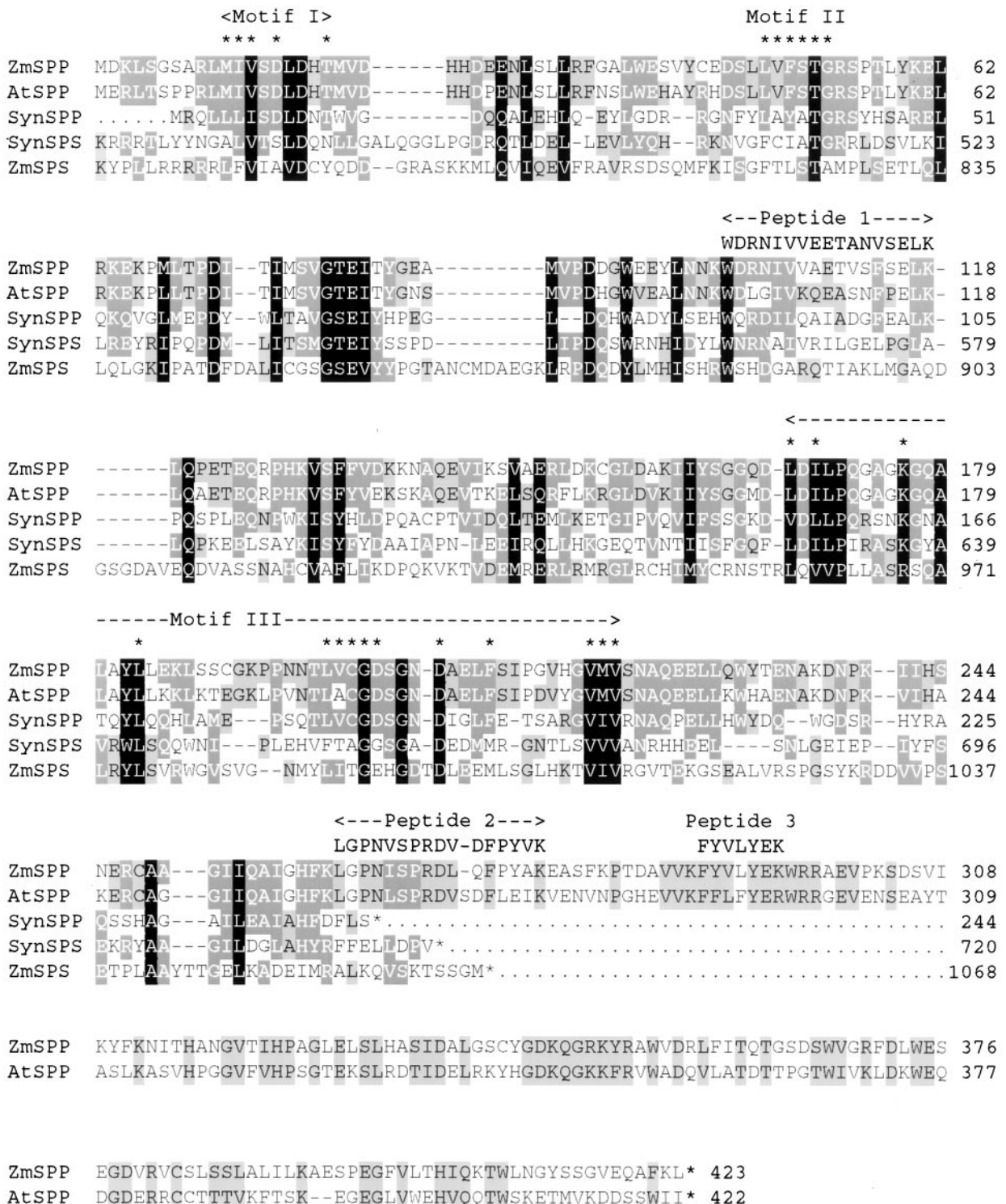
Fig. 2. Alignment of the deduced amino acid sequence of the maize SPP with SPP-like sequences and the carboxyl terminus of SPS from other species. The GenBank accession numbers of the sequences are as follows: maize (*Zea mays*) SPP (ZmSPP), AF283564; *A. thaliana* SPP (AtSPP), AF283563; *Synechocystis* sp. PCC6803 SPP (SynSPP), AF300455; *Synechocystis* sp. PCC6803 SPS (SynSPS), D64006 (residues 459–720); and maize SPS (ZmSPS), M97550 (residues 770-1068). Sequences were aligned using the PILEUP program of the GCG Wisconsin sequence analysis package. Residues conserved in five, three, and two out of the five sequences are highlighted in black, dark gray, and light gray, respectively. Asterisks indicate residues that are homologous to conserved residues in motifs I, II, and III of the HAD-type phosphatases. The sequences of peptides 1, 2, and 3 from purified rice leaf SPP are also shown in alignment with the SPP sequences.

