

This is to certify that the  
dissertation entitled

STRUCTURE FUNCTION RELATIONSHIP STUDIES OF  
*ESCHERICHIA COLI*  
ADP-GLUCOSE PYROPHOSPHORYLASE

presented by

CLARISA MARIA BEJAR

has been accepted towards fulfillment  
of the requirements for the

Ph.D. degree in Biochemistry and  
Molecular Biology

Joel Paris

Major Professor's Signature

12-1-06

Date

*MSU is an Affirmative Action/Equal Opportunity Institution*

LIBRARY  
Michigan State  
University



STRUCTURE FUNCTION RELATIONSHIP STUDIES OF  
*ESCHERICHIA COLI*  
ADP-GLUCOSE PYROPHOSPHORYLASE

By

Clarisa Maria Bejar

A DISSERTATION

Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of

DOCTOR OF PHILOSOPHY

Department of Biochemistry and Molecular Biology

2006

## ABSTRACT

### STRUCTURE FUNCTION RELATIONSHIP STUDIES OF *ESCHERICHIA COLI* ADP-GLUCOSE PYROPHOSPHORYLASE

By

Clarisa Maria Bejar

Alpha-1,4-polysaccharides are one of the main energy and carbon storage components present in most living cells. Bacteria, higher plants and algae utilize ADP-glucose as the activated form of glucose for polyglucan synthesis. This step is catalyzed by ADP-glucose pyrophosphorylase, a tightly regulated enzyme responsible for determining the rate of the entire pathway. The mechanism of allosteric regulation is linked to the carbon assimilation pathway in the cell, suggesting the evolutionary adaptation of this enzyme to respond to the specific carbon and energy necessities within the respective cell.

The allosteric properties of the enzyme depend not only on specific residues that have been identified by several genetic and biochemical studies, but also on more extended regions of the protein. In the present work, the domain organization of the *Escherichia coli* ADP-glucose pyrophosphorylase monomer and the functional and structural role of the N- and C-terminus in regulation were studied. The biochemical data presented here in combination with structural information allow a mechanism to be proposed for how the putative allosteric sites communicate with the active site to regulate enzyme activity.

Structural information on ADP-glucose pyrophosphorylases at the three-dimensional level also led to the investigation of other regions of the enzyme with predicted functional roles. Structure-guided site-directed mutagenesis was used to study the molecular basis of substrate binding and affinity. The detailed characterization of the *E. coli* enzyme's glucose 1-phosphate site architecture may have relevance not only for the members of the ADP-glucose pyrophosphorylase family, but also for other nucleotide-diphospho-sugar pyrophosphorylases. .

## ACKNOWLEDGMENTS

My time in the Department of Biochemistry and Molecular Biology at Michigan State University has been an incredible opportunity to grow both academically and personally. I not only learnt about enzymes, carbohydrate metabolism (and other scientific topics) and new experimental skills but I was also able to develop a critical scientific thinking when addressing a biological problem. The classes (how to forget BMB801!!), the seminars, the troubleshooting in the lab, the exposure to excellent scientists and all the resources offered by this department have been essential for shaping the scientist I am today. Outside the academic world I had the chance to be exposed to different cultures, languages and traditions from around the world and I am very thankful for the amazing people I have met and the wonderful experiences we have shared together.

Thanks to my PhD mentor, Dr. Jack Preiss, who gave me the possibility to join his lab and the “Pyro” team. I am very lucky and grateful for having the chance to work next to and learn from someone so experienced and knowledgeable in the field as he is.

Thanks to the professors in my PhD Advisory Committee: Dr. Christoph Benning, Dr. Shelagh Ferguson-Miller, Dr. Charles Hoogstraten, all of them from the Department of Biochemistry and Molecular Biology, and Dr. Mark Worden, from the Department of Chemical Engineering. Special thanks to Christoph for being open, for listening, for his continuous support at every moment of this PhD and for considering me as another

member of his lab. Thanks to Shelagh for her advice and her support to perform the computational modeling which has been such an enjoyable approach for me to learn. Thanks to Charlie for his positive point of view about enzymology and life, and for his English grammar lessons in the correction of this thesis. Thanks to Dr. Worden who introduced me to the Technologies for a Bio-based economy program which allowed me to get better exposure to the biotechnological industry as well as to get funding from the Department of Energy of the US that supported me during three years of my PhD.

Invaluable thanks to the former members of the Preiss lab: Drs. Miguel Ballicora, Alberto Iglesias and Alejandra Yep for their time, their advice, their support, their patience, their never-ending jokes, the mates, the Sparty's coffees and the talks about Argentinean history, literature and old TV shows. I really felt like at home with them. Special thanks to Miguel for teaching me to be like the water, to Alberto for the Almafuerte's poems and to Ale for all the time we spent working, talking, singing, talking, crying, talking, traveling, talking, cooking and talking together. To Claire Devillers for the great time during my first years in East Lansing. To Ana Bertolo, Pedro Crevillen and Tiziana Ventriglia, who visited the Preiss' lab in different summers giving us the chance to learn from their research and to laugh with them so much! And to Dr. Mirta Sivak who has always made sure that, although we are in the US, we have an Argentinean mom around taking good care of all of us.

