Resurrecting the Ancestral Enzymatic Role of a Modulatory Subunit

In the post-genomic era, functional prediction of genes is largely based on sequence similarity searches, but sometimes the homologues bear different roles because of evolutionary adaptations. For instance, the existence of enzyme and non-enzyme homologues poses a difficult case for function prediction and the extent of this phenomenon is just starting to be surveyed. Different evolutionary paths are theoretically possible for the loss or acquisition of enzyme function. Here we studied the ancestral role of a model non-catalytic modulatory subunit. With a rational approach, we “resurrected” enzymatic activity from that subunit to experimentally prove that it derived from a catalytic ancestor. We show that this protein (L subunit ADP-glucose pyrophosphorylase) evolved to have a regulatory role, losing catalytic residues more than 130 million years ago, but preserving, possibly as a by-product, the substrate site architecture. Inactivation of catalytic subunits could be the consequence of a general evolutionary strategy to explore new regulatory roles in hetero-oligomers.

An attractive model to study the evolution of allosteric control is ADP-glucose pyrophosphorylase (ADP-Glc PPase), a regulatory enzyme for the synthesis of bacterial glycogen and starch in plants. It has evolved to be regulated by different key metabolites depending on the main carbon assimilation pathway of the organism (1, 2). The quaternary structure plays an important role in its allosteric regulation. Most of the bacterial ADP-Glc PPases are homotetramers (1), whereas the plant enzymes comprise two types of homologous but distinct subunits (2–5), the small (S subunit, 50–54 kDa), and the large (L subunit, 51–60 kDa). They form a heterotetrameric structure (S2L2), as well as bacterial homotetramers (S4), as well as bacterial ADP-Glc PPases (1). However, we cannot simply rule out the second possibility because of the existence of non-catalytic homologues such as the products of the glgD genes from Gram-positive bacteria (8). There is strong evidence that the L subunit from the potato (Solanum tuberosum L.) tuber ADP-Glc PPase binds substrates. The heterotetramer (S2L2), as well as bacterial homotetramers (S4), binds four ADP-[3-14C]glucose molecules.”

EXPERIMENTAL PROCEDURES

Materials

α-[U-14C]Glucose 1-phosphate (GlcIP) was purchased from Amer sham Biosciences. [32P]PPi, was purchased from PerkinElmer Life Sciences. GlcIP, ATP, ADP-glucose (ADP-Glc), 3-phosphoglycerate (3-PGA), and inorganic pyrophosphatase were purchased from Sigma. Pfu DNA polymerase was purchased from Stratagene (La Jolla, CA, USA). Ampligase®, a thermostable DNA ligase, was purchased from Epicenter (Madison, WI). All other reagents were purchased at the highest quality available.

DNA Methods

The Macromolecular Structure, Sequencing, and Synthesis Facility (MSF) at Michigan State University performed the synthesis of oligonucleotides and automated DNA sequencing.

Site-directed Mutagenesis

Mutant D145N on the S subunit was obtained as described previously (6). Combined chain reaction (CCR) was used to introduce other mutations (11). The plasmids encoding the wild-type S subunit (pML10), the S subunit mutant D145N (pML10-D145N), or the wild-type L subunit (pMON17366) were used as templates for the mutagenesis (12, 13). In the S subunit, the residues were changed with the following oligonucleotides: K33K, 5’-GGA GCT GGG ACC AAA CTT TAT CCT CTA-3’; K43T, 5’-AAA AAA AGA GCA ACC CCA GCT GTT CCA-3’; K198R, 5’-G'TA ATT GCA GAG CCC CCA AAA GAG GAC C-3’. In the L subunit, the oligonucleotides used were: K44R, 5’-GGA GAA GGG ACC CCG TTA TTC CCA CTT A-3’; T54K, 5’-AGT AGA ACT

This paper is available on line at http://www.jbc.org

Received for publication, December 1, 2004, and in revised form, January 3, 2005 Published, JBC Papers in Press, January 4, 2005, DOI 10.1074/jbc.M413540200

Miguel A. Ballicora, Jennifer R. Dubay, Claire H. Devillers‡, and Jack Preiss§

From the Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan 48824
Evolution of Homologous Subunits

GCA AAG CCT GTP CGG-3’; K213R, S-AG TTT GCT GAA GTT CCA AAA GGT TT-3’. L subunit mutants on the residue 160 were obtained as described previously (6). The mutated fragments were excised, and the double mutant plasmid pMON17336-K41R/T54K was used under the restriction sites SpeI and SacI to obtain the triple mutants LK44R/T54K/D160N and LK44R/T54K/D160E. The coding regions of the mutant plasmids were sequenced to confirm that only the introduced mutations were introduced.

