

# A highly specific L-galactose-1-phosphate phosphatase on the path to ascorbate biosynthesis

William A. Laing<sup>\*†</sup>, Sean Bulley<sup>\*</sup>, Michele Wright<sup>\*</sup>, Janine Cooney<sup>‡</sup>, Dwayne Jensen<sup>‡</sup>, Di Barraclough<sup>\*</sup>, and Elspeth MacRae<sup>\*</sup>

<sup>\*</sup>Gene Technologies Sector, Horticultural and Food Research Institute of New Zealand, Ltd., PB 92169, Auckland, New Zealand; and <sup>‡</sup>Food Sector, Horticultural and Food Research Institute of New Zealand, Ltd., PB 3123, Hamilton, New Zealand

Communicated by William L. Ogren, Hilton Head Island, SC, October 11, 2004 (received for review July 27, 2004)

Ascorbate is a critical compound in plants and animals. Humans are unable to synthesize ascorbate, and their main source of this essential vitamin are plants. However, the pathway of synthesis in plants is yet to be established, and several unknown enzymes are only postulated to exist. We describe a specific L-galactose-1-phosphate (L-gal-1-P) phosphatase that we partially purified from young kiwifruit (*Actinidia deliciosa*) berries. The enzyme had a native molecular mass of  $\approx 65$  kDa, was completely dependent on  $Mg^{2+}$  for activity and was very specific in its ability to hydrolyze L-gal-1-P. The activity had a pH optimum of 7.0, a  $K_M(L-gal-1-P)$  of 20–40  $\mu M$  and a  $K_a(Mg^{2+})$  of 0.2 mM. The activity was inhibited by  $Mg^{2+}$  at concentrations  $>2$  mM. The enzyme from *Arabidopsis thaliana* shoots showed similar properties to the kiwifruit enzyme. The *Arabidopsis thaliana* enzyme preparation was digested with trypsin, and proteins present were identified by using liquid chromatography–MS. One of 24 proteins present in our preparation was an *Arabidopsis thaliana* protein, At3g02870, annotated *myo*-inositol-1-phosphate phosphatase in GenBank, that matched the characteristics of the purified L-gal-1-phosphate phosphatase. We then expressed a kiwifruit homologue of this gene in *Escherichia coli* and found that it showed 14-fold higher maximum velocity for L-gal-1-P than *myo*-inositol-1-P. The expressed enzyme showed very similar properties to the enzyme purified from kiwifruit and *Arabidopsis*, except that its  $K_M(L-gal-1-P)$  and  $K_a(Mg^{2+})$  were higher in the expressed enzyme. The data are discussed in terms of the pathway to ascorbate biosynthesis in plants

ascorbic acid | biosynthesis | vitamin C | *myo*-inositol-1-phosphate

Ascorbic acid is an essential compound in plants and animals, with a range of functions including scavenging reactive oxygen species (1) and maintaining human health (2). Humans cannot synthesize their own ascorbic acid (or vitamin C) (2) and are dependent on external sources of this important vitamin, mainly from plants. However, the pathway of ascorbic acid biosynthesis in plants has not been firmly established, and many enzymes have yet to be discovered (3, 4).

Three pathways of biosynthesis of ascorbic acid have been proposed in plants, one through L-galactose (L-gal) (5), another from *myo*-inositol (6–8), and a third through galacturonic acid (9). The L-gal pathway proceeds through L-gal to galactono-1,4-lactone and, thence, to ascorbate (5). The ascorbate can then be broken down to end products including oxalate (10). From mutant and biochemical studies (11–13), it is likely that the L-gal is derived from GDP-D-mannose through GDP-L-galactose to L-gal-1-phosphate (L-gal-1-P), which has been proposed to be converted to L-gal through the action of two yet-to-be-identified enzymes including an unknown specific phosphatase (14–18). In the *myo*-inositol pathway, inositol-1-P is derived from glucose-6-P. It is then converted to *myo*-inositol by a phosphatase (19–21) and, thence, to glucuronic acid by inositol oxygenase.

The enzyme activities and genes for several specific plant phosphatases have recently been identified, such as phosphoglycolate phosphatase (22), phosphoserine phosphatase (23), *myo*-inositol monophosphatase (21), sucrose-6-phosphate phos-

phatase (24), and trehalose-6-phosphate phosphatase (25), and enzyme activity for other specific phosphatases has been identified but not yet cloned; e.g., sorbitol-6-phosphate phosphatase (26) and 2-carboxy-D-arabinitol 1-phosphate phosphatase (27). There is also a wide range of information on nonspecific acid phosphatases (28–30), which would be expected to dephosphorylate L-gal-1-P among other substrates, but there appears to be no published information on a plant phosphatase that would catalyze the dephosphorylation of L-gal-1-P in a specific manner.