identity with the corresponding deduced amino acid sequences, showing highest identity with a partial EST from sorghum seedlings (Table 3). The maize *SPP1* shows much lower identity with the *SPP1*-like sequence from the cyanobacterium *Synechocystis* sp. PCC6803 (48 and 37% at the nucleotide and amino acid levels, respectively). The predicted molecular weight of the

Table 3. Comparison of maize *SPP1* with *SPP1*-like sequences from other species

Species	Accession no.	Nucleotide identity,* %	Amino acid identity,† %	Molecular weight
Maize	AF283564‡	(100)	(100)	47,213
Arabidopsis	AC007017§	68	65	47,855
Arabidopsis	AL132957§	62	52	46,657
Arabidopsis	AL132972§	65	58	48,711
Arabidopsis	AF283565‡,¶	68	65	47,855
Sorghum	AW565306‡,¶	96	95	—
Rice	AA752973‡,¶	83	71	—
Medicago	AF283566‡	69	64	47,451
Cauliflower	X97679‡,¶	64	53	—
Synechocystis	AF300455§	48	37	27,762

Sequences were obtained by searching the GenBank nonredundant nucleotide and EST databases using the TBLASTN algorithm. The similarity of sequences to the maize *SPP1* sequence was determined using the BESTFIT program of the Wisconsin sequence analysis package (Genetics Computer Group, Madison, WI).
*Comparison with maize *SPP1* nucleotide sequence.
†Comparison with deduced amino acid sequence of maize SPP.
‡cDNA clone.
§Genomic sequence.
¶Partial clone.

enzymes from higher plants ranges from 46,657 to 47,855 (Table 3). In contrast, the putative *Synechocystis* SPP has a predicted molecular weight of only 27,762 (Table 3).

Computer programs such as PSI-BLAST that use multiple sequence comparisons to match a query sequence to a database can detect many more remote homologues than programs such as BLAST, which use pairwise comparisons (17). At the third iteration, a PSI-BLAST search of the database with the maize SPP sequence indicated that the N-terminal region (~260 amino acids) of the enzyme is homologous to members of a superfamily of proteins related to the L-2-halocid dehalogenase (HAD) from *Pseudomonas* sp. (18), with expect (*E*) values ranging from about 10^{-4} to 10^{-22}. The HAD superfamily includes a number of other phosphatases, including trehalose-6-phosphatase (TPP), phosphoglycolate phosphatase, and phosphoserine phosphatase, as well as P-type ATPases, phosphomannomutase, and many hypothetical proteins (18). The overall similarity between members of the HAD superfamily is generally low; for example, the maize SPP and *Arabidopsis* TPP show only 16% overall identity. However, three motifs associated with the catalytic domain are highly conserved (18) in these proteins (Fig. 2). A further interesting and unexpected observation was that the C-terminal region of SPS sequences from higher plants and *Synechocystis* also shows homology with this N-terminal region of SPP (Fig. 2).

Table 4. Expression of maize SPP in *E. coli*

Substrate	Phosphatase activity, nmol min⁻¹ mg⁻¹ protein	
	Control	SPP
Suc6P	3	370
Suc6P + EDTA	2	53
Fru6P	14	10

The putative maize SPP cDNA clone was expressed in *E. coli* strain ER2566. Phosphatase activity with either 1.25 mM Suc6P or Fru6P as the substrate was assayed in extracts from cells carrying the empty expression vector pTYB4 (control) or pTYB4/*ZmSPP1*. SPP activity was also measured in the presence and absence of 20 mM EDTA.