Thanks to the Department of BMB for their friendly and supportive environment. Thanks to my dear friends in the department for their constant support, advice and fun

times: Hiroshi Maeda, Harini Krishnamurthy, Setsuko Wakao, Francisco Herrera, Dean Shooltz, Clarie Viellie, Karla Ziegelmann-Field, Jake McKinley, Soledad Quiroz, Cora Fong and the whole Benning lab. Thanks to Pappan and the administrative stuff (Julie, Chris, Leslie, Melinda, Jessica and Teresa) for being always so helpful and cooperative. To Joe Leykam for his advice and the interesting stories from California and to Dr. Leslie Kuhn for her support during my postdoctoral job search.

Thanks to my great friends outside the lab. Very special thanks to Gisela and Nate, for encouraging me to pursue this PhD, for their constant support, for the great times together and for Sophie. To Ale Yep (again), Javier, Valentina, Rafael, Pako, Phil, Gaby Ballicora and Ale Manzan for the many Argentinean laughs. To all the friends I met during these years, the Spanish, the Italian, the French, the Latin Americans, the Japanese... for the fun parties, the opera performances, the many trips, the wine-tasting nights and the company. They have all become my family here.

Thanks to what is most important to me: my family in Argentina. To my mom Dorita, my dad Hector, my sister Analia, my cousin Valeria, my dearest uncle Rubenacho, my aunts, uncles and cousins for their constant communication, joy and support. Thanks Lili and Gordi, Tati, Ceci, Vangi, Leonora, Irene, Gaby, Vale Palanza for being always, always, always there for me. Thanks to Internet for bringing all these people much closer.

Thanks to Brad for showing up in my life, for his support, friendship and love during this last year of my PhD and for all the exciting times to come.

## TABLE OF CONTENTS

<b>List of Tables</b> .....	x
<b>List of Figures</b> .....	xi
<b>List of Abbreviations</b> .....	xiii
<b>Chapter 1:</b>	
<b>Literature review</b> .....	1-39
1.1 Introduction to polysaccharides .....	2
1.2 Storage $\alpha$ -1,4-polysaccharides: structure and function .....	3
1.3 Elucidation of the $\alpha$ -1,4-polyglucan metabolic routes: historical background .....	9
1.4 Metabolic pathways for $\alpha$ -1,4-polysaccharides synthesis in bacteria and plants .....	11
1.4.1 ADP-glucose pyrophosphorylase .....	14
1.4.2 Bacterial glycogen / starch synthase .....	16
1.4.3 Branching enzyme .....	19
1.5 Alpha-1,4-polysaccharides in non-photosynthetic eukaryotes .....	21
1.6 ADP-glucose pyrophosphorylase: the regulatory enzyme in bacterial and plant $\alpha$ -1,4-polysaccharides biosynthesis .....	22
1.6.1 Quaternary structure .....	22
1.6.2 Regulation .....	24
1.6.3 Kinetic mechanism of action .....	29
1.6.4 Structure-function relationship: background and motivation of this thesis work .....	32
1.6.4.1 Important residues and domains .....	35
1.6.4.2 Three-dimensional structure .....	37
<b>Chapter 2:</b>	
<b>Domain organization of the ADP-glucose pyrophosphorylase from <i>Escherichia coli</i></b> .....	40-64
2.1 Abstract .....	41
2.2 Introduction .....	42
2.3 Materials and methods .....	43
2.3.1 Materials .....	43
2.3.2 Plasmids and expression vectors .....	43
2.3.3 Linker-scanning mutagenesis .....	43
2.3.4 Truncated protein constructs .....	44
2.3.5 Construction of a plasmid for expression of the C-terminus alone (Ec <sub>328-431</sub> ) .....	44
2.3.6 Bacterial strains and expression of the recombinant enzymes ..	45
2.3.7 Protein methods .....	45
2.3.8 Enzymatic assays .....	47
2.3.9 Purification of Ec-ins117 .....	47
2.3.10 Purification of Ec <sub>1-323</sub> .....	48
2.3.11 Purification of co-expressed Ec <sub>1-323</sub> + Ec <sub>328-431</sub> .....	50
2.3.12 Molecular mass determination .....	50

2.3.13	Kinetic characterization .....	51
2.4	Results and discussion .....	52
2.4.1	Purification and characterization of Ec-ins 117 .....	52
2.4.2	Purification and characterization of C-terminal truncated forms .....	58
2.4.3	Characterization of Ec <sub>1-323</sub> + Ec <sub>328-431</sub> .....	59
2.5	Conclusions .....	62

### **Chapter 3:**

#### **The structural role of the *Escherichia coli* ADP-glucose pyrophosphorylase's N-terminus in allosteric regulation .....**

		65-85
3.1	Abstract .....	66
3.2	Introduction .....	67
3.3	Materials and methods .....	69
3.3.1	Construction of N-terminal truncated enzymes .....	69
3.3.2	Expression of the recombinant enzymes .....	70
3.3.3	Purification of N-truncated ADP-Glc PPase mutants .....	72
3.3.4	Purification of co-expressed EcNΔ15- CΔ108+ Ec <sub>328-431</sub> .....	72
3.3.5	Protein concentration assays .....	73
3.3.6	Protein electrophoresis and immunoblotting .....	73
3.3.7	Enzymatic activity assays .....	73
3.3.8	Kinetic characterization .....	74
3.4	Results and discussion .....	75
3.4.1	Purification and characterization of <i>E. coli</i> ADP-Glc PPase N-terminal truncated forms .....	75
3.4.2	Characterization of EcNΔ15- CΔ108+ Ec <sub>328-431</sub> .....	81