It is very likely that kiwifruit makes ascorbate through the galactose pathway as both leaves and fruit have been shown to express the gene for the committed enzyme L-gal dehydrogenase (31), and fruit shows high levels of this enzyme (W.A.L., unpublished data). However, the preceding phosphatase step in ascorbic acid biosynthesis has not been identified or described. In this work, we describe the partial purification of a specific L-gal-1-phosphate phosphatase (L-gal-1-Pase), which would act in the L-gal pathway of ascorbic acid biosynthesis characterize the enzyme, and identify the gene. The gene has been previously identified as a *myo*-inositol-1-phosphate phosphatase (*myo*-inositol-1-Pase), raising possibilities of linkage between two different related pathways of ascorbate biosynthesis.

## Materials and Methods

**Materials.** Kiwifruit [*Actinidia deliciosa* (A. Chev.) C. F. Liang and A.R. Ferguson *deliciosa* cv. “Hayward”] berries were picked at Te Puke, New Zealand. For biochemical studies, 1-cm-diameter fruit were used and picked 4 weeks after anthesis. *Arabidopsis thaliana* (Columbia-1) plants were grown for  $\approx 5$  weeks at 20–25°C and 200  $\mu mol \cdot m^{-2} \cdot sec^{-1}$  photosynthetically active radiation. Tissues were stored frozen at  $-80^\circ C$  until use.

Most chemicals were obtained from Sigma, except the protease inhibitor tablets (Roche Applied Sciences) and L-gal-1-P (Glycoteam, Hamburg, Germany). The L-gal-1-P appeared as a single peak by HPLC and had a mass of 260 Da. Column chromatography media were obtained from Amersham Biosciences. Bradford’s reagent was provided by Bio-Rad.

**Enzyme Assays.** Two assays were used to measure the activity of the L-gal-1-Pase enzyme. The first assay measured the phosphate formed by the hydrolysis of L-gal-1-P by using a microplate-based colorimetric phosphate assay (32) and was used during purification and to characterize the enzyme. The standard conditions for this assay were 0.1 M Bis-Tris propane (pH 7.0), 2 mM  $MgCl_2$ , and 0.2 mM L-gal-1-P in 100  $\mu l$ . Between 1 and 20  $\mu l$  of extract was added to start the reaction, incubated for 15 to 30 min at 30°C, and terminated by the addition of 100  $\mu l$  of 3%

Abbreviations: L-gal, L-galactose; L-gal-1-P, L-gal-1-phosphate; L-gal-1-Pase, L-gal-1-P phosphatase; *myo*-inositol-1-P, *myo*-inositol-1-phosphate; *myo*-inositol-1-Pase, *myo*-inositol-1-P phosphatase.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AY787585 and AY787586).

<sup>†</sup>To whom correspondence should be addressed. E-mail: wlaing@hortresearch.co.nz.

© 2004 by The National Academy of Sciences of the USA

ascorbate and 0.5% (both wt/vol) ammonium molybdate in 0.5 M HCl. Background assays were obtained by adding the enzyme after the reaction was terminated. The second assay, used during purification and to verify products, coupled the production of L-gal to L-gal dehydrogenase (5, 33) and measured the reduction of NAD<sup>+</sup>. The basic assay in 100  $\mu$ l was the same as the first assay but additionally contained 0.2 mM NAD<sup>+</sup> as well as an excess of L-gal dehydrogenase either partially purified from kiwifruit berries (31) or with the cloned enzyme expressed in *Escherichia coli* (S.B. and W.A.L., unpublished data). Time courses were followed up to 30 min at 20°C by using either a Victor fluorescence plate reader (Wallac, Gaithersburg, MD) to measure NADH fluorescence (excitation, 340 nm; emission, 460 nm) or absorbance spectroscopy at 340 nm.

Enzyme kinetic analyses for  $K_m$  and  $V_{max}$  measurements, [Mg<sup>2+</sup>] response, alternative substrates, and pH optimum were replicated three times within the experiment in a standard manner. Data are presented as means and SEs as appropriate. Experiments were repeated at least twice, and a representative experiment is presented.

L-Gal dehydrogenase activity was assayed in 0.1 M Tris (pH 8.0)/0.25 mM NAD/0.25 mM L-gal in 100  $\mu$ l at 20°C by using the Victor fluorimeter to measure NADH fluorescence.

**L-Gal-1-Pase Purification.** Kiwifruit berries or *Arabidopsis thaliana* leaves were ground to a fine powder under liquid nitrogen, and  $\approx$ 70 g of frozen tissue was extracted in 100 ml (kiwifruit) or 200 ml (*Arabidopsis thaliana*) of 0.18 M Mops buffer adjusted to pH 7.0 and containing 1.0% Triton X-100, 4% polyvinylpyrrolidone, 2.5 mM DTT, 10% (vol/vol) glycerol, 10  $\mu$ M E64 (kiwifruit), and one proteinase inhibitor tablet per 50 ml. This extract was then filtered through a 20- $\mu$ m nylon mesh and centrifuged (20 min at 20,000  $\times$  g), and the supernatant was decanted. This supernatant was applied to a 5  $\times$  60-cm G75 Superdex column (GE Healthcare) run at 5 ml $\cdot$ min<sup>-1</sup> equilibrated with 0.1 M Tris (pH 8.0)/2 mM EDTA/5 mM 2-mercaptoethanol/10% glycerol. Active fractions were pooled, diluted with water and glycerol to 50 mM Tris-Cl/10% glycerol, and applied to two 5-ml HiTrap Q FF columns (GE Healthcare) in series equilibrated with 20 mM Tris-Cl (pH 8.0)/2 mM EDTA/5 mM 2-mercaptoethanol/10% glycerol. The HiTrap Q column was eluted by using a gradient of 0–0.5 M NaCl in 20 mM Tris-Cl (pH 8.0) containing EDTA, 2-mercaptoethanol, and glycerol.