Protein Methods

Protein concentration during enzyme purification was measured by using bicinchoninic acid reagent (14) from Pierce, with bovine serum albumin as standard. Interfering substances were previously removed precipitating with 12% trichloroacetic acid and 0.025% sodium deoxycholate and dissolving in 5% SDS, 0.1 M NaOH. Protein concentration of the purified enzymes was determined by UV absorbance at 280 nm using an extinction coefficient of 1.0 (6). Electrophoresis (SDS-PAGE) and immunoblotting of the protein samples were performed as described previously (6). Samples were desalted with Bio-Rad 10 DG chromatography columns. Centricon-30 devices (Amicon Inc.) were used to concentrate the enzymes.

Enzyme Assays

Pyrophosphorolysis Direction—The formation of [32P]ATP from [32P]PP, and ADP-Glc was measured as follows. The standard aqueous reaction mixture contained 50 mM Hepes buffer (pH 8.0), 7 mM MgCl₂, 2 mM dithiothreitol, 0.5 mM [14C]Glc1P (800–1000 cpm/nmol), plus enzyme in a total volume of 0.25 ml. After 10 min of incubation at 37 °C, the reaction was terminated by the addition of 3 ml of cold 5% trichloroacetic acid and dissolving in 5% SDS, 0.1 M NaOH. Protein concentration of the purified enzymes was determined by UV absorbance at 280 nm using an extinction coefficient of 1.0 (6). Electrophoresis (SDS-PAGE) and immunoblotting of the protein samples were performed as described previously (6). Samples were desalted with Bio-Rad 10 DG chromatography columns. Centricon-30 devices (Amicon Inc.) were used to concentrate the enzymes.

Expression and Purification of Mutant Enzymes

The Escherichia coli AC70R1-504 cells, which lack endogenous ADP-Glc PPase activity (12), were co-transformed with compatible plasmid pMAL (S subunit) and pMON17336 (L subunit) or their mutated derivatives for co-expression of the subunits. Transformed cells were grown in 2 liter of LB medium, induced, harvested, and sonicated in buffer A (50 mM Hepes (pH 8.0), 5 mM MgCl₂, 0.1 mM EDTA, 10% sucrose) as described previously (16). The purification was performed at 0–4 °C. The crude extract was applied onto a 12-ml DEAE-Sepharose FF column (Amersham Biosciences) equilibrated in buffer A and eluted with a linear NaCl gradient (10 column volumes, 0–0.5 M). Purest fractions were pooled, desalted, and concentrated, and applied onto a Mono Q HR 16/10 (FPLC, Amersham Biosciences) column equilibrated with buffer A and eluted with a linear NaCl gradient (16 bed volumes, 0.1–0.5 M). The post-Mono Q fractions were pooled, concentrated, resuspended in buffer B (buffer A plus 1.3 M ammonium sulfate, pH 8.0), put onto a phenyl-Sepharose CL-6B column (Amersham Biosciences) column equilibrated with buffer B. The sample was eluted with a decreasing linear gradient (20 bed volumes) of ammonium sulfate (1.3 to 0 M ammonium sulfate). The purest fractions were pooled, concentrated, and desalted. After this procedure, enzymes were >95% pure. The enzymes were followed by pyrophosphorolysis activity, except S₁₀₀₀₀₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅₋₅-
The model was checked with the programs Verify3D (33) and PRO-CHECK (34). Minor adjustments in the alignments with the templates were checked and modified iteratively to improve the three-dimension-al-one-dimensional scores provided by Verify3D along the sequence. There was no structural change, in any of the iterations, in the area of the residues analyzed in this work. The region between residues 217–232, which corresponds to an insertion, was the only one with low 3D-1D score. This is expected from insertions of more than eight residues (35). However, this area is very far from the region analyzed in the paper, in the other lobe of the catalytic domain, exposed, and not facing the active site.

Recently, the crystal structure of the tetrameric form (S) of the L subunit from potato tuber was solved (36). This structure was in an inhibited conformation, which made it unsuitable as a template for modeling an active L subunit. However, it confirmed that the templates selected in this study (RiFH and Agx2) were appropriate, because they share the same fold in the catalytic domain as previously predicted (2). RiFH and Agx2 are not allosterically regulated and are better templates to model the active conformation of the substrate site of the L subunit.