Discussion

Properties. The high specific activity and 25,000-fold purification we recorded for SPP indicate that SPP is a low abundance enzyme in leaves. We have no explanation for the 100-fold lower specific activity reported for this enzyme purified from pea shoots (9). Our results support the conclusion that the native SPP is a dimer (9). The molecular mass of the native rice SPP was estimated to be 100 kDa by gel filtration on Superdex 200 with a subunit molecular mass of 50 kDa, determined by SDS/PAGE. A higher value of about 120 kDa for the native enzyme was found by gel filtration on Sephacryl S300. These values are close to the molecular masses of 120 kDa and 55 kDa reported for the pea enzyme (9) and about 120 kDa for native SPP from rice (10).

In many species there is a shift in assimilate partitioning away from sucrose as carbohydrate accumulates in the leaves during the day (19), but the control of this process is poorly understood. In maize leaves, accumulation of sucrose alone does not result in feedback inhibition of sucrose synthesis or a change in carbon partitioning (6). It has been reported that SPP activity in crude extracts from a number of species is partially inhibited by sucrose (2), and we found that sucrose is a weak competitive inhibitor of the rice leaf SPP. To our knowledge, the concentration of Suc6P in rice leaves has not been measured, but estimates from other species range from 30–35 μM in spinach leaves to 230 μM in germinating peas (3, 20). If the concentration of Suc6P in rice leaves were within this range, then SPP would be partially inhibited by physiological concentrations of sucrose. Further work will be needed to determine whether the degree of inhibition of SPP by sucrose changes during the day, but studies on spinach leaves suggest that accumulation of Suc6P is unlikely (20). Even if SPP were inhibited enough to cause accumulation of Suc6P, it is unclear if this would have any impact on the rate of sucrose synthesis, as SPS is not significantly inhibited by Suc6P (20, 21).

Molar Ratios of SPP and SPS *in Vivo*. Recently, evidence was reported that SPP and SPS may exist as a complex in rice leaves (7, 8). If this proves to be correct, then it is likely that these enzymes occur at a molar ratio of about unity. It was therefore of interest to calculate the apparent molar ratio of SPS to SPP in leaves. This was done for rice using our values of maximum extractable activities, which averaged 2.3 μmol min⁻¹ mg⁻¹ Chl for SPP and 2.0 μmol min⁻¹ mg⁻¹ Chl for SPS, a molecular weight of 470,000 (tetramer) and specific activity of 150 μmol min⁻¹ mg⁻¹ protein for SPS (22), and, for SPP from the present studies, a molecular weight of 94,400 and specific activities indicated below.

On the basis of Coomassie staining on SDS/PAGE gels (see above), we estimated a specific activity for purified SPP of about 1,250 μmol min⁻¹ mg⁻¹ protein. With this value, the molar ratio of native tetrameric SPS to dimeric SPP was calculated to be 1.45. A higher protein value, and hence a lower specific activity giving an SPS to SPP ratio of 0.61, was obtained from the area of the A_{280} peak traced during chromatography on Superdex-200. However, this latter estimate may be less reliable because of the very low absolute absorbances recorded. Allowing for the potential errors in these calculations, the values indicate that the true molar ratio is likely to be close to unity, which is consistent with the proposal that SPS and SPP may exist as a complex *in vivo* (7, 8).

Features of the SPP Gene and Protein. The sequencing of peptides derived from the purified rice SPP allowed us to identify genes from *Arabidopsis* and the cyanobacterium *Synechocystis* sp. PCC6803, as well as cDNA clones from a number of plant species including maize and *Arabidopsis* that might code for SPP. The putative maize SPP cDNA clone was expressed in *E. coli* and

shown to encode a functional SPP enzyme. This is the first report of the identification and cloning of an *SPP* gene. It should be noted that the maize *SPP1* gene shows no similarity to bacterial genes coding for the invertase-like enzyme sucrose-6^G-phosphate hydrolase (E.C. 3.2.1.26) (23), which is involved in the phosph*enol*pyruvate-dependent sucrose phosphotransferase system used by some bacteria to take up sucrose (24).