Active fractions were combined, and ammonium sulfate was added to 1 M before application of the sample to a 5-ml HiTrap Butyl FF column. Proteins bound to the column were eluted with a reverse gradient of ammonium sulfate from 1 M to 0 M. Active fractions were then concentrated to 1 ml by using a 10,000 molecular weight cutoff centrifugal concentrator and applied to a 40  $\times$  1.6-cm G75 Superdex column equilibrated with the same buffer as the large G75 Superdex column used earlier. Active fractions were combined and used to characterize the enzyme.

SDS/PAGE was performed by using the Bio-Rad Miniprotean II system with a 4% (wt/vol) acrylamide stacking gel and a 10% (wt/vol) acrylamide separating gel by using the Tricine-based buffer (34). A broad molecular mass range of Bio-Rad prestained marker proteins were used to determine molecular mass. Gels were stained with SYPRO Ruby (Molecular Probes) followed by Coomassie blue by using standard techniques.

**Liquid Chromatography–MS Analysis of Peptides.** A partially purified L-gal-1-Pase preparation was digested with trypsin [0.5  $\mu$ g of modified sequencing grade trypsin (Roche) in 0.2 M NH<sub>4</sub>HCO<sub>3</sub> and 0.5 mM CaCl<sub>2</sub>], and gel slices containing proteins from an SDS/PAGE separation of the same preparation were subjected to in-gel digestion (35). Digested peptides were analyzed by liquid chromatography–MS by using an LCQ Deca ion-trap mass spectrometer fitted with a nanospray electrospray

ionization interface (ThermoQuest, Finnigan, San Jose, CA) and coupled to a Surveyor HPLC (ThermoQuest, Finnigan). Five microliters of the digested protein sample was injected onto a reversed phase column (PepMap C18, 75  $\mu$ m  $\times$  15 cm, 3  $\mu$ m; LC Packings, San Francisco). Tryptic peptides were separated at a flow rate of 400 nl/min with a linear gradient from 2% to 80% B (acetonitrile plus 0.1% formic acid) over 50 min. Solvent A was 0.1% aqueous formic acid.

MS/MS data were analyzed by using TURBOSEQUENT (36, 37). Spectra were searched against an *Arabidopsis thaliana* protein database. The criteria used for a positive peptide identification for a doubly charged peptide were a correlation factor >2.0, a  $\delta$  cross-correlation factor >0.1, and a high preliminary scoring. For triply charged peptides, the correlation factor was set at 2.5. All matched peptides were confirmed by visual examination of the spectra.

**Gene Cloning.** The full-length cDNA was obtained from Hort-Research's (Auckland) kiwifruit EST resource cloned into Bluescript SK(–). The 880-bp fragment was excised by using *Bam*HI and *Hind*III, cloned in frame into pET30c(+) (Novagen), and transformed into *E. coli* DH5 $\alpha$  cells. A correct clone (verified by PCR amplification with pET primers and by restriction enzyme digestion) was transformed into *E. coli* BL21 RIL cells (Novagen). The construct contains an N-terminal His<sub>6</sub>-tag. A single colony was grown in 50 ml of Terrific broth to an OD<sub>600</sub> of 0.6 at 37°C, then protein expression was induced by addition of isopropyl  $\beta$ -D-thiogalactoside to 0.3 mM. The induced culture was grown for 20 h at 20°C. Cells were harvested by centrifugation at 8,000  $\times$  g at 4°C for 5 min, resuspended in the His column binding buffer (Novagen) containing one protease inhibitor tablet per 50 ml, and broken by a high-pressure emulsifier (Emulsiflex, Avestin, Ottawa). The lysate was centrifuged at 12,000  $\times$  g for 30 min at 4°C, and the filtered (0.45  $\mu$ m) supernatant was purified through a 5-ml HiTrap Nickel column. Fractions containing protein were dialyzed at 4°C against three changes of 10 mM Bis-Tris Propane (pH 7.1), 1 mM EDTA, 5 mM 2-mercaptoethanol, and 30 mM NaCl.

## Results

**Purification of a Specific L-Gal-1-Pase from Kiwifruit.** A large peak of kiwifruit L-gal-1-Pase activity was eluted from the large G75 Superdex column after the void volume (Fig. 5, which is published as supporting information on the PNAS web site). Further analysis of this peak showed that the earlier eluting fractions consisted of a neutral pH, Mg<sup>2+</sup>, and L-gal dependent phosphatase activity, whereas activity in the later eluting fractions had a more acid pH optimum and was independent of MgCl<sub>2</sub> (data not shown). Consequently, only the earlier fractions were applied to the anion exchange column, where a clear separation of the remaining acid phosphatase and the Mg<sup>2+</sup>-dependent phosphatase was achieved (Fig. 5). The first major peak of phosphatase activity was unaffected by MgCl<sub>2</sub>, whereas the second peak was stimulated 9-fold by MgCl<sub>2</sub> (data not shown). The first peak, which had an optimum below pH 6 (data not shown), was most likely a nonspecific acid phosphatase with activity against L-gal-1-P.