**Phylogenetic Analysis**

ADP-Glc PPase sequences from Supplemental Table I were obtained from the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov). The expression data were taken from the literature. Multiple sequence alignment was performed by the Clustal program (37) available on-line (clustalw.genome.jp), and afterward, it was manually refined with the BioEdit v8.0.7 program (www.mbio.ncsu.edu/BioEdit/bioedit.html). For the phylogenetic analysis, the N terminus of all the sequences was discarded. That region is not homologous and corresponds to the leading peptide for subcellular targeting, which has been shown in the S subunits that belongs to an exon with a different evolutionary history (38). Alignment started in residue 2 of the mature recombinant S subunit from potato tuber (A^VSDQS^QS^QKT^ . . . ) and residue 13 of the mature recombinant L subunit from potato tuber (Q^3TVFVFDMPR^ . . . ). The unrooted tree was built by the neighbor-joining method (39) as implemented in the PHYLIP package (evolution-genetics.washington.edu/phylip.html), version 3.6. The tree was drawn and plotted with the program TREE-PUZZLE (41). The topology of the tree of the L subunit with the program Phylodraw (40). The outgroup chosen was the ADP-Glc PPase from *Anabaena* because cyanobacterial enzymes are the closest to plant enzymes (2). The topology of the tree of the L subunit gene family from plants was further investigated using the Maximum-likelihood method as implemented in the program TREE-PUZZLE (41). The unrooted tree was built by the neighbor-joining method (39) as implemented in the PHYLIP package (evolution-genetics.washington.edu/phylip.html), version 3.6. The tree was drawn and plotted with the program TREE-PUZZLE (41).

**RESULTS**

Identification of critical missing residues was the first important step to design an active L subunit. A sequence alignment of ADP-Glc PPases with reported activity (experimentally assayed) allowed us to focus on few candidates. The set of invariants should contain most of, if not all, the essential residues. The subset of the ones absent in the L subunit was of special interest. We chose Lys^44 and Thr^54 in the L subunit from potato tuber as the best candidates to study because the homologous residues (Arg^33 and Lys^43 in the S subunit, Fig. 1A) were differentially conserved in the active bacterial and S subunits (see a detailed analysis under “Experimental Procedures”). In addition, Lys^44 and Thr^54 are in a highly conserved region of ADP-Glc PPases (Fig. 1A).

**Resurrection of ADP-Glc PPase Catalysis**—To convert the modular L subunit into a catalytic one, we substituted Lys^44 and Thr^54 by Arg^44 and Lys^43, respectively. The mutant L^K44R/ T54K was expressed alone, but no activity was detected, most probably because the L subunit is unable to form a stable tetramer in absence of the S subunit (43). The activity of L subunit mutants cannot be readily tested with a co-expressed wild-type S subunit because of the intrinsic activity of the latter. We co-expressed the L subunit mutants with S^DI45N, an inactive S subunit in which the catalytic Asp^145 was mutated (6). In this way, the activity deriving from the S subunit was reduced more than three orders of magnitude (Table I). Co-expression of the L subunit double mutant L^K44R/T54K with S^DI45N generated an enzyme with considerable activity, 10 and 18% of the wild-type enzyme (S^WT/L^WT) in the forward and reverse direction, respectively (Table I). Single mutation K44R generated an enzyme (S^DI45NL^K44R) with no significant activity over the control (S^DI45N/L^WT). Mutation T54K (S^DI45NL^K44R/T54K) was more effective, but the combination of both mutations in the L subunit (S^DI45NL^K44R/T54K) had the most dramatic effect (Table I). Therefore, we concluded that the two residues Arg^44 and Lys^54 are needed for restoring catalytic activity to the L subunit.