There were some early reports that SPP is located in the vacuole (25, 26), but more recent studies indicated that the enzyme is cytosolic (27, 28). The predicted N-terminal sequence of the maize SPP does not have the characteristics of an endoplasmic reticulum/vacuolar signal peptide or a mitochondrial or plastidic transit peptide (29, 30), supporting a cytosolic location for the enzyme.

The N-terminal region (260 amino acids) of the maize SPP shows the three motifs typical of members of the HAD superfamily of proteins (18) (Fig. 2). Motif I (residues 11–19) includes the sequence D*XDXT, in which D* (Asp-15) is homologous to the Asp residue that functions as the acyl phosphate catalytic intermediate in other HAD-type phosphatases (31). Other conserved residues including Thr-52 in motif II (residues 48–53) and Lys-176, Asp-202, and Asp-206 in motif III (residues 167–220) are predicted to occur close to the enzyme active site (18, 31, 32). The deduced amino acid sequence of the *SPP1*-like gene from *Synechocystis* sp. PCC6803 is much shorter than that of the maize SPP and only shows similarity with this N-terminal region of the higher plant enzyme (Fig. 2). Because the *Synechocystis* SPP is the same size as other HAD phosphatases, it is likely that this N-terminal region comprises the catalytic core of the higher plant SPP.

Fig. 2 also shows an alignment of the C-terminal regions of SPS from maize and *Synechocystis* sp. PCC6803 with SPP. About three-quarters of the SPS polypeptide starting from the N terminus resembles the closely related glycosyltransferase enzyme, sucrose synthase (22). However, the remaining C-terminal region of the maize SPS shows 35% similarity to the N-terminal region of the maize SPP that is likely to be the catalytic core of the enzyme. The C-terminal region of the *Synechocystis* SPS shows 42% identity to the putative *Synechocystis* SPP. The C-terminal domains of both the maize and *Synechocystis* SPS contain residues that are homologous to Thr-52, Lys-176 (Arg in maize SPS), and Asp-206 in SPP, but not Asp-15 or Asp-202 (Fig.

2). The lack of these two Asp residues, including the critical Asp that functions as the acyl phosphate intermediate in HAD-family phosphatases, suggests that the SPP-like, C-terminal domain of SPS does not have SPP activity. This agrees with the observation that highly purified SPS shows no SPP activity (21).

The two-domain structure of SPS, an N-terminal glycosyltransferase and a C-terminal phosphatase domain, bears a remarkable resemblance to the enzymes responsible for the biosynthesis of the disaccharide trehalose, TPP, and trehalose-phosphate synthase (TPS). However, in yeast, it is TPP (encoded by the *TPS2* gene) that contains an N-terminal glycosyltransferase domain that is homologous to TPS (*TPS1*), fused to a catalytically active C-terminal TPP domain (33). The TPS-like domain in the TPP enzyme is probably not catalytically active. TPP, TPS, and two catalytically inactive, regulatory subunits (*TPS3* and *TSL1*), which are homologous to the entire TPP protein, together form a large TPS–TPP complex (33).

The conservation of the structural organization of these evolutionarily remote but catalytically similar proteins suggests that there is a significant advantage to this arrangement. The possibility of intrapolypeptide metabolite channeling can be excluded because the multidomain proteins of sucrose (SPS but not SPP) and trehalose (TPP but not TPS) synthesis do not have both catalytic activities. However, the two fused domains might allow the formation of heteropolymeric complexes and metabolite channeling within such complexes. As discussed earlier, there is some evidence that rice SPS and SPP can form a loose complex that is more active than the separate enzymes (7, 8). Possibly the SPP-like domain of SPS mediates such an interaction with SPP.