### **Chapter 4**

#### **Molecular architecture of the glucose 1-phosphate site in ADP-glucose pyrophosphorylases .....**

		86-129
4.1	Abstract .....	87
4.2	Introduction .....	88
4.3	Materials and methods .....	90
4.3.1	Materials .....	90
4.3.2	Homology modeling .....	90
4.3.3	Multiple sequence alignment .....	91
4.3.4	Site directed mutagenesis .....	91
4.3.5	Bacterial strains and expression of the recombinant ADPGlc PPases .....	92
4.3.6	Purification of pMAB3-single mutants .....	92
4.3.7	Protein methods .....	93
4.3.8	Enzymatic assays .....	95
4.3.9	Kinetic characterization .....	95
4.3.10	Thermal stability .....	96
4.4	Results .....	97
4.4.1	Homology modeling .....	97

4.4.2	Selection of residues for analysis .....	101
4.4.3	Expression and purification of pMAB3-single mutants .....	105
4.4.4	Kinetic characterization .....	105
4.4.5	Thermal stability .....	113
4.5	Discussion .....	115
 <b>Chapter 5</b>		
	<b>Conclusions</b> .....	130-156
5.1	Structural and functional role of the N- and C-termini in allosteric regulation .....	133
5.2	Structural basis of substrate binding .....	141
5.3	Structural basis of the catalytic mechanism .....	150
5.4	Future directions .....	154
	 <b>References</b> .....	 157-182

## LIST OF TABLES

<b>Table 1.1.</b> Carbon metabolism and structural and regulatory properties of ADP-Glc PPases from different .....	25
<b>Table 2.1.</b> Specific activities of mutant and wild-type <i>E. coli</i> ADP-Glc PPases ....	49
<b>Table 2.2.</b> Apparent affinity for the different effectors of the mutant Ec-ins117 and wild-type ADP-Glc PPase .....	57
<b>Table 2.3.</b> Kinetic parameters in the synthesis direction of <i>E. coli</i> ADP-Glc PPase wild type and Ec <sub>1-323</sub> + Ec <sub>328-431</sub> form .....	61
<b>Table 3.1.</b> Specific activities of wild-type and N-terminal truncated <i>E. coli</i> ADP-Glc PPases .....	76
<b>Table 3.2.</b> Specific activities of wild-type and mutants <i>E. coli</i> ADP-Glc PPases ...	82
<b>Table 3.3.</b> Kinetic parameters of wild-type and mutants <i>E. coli</i> ADP-Glc PPases .	83
<b>Table 4.1.</b> Complementary oligonucleotides used to introduce single mutations ...	94
<b>Table 4.2.</b> Comparison of specific activities and apparent affinity for the substrate Glc1P of the <i>E. coli</i> wild type and mutant enzymes .....	106
<b>Table 4.3.</b> Kinetic parameters of the <i>E. coli</i> wild type and mutant ADP-Glc PPases .....	110
<b>Table 4.4.</b> Thermal stability of the wild type and single mutants .....	114

## LIST OF FIGURES

<b>Figure 1.1.</b> Schematic organization of $\alpha$ -1,4-polyglucans structure .....	5
<b>Figure 1.2.</b> Metabolic routes involved in the biosynthesis of $\alpha$ -1,4-polysaccharides in different systems .....	12
<b>Figure 1.3.</b> Schemes of the sequential ordered Bi Bi kinetic mechanism proposed for ADP-Glc PPases .....	30
<b>Figure 1.4.</b> Three-dimensional structures of NDP-sugar PPases monomers .....	34
<b>Figure 2.1.</b> <i>E. coli</i> ADP-Glc PPase constructs .....	46
<b>Figure 2.2.</b> SDS-PAGE of <i>E. coli</i> wild-type ADP-Glc PPase and Ec-ins117 mutant .....	53
<b>Figure 2.3.</b> N-terminal sequence analysis of the Ec-ins117 mutant .....	55
<b>Figure 2.4.</b> Sequence of the insertion 117 in the <i>E. coli</i> ADP-Glc PPase .....	56
<b>Figure 2.5.</b> Domain organization of the ADP-Glc PPase from <i>E. coli</i> .....	64
<b>Figure 3.1.</b> Summary of the <i>E. coli</i> ADP-Glc PPase constructs studied and their activation properties .....	71
<b>Figure 3.2.</b> N-terminal <i>E. coli</i> ADP-Glc PPase sequence and predicted secondary structures .....	77
<b>Figure 3.3.</b> FBP-activation of wild-type and the N-terminal truncated <i>E. coli</i> ADP-Glc PPases .....	79
<b>Figure 3.4.</b> AMP inhibition kinetics of wild-type and mutants <i>E. coli</i> ADP-Glc PPases .....	80
<b>Figure 3.5.</b> FBP activation kinetics of wild-type and mutants <i>E. coli</i> ADP-Glc PPases .....	84
<b>Figure 4.1.</b> Structural model of <i>E. coli</i> ADP-Glc PPase .....	98
<b>Figure 4.2.</b> <i>E. coli</i> ADP-Glc PPase- substrate interaction .....	102
<b>Figure 4.3.</b> Sequence alignment of the <i>E. coli</i> ADP-Glc PPase and its homologues .....	103