The resurrection of catalytic activity in the L subunit by incorporation of Arg^44 and Lys^54 predicted that the homologous residues in the S subunit are also critical. Replacement of those two residues by the homologous in the L subunits (mutations R33K and K43T) decreased the activity one and two orders of magnitude, respectively, in both directions, confirming the hypothesis (Table I). The mutant enzymes were still activated by 3-PGA and inhibited by orthophosphate (P_1) (data not shown), which are the main regulators of plant enzymes (2). The wild-type enzyme and S^DI45NL^K44R/T54K had very similar kinetic properties.
**Evolution of Homologous Subunits**

The reaction in the forward (synthesis of ADP-Glc) direction uses ATP and Glc1P as substrates releasing PPi and ADP-Glc as products. The reaction can also be assayed in the reverse (pyrophosphorolysis of ADP-Glc) direction. Enzymes were obtained by co-expression of the S and L subunits and purified to homogeneity, and the activity was assayed in both directions of the reaction as described under "Experimental Procedures." WT, wild type.

| Subunits | S | L | Forward (synthesis) | Reverse (pyrophosphorolysis) |
|----------|---|---|---------------------|-----------------------------|
|          |   |   | $s$ $^{-1}$         | $s$                         |
| WT + WT  | 108 ± 3 | 165 ± 7 |
| D145N + WT | 0.057 ± 0.003 | 0.125 ± 0.007 |
| D145N + K44R | 0.104 ± 0.003 | 0.111 ± 0.007 |
| D145N + T54K | 3.1 ± 0.3 | 1.9 ± 0.1 |
| D145N + K44R/T54K | 10.8 ± 0.7 | 30.3 ± 2.4 |
| R33R + WT | 12.1 ± 0.3 | 13.8 ± 0.3 |
| K43T + WT | 1.06 ± 0.03 | 0.94 ± 0.03 |

- $a$ Activator (4 mM 3-PGA) and substrates (2 mM ATP and 0.5 mM Glc1P) were saturating.
- $b$ Reverse (4 mM 3-PGA) and substrate ADP-Glc (2 mM) were saturating. The concentration of PPi (1.4 mM) was the highest to avoid precipitation in the assay mixture.

| Parameter | $S_{0.5}$ (μM) | $k_{cat}$ | $S_0$ (μM) | $k_{cat}$ |
|-----------|----------------|-----------|------------|-----------|
| ATP       | 86 ± 7 (1.6)   | 97 ± 7 (1.6) | 170 ± 13 (1.8) |
| Glc1P     | 27 ± 2 (1.1)   | 27 ± 2 (1.1) | 11 ± 1 (1.2) |
| 3-PGA     | 60 ± 5          | 135 ± 11 (0.9) | 13.6 ± 0.7 (1.3) |
| ADP-Glc   | 200 ± 20 (1.4) | 200 ± 20 (1.4) | 70 ± 6 (1.5) |
| PP        | 37 ± 2 (1.0)   | 37 ± 2 (1.0) | 98 ± 10 (1.0) |
| 3-PGA     | 3.1 ± 0.3      | 3.1 ± 0.3  | 7.6 ± 1.0  |
| ADP-Glc   | 1.1 ± 0.3 (1.1)| 1.1 ± 0.3 (1.1) | 4.0 ± 0.5 (0.9) |
| 3-PGA (+ 2 mM P) | 96 ± 19 (1.9) | 96 ± 19 (1.9) | 101 ± 13 (1.1) |

- $a$ Coefficients of the Hill equation ($n_H$) are in parentheses.
- $b$ Activation values (-fold) are the ratio between the activity at saturated concentration of activator (3-PGA) and the activity in absence of activator.

The fact that only two mutations in the L subunit restored enzyme activity is the ultimate evidence that the L subunit derives from a catalytic ancestor. To confirm that the catalysis occurs in the L subunit of S$_{D145N}$L$_{K44R/T54K}$, we disrupted the substrate site in each subunit and compared the kinetic properties. Previously, in the wild-type enzyme, replacement of Lys$_{44}$ in the S subunit decreased the substrate affinity (Glc1P), whereas disruption of the homologous residue Lys$_{54}$ (Fig. 1) in the L subunit did not have the same effect (9). In the case of S$_{D145N}$L$_{K44R/T54K}$, the mutation K213R on the L subunit severely decreased the substrate affinity for Glc1P, whereas K198R on the S subunit did not (Table III). This indicated that the L subunit double mutant, and not S$_{D145N}$, was catalytic. In the wild-type enzyme, Lys$_{54}$ does not seem to play any important role; but in S$_{D145N}$L$_{K44R/T54K}$, it recovered its ancestral ability to confer to the enzyme a high apparent affinity for Glc1P. Previous results showed that Asp$_{145}$ in the S$_{WT}$ subunit is essential for catalysis, whereas the homologous Asp$_{160}$ in the L$_{WT}$ subunit is not (6). Conversely, in this study, mutation D160N or D160E in the L$_{K44R/T54K}$ subunit abolished the activity (Table IV), which shows the ancestral essential role of this residue and confirms that the catalysis of S$_{D145N}$L$_{K44R/T54K}$ occurs in the L subunit.