Further purification of the second HiTrap Q peak was gained through fractionation on a hydrophobic interaction column and through rechromatography on a G75 Superdex column (Fig. 6, which is published as supporting information on the PNAS web site). The peak activity on this latter column eluted the fraction after BSA, suggesting a native molecular mass close to 65,000 Da.

Further purification (including chromatofocusing on a MonoP column and native gel electrophoresis) resulted in increases in protein purity but catastrophic losses in activity. The enzyme would not bind to a HiTrap SP column at pH 5.2. Consequently,

**Table 1. Substrate specificity for the partially purified L-gal-1-Pase activity**

| Substrate                  | Partially purified kiwifruit enzyme |       | Partially purified <i>Arabidopsis</i> enzyme |       | <i>E. coli</i> expressed kiwifruit enzyme EST 233909 |       |
|----------------------------|-------------------------------------|-------|--|-------|--|-------|
|                            | nmol/min                            | %     | nmol/min                                     | %     | nmol/min   | %     |
| L-Gal-1-P                  | 0.220                               | 100.0 | 0.163  | 100.0 | 1.61*  | 100.0 |
| D-Gal-1-P                  | 0.011                               | 4.6   |  |       | 0.066  | 4.1   |
| D-Mannose-1-P              | 0.008                               | 3.3   |  |       | 0.059  | 3.7   |
| Dihydroxyacetone phosphate | -0.001                              | -0.3  |  |       |  |       |
| Fructose-1, 6-bisP         | -0.001                              | -0.4  |  |       |  |       |
| Fructose-1-P               | -0.003                              | -1.3  |  |       |  |       |
| Fructose-6-P               | 0.000                               | -0.2  |  |       |  |       |
| Fructose-2, 6-bisP         | 0.020                               | 9.8   |  |       |  |       |
| Guanosine diphosphate      | 0.007                               | 3.6   |  |       |  |       |
| Glucose-6-P                | 0.000                               | -0.1  |  |       | -0.001   | 0.0   |
| Glucose-1-P                | 0.016                               | 6.7   |  |       |  |       |
| Glucose-1-6-bisP           | 0.002                               | 1.0   |  |       |  |       |
| Guanosine triphosphate     | -0.004                              | -2.1  |  |       |  |       |
| Mannose-6-P                | -0.001                              | -0.3  |  |       |  |       |
| p-Nitrophenyl phosphate    | 0.009                               | 4.2   | 0.001  | 0.4   | -0.014   | -0.9  |
| Phosphoglyceric acid       | 0.001                               | 0.3   |  |       | 0.000  | 0.0   |
| Phosphoglycolate           | 0.008                               | 3.6   | 0.009  | 5.9   |  |       |
| Ribose-5-P                 | 0.003                               | 1.5   |  |       | 0.014  | 0.9   |
| Ribulose-1, 5-bisP         | -0.003                              | -1.2  |  |       |  |       |
| Sucrose-6-P                | 0.006                               | 2.7   | 0.003  | 1.9   | -0.001   | -0.1  |
| Trehalose-6-P              | -0.001                              | -0.6  | 0.003  | 1.7   | -0.003   | -0.2  |
| Myo-inositol-1-P           |                                     |       |  |       | 0.353  | 21.9* |

Reactions were run in 100 mM Bis-Tris Propane buffer (pH 7.0), with 2 mM MgCl<sub>2</sub>, and 0.5 mM phosphorylated substrate for 30 min (kiwifruit and *Arabidopsis*) or 11 min (*E. coli*) at 30°C. SEs were ≈2% of the mean for these assays. Empty positions indicate that the substrate was not assayed with that enzyme.

\*The test for alternative substrates used excess *E. coli* expressed enzyme in this assay to maximize sensitivity for alternative substrates, and the control rate was beyond the linear portion of the activity versus the amount of enzyme curve (data not shown). The true level of L-Gal-1-Pase activity was about three times this value. In separate measurements, we established that under saturating substrate concentration and 2 mM MgCl<sub>2</sub> at pH 7.0, myo-inositol-1-P was hydrolysed at 7.1% the rate of L-gal-1-P.

characterization of the enzyme was carried out on combined fractions from the small G75 Superdex column. At this stage of the purification, at least 10 proteins with molecular masses between 30,000 and 100,000 Da were apparent by SDS/PAGE (data not shown).