<b>Figure 4.4.</b> Steady-state kinetic measurement for Glc1 <i>P</i> dependence for wild type and E194A, E194D and E194Q mutant enzymes .....	107
<b>Figure 4.5.</b> Superposition of the amino acids in the Glc1 <i>P</i> site from three NDP-Glc PPases .....	118
<b>Figure 4.6.</b> Hydrogen bond network involving Ser <sup>212</sup> in the <i>E. coli</i> ADP-Glc PPase Glc1 <i>P</i> site .....	121
<b>Figure 4.7.</b> Stick representation of the Ser <sup>212</sup> mutants' modeled active sites .....	122
<b>Figure 5.1.</b> Three-dimensional structure of the potato tuber ADP-Glc PPase .....	135
<b>Figure 5.2.</b> Three-dimensional model of <i>E. coli</i> ADP-Glc PPase monomer .....	139
<b>Figure 5.3.</b> Overlap of the three <i>E. coli</i> ADP-Glc PPase models: non-ligand-, ATP- and ADP-Glc-bound forms .....	142
<b>Figure 5.4.</b> Conformational changes in ADP-Glc PPase active site upon ligand binding .....	145
<b>Figure 5.5.</b> Conformational changes experienced by Glu <sup>194</sup> and Lys <sup>195</sup> upon ligand binding .....	148
<b>Figure 5.6.</b> Chemical reaction catalyzed by the ADP-Glc PPase .....	151

**Images in this dissertation are presented in color.**

## LIST OF ABBREVIATIONS

**3-PGA:** 3-phosphoglyceric acid

**ADP-Glc PPase:** adenosine 5'-diphosphoglucose pyrophosphorylase

**ADP-Glc:** adenosine 5'-diphosphoglucose

**BSA:** bovine serum albumin

**CDP-Glc PPase:** cytidine diphosphoglucose pyrophosphorylase

**CrATP:** chromium-ATP

**DHAP:** dihydroxyacetone phosphate

**dTDP-Glc PPase:** deoxythymidine diphosphoglucose pyrophosphorylase

**EDTA:** ethylenediaminetetraacetic acid

**F6P:** fructose 6-phosphate

**FBP:** fructose 1,6-bisphosphate

**Gal1P:** galactose 1-phosphate

**Glc1P:** glucose 1-phosphate

**Glc6P:** glucose 6-phosphate

**IPTG:** isopropyl- $\beta$ -D-thiogalactoside

**NDP-sugar PPase:** nucleotide-diphospho-sugar pyrophosphorylases

**PEP:** phosphoenolpyruvate

**P<sub>i</sub>:** inorganic phosphate

**PLP:** pyridoxal 5'-phosphate

**PP<sub>i</sub>:** inorganic pyrophosphate

**SDS:** sodium dodecyl sulfate

**UDP-Glc PPase:** urididine 5'-diphosphoglucose

**TCA:** tricarboxylic acid

# **CHAPTER 1**

## **Literature review**

## 1.1. INTRODUCTION TO POLYSACCHARIDES

Carbohydrates, proteins, lipids and nucleic acids are the four major classes of biomolecules in nature. Carbohydrates or saccharides (Greek: *sakcharon*, sugar) account for the most abundant group of compounds on earth and they play different roles, serving as energy storage, metabolic fuels, main intermediates in the various carbon metabolism routes, as well as structural components of the cells. These biomolecules are present in all living organisms as mono-, di-, oligo- and polysaccharides, and their structural units have multiple hydroxyl groups and are linked by *O*-glycosidic bonds (1, 2).

Monosaccharides are monomeric units of polyhydroxyaldehydes or polyhydroxyketones. The hexoses D-glucose and D-fructose are the most important of these simple sugars, whereas phosphorylated sugars like glucose 1-phosphate (Glc1P), glucose 6-phosphate (Glc6P), 3-phosphoglyceric acid (3-PGA) and dihydroxyacetone phosphate (DHAP), are another class of monosaccharide essential for the cellular metabolism. Reactive forms of monosaccharides, such as ADP-glucose (ADP-Glc) and UDP-glucose (UDP-Glc), known as sugar nucleotides, are also needed for the formation of *O*- and *N*-glycosidic linkages. Disaccharides consist of two monomeric sugars joined by an *O*-glycosidic bond. Sucrose (glucose- $\alpha$ -1,2- $\beta$ -fructose), maltose (glucose- $\alpha$ -1,4-glucose) and lactose (galactose- $\beta$ -1,4-glucose) are the three most abundant examples. Oligosaccharides are relatively short polymers containing a small number of monosaccharides (typically three to six). They are generally found either *O*- or *N*-linked to amino acid side chains and lipids and participate in molecular targeting and cell-cell