1. Hawker, J. S. & Hatch, M. D. (1966) *Biochem. J.* **99,** 102–107.
2. Hawker, J. S. & Smith, G. M. (1984) *Phytochemistry* **23,** 245–249.
3. Lunn, J. E. & ap Rees, T. (1990) *Biochem. J.* **267,** 739–743.
4. Stitt, M., Huber, S. & Kerr, P. (1987) in *Biochemistry of Plants, Vol 10: Photosynthesis*, eds. Hatch, M. D. & Boardman, N. K. (Academic, San Diego), pp. 327–409.
5. Lunn, J. E. & Furbank, R. T. (1997) *Planta* **202,** 106–111.
6. Lunn, J. E. & Hatch, M. D. (1997) *Aust. J. Plant Physiol.* **24,** 1–8.
7. Salerno, G. L., Echeverria, E. & Pontis, H. G. (1996) *Cell. Mol. Biol.* **42,** 665–672.
8. Echeverria, E., Salvucci, M. E., Gonzalez, P., Paris, G. & Salerno, G. (1997) *Plant Physiol.* **115,** 223–227.
9. Whitaker, D. P. (1984) *Phytochemistry* **23,** 2429–2430.
10. Echeverria, E. & Salerno, G. (1994) *Plant Sci.* **96,** 15–19.
11. Ashton, A. R. & Polya, G. M. (1978) *Biochem. J.* **175,** 501–506.
12. Ames, B. N. (1966) *Methods Enzymol.* **8,** 115–119.
13. Bradford, M. (1976) *Anal. Biochem.* **72,** 248–254.
14. Laemmli, U. K. (1970) *Nature (London)* **227,** 680–685.
15. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Plainview, NY), 2nd Ed.
16. Lunn, J. E., Price, G. D. & Furbank, R. T. (1999) *Plant Mol. Biol.* **40,** 297–305.
17. Park, J., Karplus, K., Barrett, C., Hughey, R., Haussler, D., Hubbard, T. & Chothia, C. (1998) *J. Mol. Biol.* **284,** 1201–1210.

18. Aravind, L., Galperin, M. Y. & Koonin, E. V. (1998) *Trends Biochem. Sci.* **23,** 127–129.
19. Lunn, J. E. & Hatch, M. D. (1995) *Planta* **197,** 385–391.
20. Krause, K. P. & Stitt, M. (1992) *Phytochemistry* **31,** 1143–1146.
21. Lunn, J. E. & ap Rees, T. (1990) *Phytochemistry* **29,** 1057–1063.
22. Huber, S. C. & Huber, J. (1996) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **47,** 431–444.
23. Rauch, P. J. G. & de Vos, W. M. (1992) *Gene* **121,** 55–61.
24. Thompson, J. & Chassy, B. M. (1981) *J. Bacteriol.* **147,** 543–551.
25. Hawker, J. S., Smith, G. M., Phillips, H. & Wiskich, J. T. (1987) *Plant Physiol.* **84,** 1281–1285.
26. Echeverria, E. & Salvucci, M. E. (1991) *Plant Physiol.* **96,** 1014–1017.
27. Echeverria, E. & Salerno, G. (1993) *Physiol. Plant.* **88,** 434–438.
28. Echeverria, E. (1995) *Physiol. Plant.* **95,** 559–562.
29. Nielsen, H., Engelbrecht, J., Brunak, S. & von Heijne, G. (1997) *Protein Eng.* **10,** 1–6.
30. Emanuelsson, O., Nielson, H. & von Heijne, G. (1999) *Protein Sci.* **8,** 978–984.
31. Collet, J.-F., Stroobant, V., Pirard, M., Delpierre, G. & Van Schaftingen, E. (1998) *J. Biol. Chem.* **273,** 14107–14112.
32. Toyoshima, C., Nakasako, M., Nomura, H. & Ogawa, H. (2000) *Nature (London)* **405,** 647–655.
33. Bell, W., Sun, W., Hohmann, S., Wera, S., Reinders, A., De Vigilio, C., Wiemken, A. & Thevelein, J. M. (1998) *J. Biol. Chem.* **273,** 33311–33319.