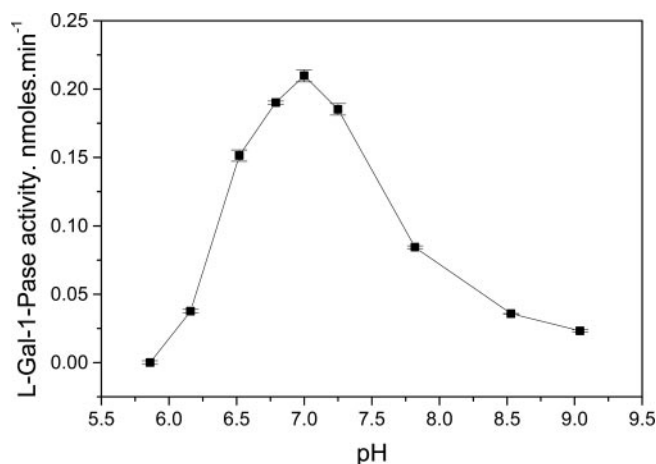
**Characterization of the Kiwifruit Phosphatase Activity.** L-Gal-1-Pase activity was linear with the amount of enzyme until at least 10 nmol of phosphate had been generated in the assay, which represented ≈20% of the substrate present. The activity was also linear with time for at least 30 min.

**Substrate Specificity of the Kiwifruit Enzyme.** We verified that the product of the reaction was L-gal by coupling the phosphatase activity to the reduction of NAD by using semipurified L-gal dehydrogenase from kiwifruit. This enzyme has been shown to be active with L-gal as the substrate with some activity with L-sorbose (*Arabidopsis*), L-fucose (*Arabidopsis*, but not kiwifruit), and L-gulose but no activity against a wide range of other sugars (31, 33). In all steps of purification, the coupled reaction paralleled the direct phosphate assay (data not shown).

The ability of the L-gal-1-Pase to hydrolyze a range of phosphorylated compounds was measured by using the standard phosphate assay. The results are the means of two experiments, and no substrate other than L-gal-1-P was significantly and reproducibly hydrolyzed by the partially purified enzyme (Table 1). Several substrates (fructose-2,6-bisP, ribulose-1,5-bisP, and dihydroxyacetone phosphate) were partially hydrolyzed (≈20% of the available substrate) by the acid stop mix and showed high backgrounds, explaining an apparent ability of the enzyme to

hydrolyze fructose-2,6-bisP. However, this substrate showed a near-zero rate in other assays. Where there was slight activity, the enzyme showed a preference for phosphate in the C1 position.

The pH optimum showed a well defined maximum at pH 7 with a falloff to half of the maximum rate within ±0.6 pH of the maximum (Fig. 1).



**Fig. 1.** pH response of L-gal-1-Pase. Assays were run in 0.2 M Bis-Tris propane buffer (pH as specified on the graph), 2 mM MgCl<sub>2</sub>, and 0.5 mM L-gal-1-P for 30 min at 30°C. Data are the mean ± SE of three replicates. Separate experiments showed very similar responses.



*myo*-inositol-1-Pase that fell in group 1 (Fig. 4), in *E. coli*. This protein showed some *myo*-inositol-1-Pase activity but had  $\approx 14$  times more L-gal-1-Pase activity and very little activity against a wide range of other phosphorylated compounds (Table 1). The expressed enzyme had a maximum specific activity against L-gal-1-P of  $16.4 \pm 1.0 \mu\text{mol } \pi\text{-generated per min}^{-1}$  per mg of protein (as calculated from the L-gal-1-P response curve), whereas an extract from *E. coli* expressing the empty vector after His-column purification showed no phosphatase activity against L-gal-1-P, *myo*-inositol-1-P, or other substrates tested. Coupled assays with *E. coli* expressed kiwifruit L-gal dehydrogenase, which uses L-gal to reduce  $\text{NAD}^+$ , showed that the products of the reaction were L-gal as well as inorganic phosphate (data not shown).

Kinetic analysis of the cloned and expressed kiwifruit enzyme showed it had a  $K_m(\text{L-gal-1-P})$  of  $0.15 \pm 0.02 \text{ mM}$ , a  $K_m(\text{myo-inositol-1-P})$  of  $0.33 \pm 0.02 \text{ mM}$ , and a  $K_a(\text{Mg}^{2+})$  of  $0.47 \pm 0.16 \text{ mM}$  with a  $K_i(\text{Mg}^{2+})$  of  $13.4 \pm 5.9 \text{ mM}$  (data not shown). The high SEs for Mg inhibition reflect the much lower inhibition by  $\text{Mg}^{2+}$  seen with this enzyme compared with the enzymes extracted from plants. The enzyme showed a pH optimum of 7.0 (data not shown), very similar to the *Arabidopsis* and kiwifruit enzymes. LiCl inhibited the L-gal-1-Pase activity, although with a  $K_i$  of  $3.7 \pm 0.4 \text{ mM}$  (data not shown), which is considerably higher than that observed with the expressed tomato *myo*-inositol-1-Pase (21). The enzyme also appeared to be a dimer by gel filtration. The predicted subunit molecular mass of the expressed enzyme, including the His-tag, is  $\approx 34,800 \text{ Da}$ . On a G75 Sephadex column, the enzyme eluted (data not shown) just before a BSA standard (68,000 Da), as would be expected for a dimer ( $\approx 70,000 \text{ Da}$ ).