recognition, among other roles. Polysaccharides, which are also known as glycans, are large linear or branched molecules consisting of a variable number and type of sugar monomers linked in specific manners. Based on their chemical composition they can be classified in homopolysaccharides, with one type of component monomeric unit, and heteropolysaccharides, which are formed by more than one sort of monosaccharide. In addition, polysaccharides can have either a structural role or serve as an energy source or as carbon storage compounds. Examples of homopolysaccharides are glycogen and starch (the most abundant cellular storage forms) and cellulose and chitin (the principal structural components of the plant cell wall and invertebrate exoskeleton, respectively). On the other hand, among heteropolysaccharides, hyaluronic acid, which is present in connective tissues, is composed by units of D-glucuronic acid and *N*-acetyl-D-glucosamine, whereas the bacterial cell wall component peptidoglycan (or murein) consists of repetitive molecules of *N*-acetylmuramic acid and *N*-acetyl-D-glucosamine (1, 2).

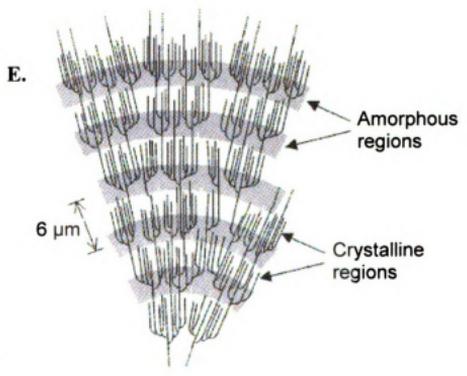
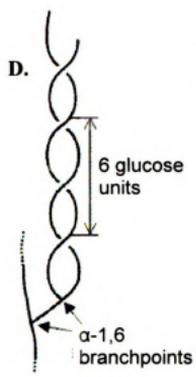
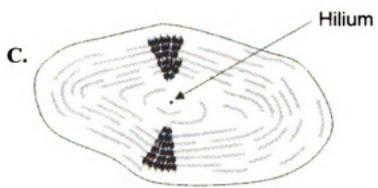
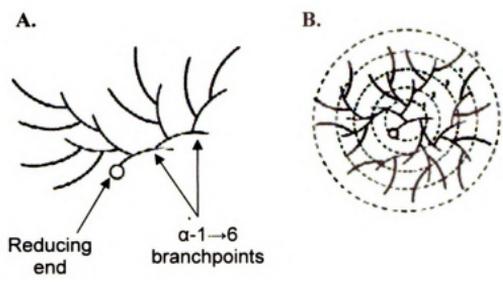
## **1.2. STORAGE $\alpha$ -1,4-POLYSACCHARIDES: STRUCTURE AND FUNCTION**

Alpha-1,4-polyglucans are one of the main strategies developed by living organisms to store carbon and energy in a readily available manner to cope with starvation conditions temporarily present in the environment (3-9). The biosynthetic process consists of the polymerization of large amounts of D-glucose surplus to form glycans of high molecular weight. One of the principal advantages of using polysaccharides as reserve compounds is that, due to their physiochemical properties they

produce negligible osmotic pressure, thus allowing for the accumulation of large amounts of carbon without disturbing the water relations within the cell (3, 5, 7-10). Another advantage is that the large amount of glycosidic linkages, present in amylopectin and glycogen as  $\alpha(1\rightarrow4)$  and  $\alpha(1\rightarrow6)$  bonds, generates a particular structural organization that allows for the access of degrading enzymes at the non-reducing ends and at internal linkages (Fig. 1.1). Enzymes such as amylases and phosphorylases can act simultaneously at different branches, thus speeding the conversion of large polymers into single units of D-glucose derivatives (7, 11). The use of  $\alpha$ -1,4-polyglucans as reserves is broadly seen in nature, including the starch in photosynthetic eukaryotes and the glycogen present in bacteria, (including cyanobacteria), archea and non-photosynthetic eukaryotes.

Glycogen is an essentially homogeneous water-soluble polymer of high molecular weight ( $10^7$ - $10^9$ ) containing an average of 10-14  $\alpha(1\rightarrow4)$ -linked D-glucose units, interlinked by  $\alpha(1\rightarrow6)$ -glycosidic linkages, to form a highly branched compound (Fig. 1.1.A) (5, 12, 13). The molecule has a spherical shape, with the chains organized in concentric tiers (Fig. 1.1.B) (14). The structure of the bacterial glycogen has been less studied than that of the mammalian one, but the available data point out the general similarity between them (13).

**Figure 1.1. Schematic organization of  $\alpha$ -1,4-polyglucans structure.** Drawing of glycogen showing: **(A)** the homogeneous structure and **(B)** its spherical shape formed by concentric tiers (marked by dashed lines) (13). The branches in the most external tiers are the most readily available for hydrolysis. Starch accumulates in granules **(C)** as a result of the growing helices of amylopectin **(D)** which determine regions of crystalline and amorphous starch **(E)**, viewed as concentric rings (M. F. Chaplin, Starch, [<http://www.lsbu.ac.uk/water/hysta.html>] adapted from T. P Coultate, Food: The Chemistry of its Components, 4th Ed. Accessed on 13 Nov 2006).