**Structural Basis for the Ancestral Function**—A comparative model of L$_{K44R/T54K}$ illustrates the predicted role of Arg$_{44}$ and Lys$_{54}$ (Fig. 1B). In the model, Asp$_{44}$, which is homologous to the catalytic Asp$_{45}$ in the S subunit and catalytic Asp$_{45}$ in the E. coli ADP-Glc PPase (6, 25), interacts with Lys$_{54}$. This type of interaction (Lys$_{54}$-Asp$_{45}$) has also been observed in crystal structures of enzymes that catalyze similar reactions, such as dTDP-glucose pyrophosphorylase (dTDP-Glc PPase) and UDP-N-acetylglucosamine pyrophosphorylase (UDP-GlcNAC PPase), and postulated to be important for catalysis by correctly orienting the Asp (28, 44, 45). In addition, Lys$_{54}$ interacts with the oxygen bridging the alpha- and beta-phosphates as it has been observed in the crystal structure of E. coli dTDP-Glc PPase (28). That contact may neutralize a negative charge density stabilizing the transition state and making the pyrophosphate a better leaving group. Arg$_{44}$ interacts in the model with the beta- and gamma-phosphates of ATP, which correspond to the PPi product (Fig. 1). Similarly, Arg$_{15}$ in the E. coli dTDP-Glc PPase was postulated to contribute to the departure of PPi (44). The kinetic data agreed with the predicted interaction of PPi with Arg$_{44}$. A Lys in that position, in both the catalytic L subunit and the non-catalytic L subunit, Lys$_{54}$ and Thr$_{54}$ do not interact as Arg$_{54}$ and Lys$_{54}$ (Fig. 1).

**Phylogenetic Analysis**—A phylogenetic tree of the ADP-Glc PPases from photosynthetic eukaryotes provides information about the origin of the subunits. In multicellular organisms, S and L subunits form two and four distinct groups, respec-
SR33KLWT, S D145NLK44R/T54K, and S D145NLT54K. Complete saturation of the wild-type enzyme and the mutants (subunit) on the pyrophosphate saturation curves. Concentrations of PPi lead to precipitation and cloudiness of the reaction mixture. The kinetic constants were calculated as described under "Experimental Procedures." Pyrophosphorolysis activity was measured in crude extracts. Concentrations of the activator were saturating. The increase values were the relative values of $S_{0.5}$ compared to the enzyme $S_{145NLT44R/T54K}$ (Table II).

### Table III
*Effect of the mutations on the Glc1P site*

| Subunits | Glc1P, $S_{0.5}$ | Increase | $k_{cat}$ (Forward) | $k_{cat}$ (Reverse) |
|----------|-----------------|----------|---------------------|---------------------|
| D145N + K44R/T54K/K213R | 910 ± 110 | 82-fold | 4.4 ± 0.3 | 14.1 ± 0.3 |
| D145N/K198R + K44R/T54K | 13.2 ± 0.5 | 1.2-fold | 7.4 ± 0.3 | 20.5 ± 0.7 |

### Table IV
*Effect of mutations on the residue Asp100 on the L subunit*

Enzymes were obtained by co-expression of the L and S subunits as described under “Experimental Procedures.” Pyrophosphorolysis activity was measured in crude extracts. Concentrations of the activator 3-PGA (4 mM) and substrate ADP-Glc (2 mM) were saturating. The saturating concentration of substrate PPi (1.4 mM) was the highest concentration to avoid precipitation.

![Graph](image)

**Fig. 2.** Effect of the mutations R33K (S subunit) and K44R (L subunit) on the pyrophosphate saturation curves. Activity was analyzed in the pyrophosphorolysis direction as indicated under “Experimental Procedures” for the wild-type enzyme and the mutants $S_{R33K/LWT}$, $S_{D145N/LK44R/T54K}$, and $S_{D145NL/LT54K}$. Complete saturation of substrate PPi could not be reached for the mutants $S_{R33K/LWT}$ and $S_{D145NL/LT54K}$. Higher concentrations of PPi lead to precipitation and cloudiness of the reaction mixture. The kinetic constants were calculated as described under “Experimental Procedures.”