## Discussion

Ascorbic acid is a critical chemical in plants, for both plant and human health, and understanding the path of biosynthesis is important in achieving the significant goal of increasing plant ascorbic acid levels. In this work, by discovering a phosphatase highly specific to the L-gal pathway of ascorbic acid biosynthesis, we further establish the significance of this pathway in higher plants, as well as open new possibilities to manipulate ascorbic acid levels in plants. In addition, we raise the possibility that two pathways of ascorbic acid biosynthesis, namely the L-gal and *myo*-inositol pathways, may be linked through the action of a common phosphatase enzyme.

We have partially purified enzymes from kiwifruit berries and *Arabidopsis thaliana* shoots that show significant specificity in hydrolyzing L-gal-1-P to L-gal and inorganic phosphate as shown by direct assays of both products. The activities show a low  $K_m$  for both L-gal-1-P and  $\text{Mg}^{2+}$  and a pH optimum of 7.0. These properties make them very viable candidates to be the phosphatase in the pathway to ascorbate biosynthesis because there appears to be no other use for L-gal in plants other than ascorbate biosynthesis (15, 18).

**L-Gal-1-Pase Gene.** We have also identified a kiwifruit cDNA as coding the protein L-gal-1-Pase. This clone shows significant homology to an *Arabidopsis thaliana* gene (*At3g02870*), identified in GenBank as a putative *myo*-inositol-1-Pase protein, in a partially purified protein preparation from *Arabidopsis* shoots. The *E. coli* expressed enzyme had very similar properties and highly specific ability to hydrolyze L-gal-1-P as found in both the *Arabidopsis* and kiwifruit partially purified enzymes. These properties of the partially purified enzymes included a native mass close to 65,000 Da (presumably a homodimer), a low pI (excluded by SP column at pH 5.2), an absolute  $\text{Mg}^{2+}$  dependency with inhibition at higher  $\text{MgCl}_2$  concentrations, a very strong preference for L-gal-1-P as substrate, and a pH optimum close to 7.0. The main differences between the recombinantly

expressed and recombinantly extracted enzymes were that the recombinant enzyme had a higher  $K_a$  for  $\text{Mg}^{2+}$  and a higher  $K_m$  for L-gal-1-P than the extracted enzymes, but this may reflect the bacterial folding environment or the presence of the His-tag and other sequence contributed by the pET vector as well as another 14 residues from the 5' UTR of the EST in front of the N-terminal of the expressed enzyme. The *Arabidopsis* enzyme does not have a targeting sequence nor does the Kiwifruit enzyme. The next enzyme in the pathway, L-gal dehydrogenase, has no targeting sequence either (31).

**Putative *myo*-Inositol-1-Pase Genes in Apple and Kiwifruit.** We observed that apple, kiwifruit, and *Arabidopsis thaliana* putative *myo*-inositol-1-Pase sequences partition into three clusters (groups 1–3 in Fig. 4). We have analyzed cDNA sequences that belong to the three clusters. Within apple, there were 23 ESTs belonging to the *myo*-inositol phosphatase family in total, of which 18 belong to group 1, the clade with demonstrated gal-1-Pase activity. Of these group-1 ESTs, 13 ESTs were from libraries made from fruit tissue, whereas five ESTs came from shoot and xylem libraries. Because HortResearch has sequenced  $\approx 160,000$  apple ESTs, these 23 ESTs represent a frequency of 0.014%. The group-2 and group-3 apple ESTs came from fruit (3 ESTs) and shoot (5 ESTs) tissues. In the case of kiwifruit, there were also 23 ESTs identified (a frequency of 0.017%). Group 1 ESTs came from fruit (15) and buds (4). Group-2 and group-3 ESTs also came from fruit (3) and buds (1). Thus, overall, group-1 (L-gal-1-Pase-like) ESTs were the most commonly found ESTs (total 37 of 46), predominantly from fruit (28) and, to a lesser extent, shoot (9) tissues. This result is compatible with ascorbic acid biosynthesis occurring via the L-gal pathway in fruit tissue.

**The Inositol-1-Monophosphatase Gene Family.** This inositol-1-monophosphatase family is a widely distributed family (39, 40) and includes the *E. coli* phosphomonoesterases such as SuhB (41). SuhB preferentially hydrolyses *myo*-inositol-1-P but can also hydrolyze glycerophosphate, *myo*-inositol-2-P, and adenosine-2'-monophosphate. The *myo*-inositol family also includes members of the *N*-amidino-scylo-inosamine-4-phosphate phosphatase and fructose bisphosphatase families, and enzymes with the ability to cleave D-gal-1-P and fructose-1-6-bisphosphate are described in refs. 42–44.

In plants, two genomes have now been fully sequenced: rice and the model plant *Arabidopsis thaliana*. Both have one putative *myo*-inositol phosphatase gene belonging to group 1 in our phylogenetic tree and one *myo*-inositol phosphatase-like gene in each of the other groups. Tomato (21, 45) and potato have three putative *myo*-inositol phosphatase genes in group 1 (Fig. 4), the same cluster in which we found an EST with L-gal-1-Pase activity (EST 233909, Fig. 4). There is a GenBank Solanaceae sequence that belong to each of the other two groups.