In mammals, the major deposits of glycogen are in liver and skeletal muscle but many cell types are capable of glycogen synthesis and limited accumulation, like cardiac and smooth muscle, the kidney, the brain and even adipose tissue (11). Its accumulation and utilization are hormonally linked to the nutritional status. Glycogen synthesis in the liver occurs after food intake as a consequence of the increased glucose level in blood which stimulates the secretion of the pancreatic hormone insulin, whereas decreased levels of glucose in the blood stream stimulate secretion of another pancreatic hormone, glucagon, which promotes the breakdown of the hepatic glycogen and its conversion to glucose by gluconeogenesis (1, 2). Therefore, the synthesis and degradation of this polysaccharide in liver determine the blood glucose homeostasis which is important for providing the different tissues with energy supply.

In bacteria, the precise role that glycogen may play is still unclear; however its presence is hypothesized to preserve cellular components susceptible to turnover in stationary phase for production of energy during starvation (15). These microorganisms usually accumulate glycogen as a result of limited growth conditions in the presence of an excess source of carbon, generally when they reach the stationary phase of growth (16). Exceptions to this are *Streptococcus mitis* (17), *Mycobacterium smegmatis* (18) and *Rhodobacter capsulata* (19) that have been reported to do it during the exponential phase of growth, and *Bacillus subtilis* (20) and *Streptomyces coelicolor* (21) that accumulate it prior to the onset or early in sporulation to supply the necessary resources for differentiation. In *Salmonella enteritidis*, glycogen storage was described as related to biofilm formation and virulence of this strain (22), whereas in *Streptococcus mutans*

accumulation of the polysaccharide was shown to be associated with cariogenic capacity of the bacterium (23).

Higher plants and algae accumulate starch (9). It is composed of two  $\alpha$ -1,4-polyglucans: amylose and amylopectin. Amylose is referred to as a smaller (about 1500 D-glucose residues), mainly linear  $\alpha(1\rightarrow4)$ -linked polysaccharide. Amylopectin is the major component of starch and, like glycogen, is an  $\alpha(1\rightarrow6)$ -branched  $\alpha(1\rightarrow4)$ -polyglucan (Fig. 1.1.D), although the degree of branching is only 5-6% (compared to the glycogen 8-12%) (7, 24). Starch is organized in concentric growing rings around a central hilum to form large granules that range from 0.1  $\mu\text{m}$  to 150  $\mu\text{m}$  in diameter, depending on the species (Fig. 1.1.C) (7, 12). Starch granules are microcrystalline structures of complex organization, with different polymorphisms depending on the source (7, 24-26). Regions of amorphous and crystalline starch that are formed by the alignment of helical external chains of the amylopectin molecule alternate in the granule (Fig. 1.1.E). Starch granules are found in almost all green plants in various types of plant tissues and organs (leaves, roots, shoots, fruits, grains, and stems) (7). Accumulation of starch in leaves occurs in the chloroplast due to carbon fixation during photosynthesis, a phenomenon that was demonstrated by Sachs in the XIX century (27). Exposure to low light or to extended periods of dark (24-48 h) causes disappearance of starch from photosynthetic tissues, as demonstrated by iodine staining (28) and light or electron microscopy (29). Starch is degraded to products used mainly for sucrose synthesis, which serves as carbon supply for sink tissues. In stomatal guard cells, however, starch is synthesized in dark periods and degraded during the day while stomata are open (30-32). In cereals, starch is

mainly synthesized and accumulated within the amyloplasts present in the endosperm (7). Deposits of this polysaccharide in potato tuber, maize and cassava, yam and sweet potato roots can range between 65% and 90% of their dry matter. In storage organs, such as fruits or seeds, accumulation of starch occurs during development or maturation of the tissue (7); whereas degradation takes place at the time of fruit ripening and seed or tuber germination, generating metabolites used as sources of carbon and energy.

### **1.3. ELUCIDATION OF THE $\alpha$ -1,4-POLYGLUCAN METABOLIC ROUTES: HISTORICAL BACKGROUND**

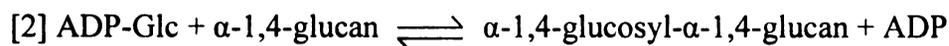
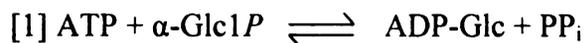
By the late 1930s, it was assumed that the synthesis and degradation of the mammalian glycogen proceed by a common metabolic route. This was based on observations made by Carl and Gerti Cori who obtained the first *in vitro* synthesis of a polyglucan from Glc1P by the glycogen phosphorylase, the enzyme responsible for the mammalian glycogen degradation (33). It was only one decade later that a different biosynthetic pathway was postulated after observation of large deposits of glycogen in the skeletal muscle of patients with the “Mc Ardle’s disease”, caused by a glycogen phosphorylase B deficiency (34). Those patients were able to synthesize glycogen but not to degrade it, leading to the conclusion that *in vivo* the synthesis and degradation of the polyglucan occur through separate metabolic routes. What the Coris had observed *in vitro* was merely the reversible reaction of the glycogen phosphorylase.