**DISCUSSION**

The resurrection of the enzymatic activity of the L subunit ADP-Glc PPase by only two mutations reveals the catalytic ancestry. In addition, it points out the important role of the residues Arg44 and Lys54 for catalysis. A comparison with other pyrophosphorylases agrees with this data. The importance of Arg44 and Lys54 in the Active L Subunit Mutant Is Possibly Widespread in Other Pyrophosphorylases—The identity between ADP-Glc PPases and other nucleotide-diphosphate-sugar pyrophosphorylases is very low (~20%). However, prediction of the secondary structure and further structure-based alignment of the sequences emphasized the similarities (1, 2, 25). Residues Arg44 and Lys54 of the active L subunit mutant are not only present in active ADP-Glc PPases but also in other pyrophosphorylases in structurally similar positions (25, 44, 45). In addition, in some cases these residues were shown to be important. For instance, Lys25 in the UDP-GlcNAC PPase (GlmU), homologous to Lys54 in the mutant $L_{K44R/T54K}$, was mutated to Ala and the specific activity decreased 8-fold (45). In this work, homology modeling predicted that Arg44, an important residue for activity in the mutant $L_{K44R/T54K}$, interacts with PPi, making it a better leaving group in the forward (synthesis) direction. In the reverse (pyrophosphorolysis) direction, it is predicted to have a role in PPi binding as a substrate. In other pyrophosphorylases, a non-conservative mutation from Arg to Ala, at the homologous position, reduced the specific activity in the nucleotide-sugar synthesis direction two to three orders of magnitude (45, 46). Unfortunately, the effect on the PPi apparent affinity has not been examined. The presence of homologous residues to Arg44 and Lys54 suggest that they may be also important in other distantly related pyrophosphorylases, such as the enzyme 4-diphosphocytidyl-2-C-methylerythritol synthetase (47). In this case, despite the very low percentage of identity, the homology can also be identified with a structural alignment.

Evolutionary History of ADP-Glc PPase L Subunits—We cannot ascertain when the L subunit from potato tuber became modulatory only by phylogenetic analysis, which is based upon primary structure rather than function. However, with structure-function relationship information, we could theoretically trace function or the absence of it in the tree. Both groups III and IV lack the residues homologous to Arg33 and Lys43 (SWT numbering, Fig. 1A). Group III has Lys, His, and Gln, and group IV has Gln, respectively, in place of a homologous Arg33 (Fig. 3). A distinctive characteristic in both groups is that a Thr replaces the catalytic Lys43 (Fig. 3). The most likely scenario supported by phylogenetic analysis (Fig. 3) is that the shared ancestor of both branches lacked a Lys in that position and, consequently, was already non-catalytic. Because one branch (III) is composed of dicot and the other (IV) of monocot plants, we can infer that the L subunit from groups III and IV ceased to be catalytic before monocots split from dicots among the angiosperms more than 130 million years ago (48, 49). Conceivably, they had already acquired adaptive modulatory properties at that point in evolution given that catalysis had been already lost (Fig. 3).
Evolution of Homologous Subunits

Modulatory L subunits from branches I and II of the phylogenetic tree may be non-catalytic too (7), but the reasons for this lack of activity are not clear. They contain in their sequence the important residues analyzed in this work (Fig. 3), as well as all the others conserved in the catalytic subunits. This presence predicts a catalytic role; however, the only biochemical evidence gathered so far has not confirmed it. None of the L subunits from A. thaliana (groups I–III, supplemental Table I), when expressed in presence of an inactive S subunit (APS2), originated a form with detectable activity (7). Despite no apparent essential residue seems to be missing, other subtle ways of enzyme inactivation during the course of evolution can explain those results (50). At this time, we cannot discard the alternative hypothesis that A. thaliana L subunits from groups I and II are catalytic but could not interact properly with APS2.

In the literature, no other L subunit from these two groups has been expressed and tested for intrinsic activity. Based on sequence comparison, we cannot rule out the possibility of finding a catalytically competent subunit among groups I and II.

Phylogenetic trees based on protein sequences are very powerful tools to test evolutionary hypothesis, but they do not show the functional information that those sequences carry. Gene synthesis of hypothetical ancestors and characterization of their products may tackle this problem (51). As performed in this paper, analyses of structure-function relationships should complement phylogenetic analysis for tracing and understanding how and when certain enzymes acquired new roles and lost catalytic function. The use of this information could potentially increase the accuracy of function prediction for these difficult cases (50, 52).