Discovering *myo*-inositol-1-Pase activity by the expressed tomato genes was based on screening a phage library for 5-bromo-4-chloro-3-indoyl phosphate phosphatase activity at neutral pH (21). The one positive clone (LeIMP) was then used to screen a young fruit cDNA library resulting in isolation of the two other clones (21). The three clones show greater than  $\approx 80\%$  similarity, whereas the three *Arabidopsis thaliana* clones only show 20–40% similarity. This finding may explain why putative *myo*-inositol-1-Pases belonging to groups 2 and 3 were not cloned by these authors from tomato. The activity reported for the tomato clones used a very sensitive assay (21), and as we show, our expressed L-gal-1-Pase can hydrolyze *myo*-inositol-1-P at  $\approx 7\%$  of the rate of L-gal-1-P hydrolysis. Gillespy *et al.* (21) did not test specificity of activity against other substrates and did not characterize the enzyme's properties except to suggest that it was a dimer,  $\text{Mg}^{2+}$ -dependent, and  $\text{Li}^+$ -sensitive (21). Curiously, whereas

Gillespy *et al.* isolated *LeIMP1* from a young fruit library, this gene is more highly expressed in mature fruit and flowers and increases in expression with exposure to light (21). Such an increase in expression with exposure to light could suggest that this gene may also have L-gal-1-Pase activity due to activation of ascorbate synthesis. Although we currently do not know the activity of the proteins encoded by genes found in the other two groups, it may be that the group-1 genes are actually all L-gal-1-Pases and the real *myo*-inositol-1-Pase has yet to be identified.

**The Function of a Phosphatase in the Biosynthesis of Ascorbate.** Two pathways of ascorbic acid biosynthesis involve a phosphatase step: the L-gal pathway (5) and the *myo*-inositol pathway (8, 21). *Myo*-inositol metabolism is also of significance in the widespread phosphatidylinositol signaling pathway (39). It is of note that the one enzyme we describe shows activity against both of these substrates, but it is much more effective against L-gal-1-P than *myo*-inositol-1-P. It is possible that both pathways involve the one gene and enzyme, or more likely, that the L-gal pathway has seconded the *myo*-inositol-1-Pase into a new role through gene

duplication. This finding would imply that the group-2 or group-3 members of the *myo*-inositol-1-Pase family are the true phosphatases for *myo*-inositol-1-P, and that the group-1 phosphatases are the seconded enzyme.

In conclusion, we have partially purified and characterized a specific L-gal-1-Pase that we propose is significant in the biosynthesis of ascorbic acid, further supporting the importance of the L-gal route to ascorbate biosynthesis, and we have identified the gene that encodes this protein and discussed its relationship with other genes.

**Note Added in Proof.** Since submission of this paper, we have shown that the *Arabidopsis thaliana* group 1 clone At3g02870 expressed in *E. coli* shows  $\approx 12$  times as much L-gal-1-Pase activity as *myo*-inositol-1-Pase activity (data not shown).

We thank Dr. Andrew Gleave (HortResearch) for providing sequencing facilities; Dr. Ross Crowhurst (HortResearch), manager of the EST sequence database and bioinformatics suite; and Dr. Lesley Beuning (HortResearch), who oversaw tissue collection and library construction for kiwifruit and apple ESTs. This work was supported by Foundation for Science and Technology Grant CO6X0006.