The elucidation of the metabolic route for polysaccharide accumulation begun in the late 1940s, when L.F. Leloir and his collaborators in Buenos Aires, Argentina, discovered that the UDP-Glc had another role besides being a cofactor for the transformation of the galactose 1-phosphate (Gal1P) into Glc1P (35-37). Cellular fractions from several prokaryotic and eukaryotic sources were reported as able to catalyze glycosidic transferences from UDP-Glc, and other sugar nucleotides that work as activated forms of glucose, to their respective acceptors. First, the incorporation of UDP-Glc in starch from plant extracts was reported (38); however, later studies on specificity for synthetic sugar-nucleotides showed that ADP-Glc was a better substrate for the synthesis of starch in higher plants (39). These results led to the isolation of ADP-Glc from maize extracts (40) and to the first identification of an enzyme involved in ADP-Glc synthesis, the ADP-glucose pyrophosphorylase (41).

Studies performed with mammalian liver extracts clearly determined that the glucosyl donor for glycogen synthesis was UDP-Glc (42). On the other hand, the metabolic reactions involved in the synthesis of bacterial glycogen were elucidated after Sigal and collaborators reported UDP-Glc synthesis deficient *Escherichia coli* strains that accumulated normal glycogen levels (43). The finding of large amounts of ADP-Glc-specific glycogen synthase and ADP-Glc pyrophosphorylase in those microorganisms ruled out that UDP-Glc was an important precursor of glycogen synthesis in enterobacteria and strongly suggested that ADP-Glc was the main glucosyl donor for polysaccharide biosynthesis in bacteria and plants (7).

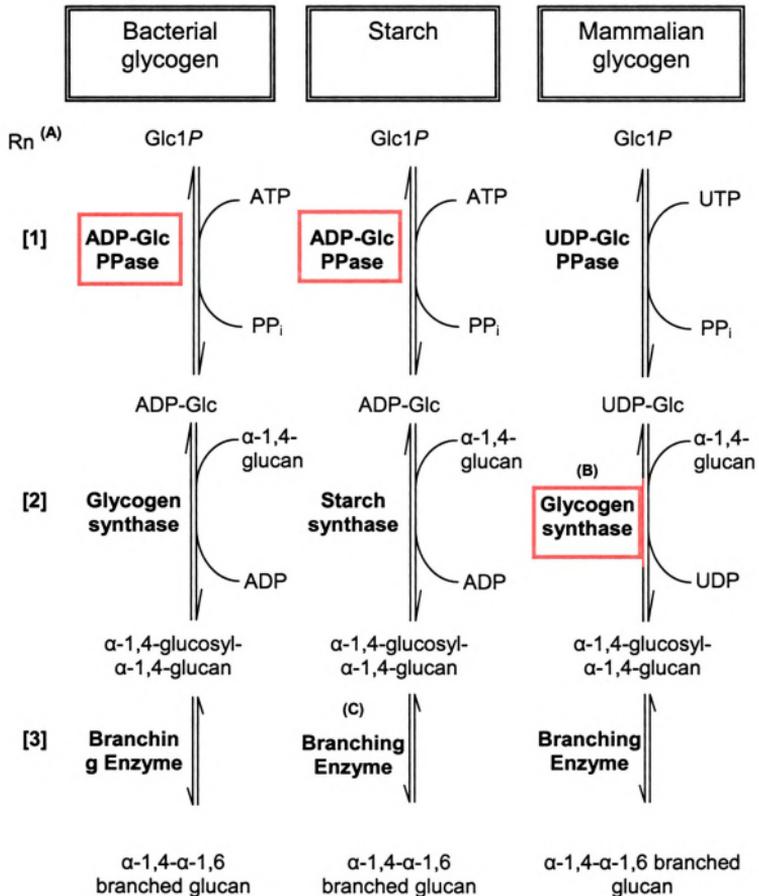
## 1.4. METABOLIC PATHWAYS FOR $\alpha$ -1,4-POLYSACCHARIDES SYNTHESIS IN BACTERIA AND PLANTS

The synthesis of the  $\alpha$ -1,4 linked,  $\alpha$ -1,6 branched polysaccharides can be divided in the general following steps: [1] the synthesis of the glucosyl donor molecule from the sugar-phosphate, [2] the transference of a single glucosyl moiety from the sugar nucleotide to the glycan synthesis primer, and [3] the polysaccharide structure remodeling. The biosynthesis of bacterial glycogen and plant starch occurs by utilizing ADP-Glc as the glucosyl donor. The reactions taking place are as follows:



The enzymes catalyzing the different steps, [1] ADP-glucose pyrophosphorylase, [2] glycogen and starch synthase and [3] branching enzymes, are described in sections 1.4.1, 1.4.2 and 1.4.3 and a more complete analysis of the biochemical characterization of the ADP-glucose pyrophosphorylases is presented in section 1.6. A comparison between the biosynthesis of  $\alpha$ -1,4-polysaccharides in bacteria, plants and mammals is depicted in Fig. 1.2

**Figure 1.2. Metabolic routes involved in the biosynthesis of  $\alpha$ -1,4-polysaccharides in different systems.** ADP-Glc PPase and UDP-Glc PPase catalyze reaction [1], glycogen and starch synthases catalyze reaction [2] and branching enzyme catalyze reaction [3], described in section 1.4. Bacterial glycogen and starch utilize ADP-Glc as the activated form of glucose. **(A)** Rn: Reaction. **(B)** *De novo* reaction is:  $\text{UDP-Glc} + \text{apo-glycogenin} \rightleftharpoons (\text{glycosyl})\text{glycogenin} + \text{UDP}$ ; further glucose-residues incorporation occurs as depicted in reaction [2] in the diagram. **(C)** In amylopectin synthesis. Enzymes in the red boxes catalyze the rate limiting steps in the respective pathways and are tightly regulated.