Enzyme Inactivation as a General Strategy—Divergence from an ancestral catalytic subunit to form hetero-oligomers with non-catalytic and catalytic subunits could have been a strategy used by nature to add complexity and explore new regulatory functions. There are phylogenetic evidences suggesting this could be a general evolutionary mechanism. In a superfamily of enzymes, if only one branch (“odd one out”) in the tree is non-catalytic, the common ancestor most probably was catalytic as the majority of the branches. Six cases of hetero-oligomers with solved three-dimensional structures that meet this criterion have been observed (50). In this work, we experimentally tested this hypothesis for a plant ADP-Glc PPase and found that the common ancestor of catalytic and non-catalytic subunits was catalytic. This enzyme is a very attractive model because the “inactivated” subunit acquired differential modulatory properties in plant tissues, expanding the plasticity for regulation of a key synthetic pathway. After gene duplication, the ADP-Glc PPase subunits acquired asymmetric functions. The S subunit remained catalytic with defective allosterism whereas the L subunit became modulatory (i.e. modifying the properties of the S subunit) and lost catalytic function.

Considering the ancient inactivation of L subunits and their higher divergence (Fig. 3), it is remarkable that the general architecture of their substrate site has not changed much by evolution. The wild-type enzyme and S<sub>Δ145NLK44R/T54K</sub> had comparable apparent affinities for the substrates. This preservation occurred despite substrate binding does not seem to be an adaptive pressure because is not needed for the modulatory roles of the L subunit. It was shown previously that mutations in the L subunit of a critical residue for Glc1P binding (Lys<sub>213</sub>) do not affect the modulatory properties of the potato tuber ADP-Glc PPase (9). In addition, in the wheat endosperm enzyme, Glu has replaced that Lys. Perhaps, the active-site architecture has been kept as a by-product of evolution or spanneld (53) because some disruptions could indirectly affect the modulatory properties. For instance, Asp<sup>160</sup> in the wild-type L subunit is conserved in all ADP-Glc and other nucleotide-sugar PPases (6, 25, 44). In the catalytic subunits, it has a catalytic role but replacement in the wild-type L subunit of Asp<sup>160</sup> altered the modulatory properties of the tetramer (6). This residue, Asp<sup>160</sup>, “recovered” its critical role for activity in the L<sub>Δ145NLK44R/T54K</sub> mutant (Table IV).

Adaptation from enzymes to non-enzymes and vice versa is a very exciting topic in protein evolution because it involves the generation of novel functions (50). Not only subunits from hetero-oligomers, but also many enzymes and non-enzymes homologues, are being detected. Based on the odd one out criterion, it has been reported some cases of non-enzymes that became enzymes, but the opposite scenario was observed more often.
often (50, 54). Some examples may be critical for the appearance of higher levels of complexity and control. For instance, a thrilling proposal is that many transcription regulators have an ancient catalytic past (55, 56). Supporting experimental evidence, as performed in the present study, will be enlightening to confirm these hypotheses and understand how key regulatory roles have evolved at a molecular level.

Acknowledgments—We thank C. M. Bejar and A. Yep for stimulating discussions and S. Wakao and Dr. E. S. Trombetta for helpful criticisms and revision of the manuscript.

REFERENCES

1. Ballicora, M. A., Iglesias, A. A., and Preiss, J. (2003) Microbiol. Mol. Biol. Rev. 67, 213–225.
2. Ballicora, M. A., Iglesias, A. A., and Preiss, J. (2004) Photosynth. Res. 79, 1–24.
3. Morell, M. K., Blom, M., Knowles, V., and Preiss, J. (1987) Plant Physiol. 85, 182–187.
4. Okita, T., Nakata, P. A., Anderson, J. M., Sowokinos, J. R., Morell, M., and Preiss, J. (1998) J. Biol. Chem. 273, 18228–18236.
5. Okita, T., Nakata, P. A., Anderson, J. M., Sowokinos, J. R., Morell, M., and Preiss, J. (1999) Plant Physiol. 120, 765–770.
6. Okita, T., Nakata, P. A., Anderson, J. M., Sowokinos, J. R., Morell, M., and Preiss, J. (1999) J. Biol. Chem. 274, 17996–18006.
7. Okita, T., Nakata, P. A., Anderson, J. M., Sowokinos, J. R., Morell, M., and Preiss, J. (1999) Microbiol. Mol. Biol. Rev. 63, 64–71.