1. Apel, K. & Hirt, H. (2004) *Annu. Rev. Plant Biol.* **55**, 373–399.
2. Naidu, K. A. (2003) *Nutr. J.* **2**, 7.
3. Smirnov, N., Conklin, P. L. & Loewus, F. A. (2001) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **52**, 437–467.
4. Smirnov, N., Running, J. A. & Gatzek, S. (2004) in *Vitamin C: Its Function and Biochemistry in Animals and Plants*, eds. Asard, H., May, J. M. & Smirnov, N. (Bios Scientific, London), pp. 7–29.
5. Wheeler, G. L., Jones, M. A. & Smirnov, N. (1998) *Nature* **393**, 365–369.
6. Loewus, F. A. & Kelly, S. (1961) *Arch. Biochem. Biophys.* **95**, 483–493.
7. Isherwood, F. A. & Mapson, L. W. (1956) *Biochem. J.* **64**, 13–22.
8. Lorence, A., Chevone, B. I., Mendes, P. & Nessler, C. L. (2004) *Plant Physiol.* **134**, 1200–1205.
9. Agius, F., Gonzalez-Lamothe, R., Caballero, J. L., Munoz-Blanco, J., Botella, M. A. & Valpuesta, V. (2003) *Nat. Biotechnol.* **21**, 177–181.
10. Banhegyi, G. & Loewus, F. A. (2004) in *Vitamin C: Function and Biochemistry in Animals and Plants*, eds. Asard, H., May, J. M. & Smirnov, N. (Bios Scientific, London), pp. 31–48.
11. Hancock, R. D., McRae, D., Haupt, S. & Viola, R. (2003) *BMC Plant Biol.* **3**, 7.
12. Conklin, P. L., Norris, S. R., Wheeler, G. L., Williams, E. H., Smirnov, N. & Last, R. L. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 4198–4203.
13. Wolucka, B. A., Persiau, G., Van Doorselaere, J., Davey, M. W., Demol, H., Vandekerckhove, J., Van Montagu, M., Zabeau, M. & Boerjan, W. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 14843–14848.
14. Conklin, P. L. (2001) *Plant Cell Environ.* **24**, 383–394.
15. Davey, M. W., Van Montagu, M., Inze, D., Sanmartin, M., Kanellis, A., Smirnov, N., Benzie, I. J. J., Strain, J. J., Favell, D. & Fletcher, J. (2000) *J. Sci. Food Agric.* **80**, 825–860.
16. Smirnov, N. (2000) *Curr. Opin. Plant Biol.* **3**, 229–235.
17. Smirnov, N. & Wheeler, G. L. (2000) *Crit. Rev. Biochem. Mol. Biol.* **35**, 291–314.
18. Smirnov, N. (2001) *Vitam. Horm.* **61**, 241–266.
19. Gumber, S. C., Loewus, M. W. & Loewus, F. A. (1984) *Plant Physiol.* **76**, 40–44.
20. Loewus, M. W. & Loewus, F. A. (1982) *Plant Physiol.* **70**, 765–770.
21. Gillaspay, G. E., Keddle, J. S., Oda, K. & Gruijsem, W. (1995) *Plant Cell* **7**, 2175–2185.
22. Mamedov, T. G., Suzuki, K., Miura, K., Kucho Ki, K. & Fukuzawa, H. (2001) *J. Biol. Chem.* **276**, 45573–45579.
23. Ho, C. L., Noji, M. & Saito, K. (1999) *J. Biol. Chem.* **274**, 11007–11012.
24. Lunn, J. E., Ashton, A. R., Hatch, M. D. & Heldt, H. W. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 12914–12919.
25. Vogel, G., Aeschbacher, R. A., Muller, J., Boller, T. & Wiemken, A. (1998) *Plant J.* **13**, 673–683.
26. Zhou, R., Cheng, L. & Wayne, R. (2003) *Plant Sci.* **165**, 227–232.
27. Kingston-Smith, A. H., Major, I., Parry, M. A. & Keys, A. J. (1992) *Biochem. J.* **287**, 821–825.
28. Turner, W. L. & Plaxton, W. C. (2001) *Planta* **214**, 243–249.
29. Schenk, G., Guddat, L. W., Ge, Y., Carrington, L. E., Hume, D. A., Hamilton, S. & de Jersey, J. (2000) *Gene* **250**, 117–125.
30. Bozzo, G. G., Raghothama, K. G. & Plaxton, W. C. (2002) *Eur. J. Biochem.* **269**, 6278–6286.
31. Laing, W. A., Frearson, N., Bulley, S. & MacRae, E. (2004) *Funct. Plant Biol.* **31**, 1015–1025.
32. Chifflet, S., Torriglia, A., Chiesa, R. & Tolosa, S. (1988) *Anal. Biochem.* **168**, 1–4.
33. Gatzek, S., Wheeler, G. L. & Smirnov, N. (2002) *Plant J.* **30**, 541–553.
34. Schagger, H. & von Jagow, G. (1987) *Anal. Biochem.* **166**, 368–379.
35. Jimenez, C. R., Huang, L., Qiu, Y. & Burlingame, A. L. (2003) in *Current Protocols in Protein Science*, eds. Coligan, J. E., Dunn, B. M., Ploegh, H. L., Speicher, D. W. & Wingfield, P. T. (Wiley, Hoboken, NJ), pp. 16.4.1–16.4.5.
36. Eng, J., McCormack, A. L. & Yates, J. R. (1994) *J. Am. Soc. Mass Spectrom.* **5**, 976–989.
37. Yates, J. R., Eng, J. K., McCormack, A. L. & Schieltz, D. (1995) *Anal. Chem.* **67**, 1426–1436.
38. Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G. & Gibson, T. J. (1998) *Trends Biochem. Sci.* **23**, 403–405.
39. Majerus, P. W., Kisseleva, M. V. & Norris, F. A. (1999) *J. Biol. Chem.* **274**, 10669–10672.
40. York, J. D., Ponder, J. W. & Majerus, P. W. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 5149–5153.
41. Matsuhisa, A., Suzuki, N., Noda, T. & Shiba, K. (1995) *J. Bacteriol.* **177**, 200–205.
42. Parthasarathy, R., Parthasarathy, L. & Vadnal, R. (1997) *Brain Res.* **778**, 99–106.
43. Bhat, P. J. (2003) *Med. Hypotheses* **60**, 123–128.
44. Stec, B., Yang, H., Johnson, K. A., Chen, L. & Roberts, M. F. (2000) *Nat. Struct. Biol.* **7**, 1046–1050.
45. Styer, J. C., Keddle, J., Spence, J. & Gillaspay, G. E. (2004) *Gene* **326**, 35–41.