### 1.4.1. ADP-Glc pyrophosphorylase

Bacterial glycogen and starch biosynthesis route share a common first reaction (Reaction [1]) catalyzed by the ADP-Glc pyrophosphorylase (ATP:  $\alpha$ -D-glucose 1-phosphate adenylyltransferase, E.C. 2.7.7.27; ADP-Glc PPase). This enzyme is coded by a single gene in most bacteria except those from the genus *Bacillus* (*glgC* in *E. coli*). But in plants, at least two genes are found (8, 9). The reaction was first reported by Espada in 1962 (41) and, subsequently, Preiss and collaborators described its requirement of a divalent metal ion,  $Mg^{2+}$  or  $Mn^{2+}$ , as well as its *in vitro* reversibility, with equilibrium close to 1 (44). *In vivo*, however, the presence of inorganic pyrophosphatases in the cell and the utilization of ADP-Glc for polysaccharide synthesis drive the equilibrium towards the synthesis of the sugar-nucleotide (3). In addition, the ADP-Glc PPases from many systems were found to be highly regulated by allosteric metabolic intermediates (4, 7-9, 16, 45-47).

The physiological role of ADP-Glc PPase as the regulatory enzyme in the bacterial glycogen and plant starch biosynthetic pathways is supported by solid experimental evidence. Mutants of *E. coli* and *Salmonella enterica* serovar Typhimurium with affected capacity of glycogen accumulation after chemical mutagenesis, are the basis of the previous statement. These mutant strains expressed ADP-Glc PPases with altered regulatory properties and the point mutations found through gene sequencing were the subject of further biochemical characterization. For instance, the *E. coli* B mutants that accumulated glycogen at a faster rate than their parent strain expressed a

mutated ADP-Glc PPase with higher apparent affinity for the activator fructose 1,6-bisphosphate (FBP) and lower affinity for the inhibitor, AMP (48-50). On the other hand, *E. coli*-glycogen deficient mutants, that synthesized glycogen at 40% the rate of wild type, had a lower affinity for the activator, FBP (51, 52).

A similar situation was observed in oxygenic photosynthetic organisms, where the main activator and inhibitor are 3-PGA and inorganic orthophosphate ( $P_i$ ). Starch deficient mutants of the unicellular algae, *Chlamydomonas reinhardtii*, expressed an ADP-Glc PPase insensitive to 3-PGA activation and  $P_i$  inhibition (53). Work with *Arabidopsis* (54-57), maize (58), wheat (59) and potato (60) mutants also supported the key regulatory role of ADP-Glc PPase in starch biosynthesis. More recently, metabolic control analysis in potato tuber showed that the enzyme catalyzes a near rate-limiting step in the pathway of starch synthesis (61). Therefore, because of the demonstrated regulatory role of ADP-Glc PPase in starch biosynthesis *in vivo*, several studies have been aimed to increase the polysaccharide yield and overall plant productivity by genetic manipulation of this enzyme's activity (58-60, 62, 63).

Overall, the finding that regulation in bacteria and plants occurs at the ADP-Glc synthesis step agrees with the concept that effective regulation of a biosynthetic pathway occurs at its first unique step.

In section 1.6., the particular regulatory properties of the ADP-Glc PPases from different sources and information on their functional and structural characterization known to date are presented.

#### **1.4.2. Bacterial glycogen / starch synthase**

The second step in the synthesis of bacterial glycogen and plant starch (Reaction [2]) consists in the transfer of the activated glucose to the non-reducing end of a growing  $\alpha$ -1,4-linked glucan. Bacterial glycogen synthases (ADP-Glc: $\alpha$ -1,4-glucan, 4-glucosyltransferase; E.C. 2.4.1.21; GS) and plant starch synthases (ADP-Glc: $\alpha$ -1,4-glucan, 4-glucosyltransferase; E.C. 2.4.1.21; SS) catalyze this chemical reaction. In bacteria (e.g., *E. coli*) only one GS and GS gene (*glgA*) have been found (64).

Both GS and SS are non-regulated enzymes that use exclusively ADP-Glc as a glucosyl donor, have molecular masses of 48-55 kDa and organize in dimers in the native form (65-68). The sequence identity shared by the bacterial GS and the plant and algae SS is between 30 and 36%, and it has been suggested that they could have a common three-dimensional fold and catalytic mechanism. Both are glycosyltransferases that operate retaining the anomeric configuration of the glucose transferred from ADP-Glc to the non-reducing end of glycogen and amylose or amylopectin, respectively.

Secondary structure predictions and threading techniques postulated that GS and SS have a glycosyltransferase-type B (GT-B) fold consisting of two distinctive

Rossmann-like- $\beta$ - $\alpha$ - $\beta$  domains separated by a big cleft (69, 70) and in more recent reports the *E. coli* GS three-dimensional structure was modeled from other related GT-B glycosyltransferases. The model obtained was in its active conformation (“closed” form) and a structure-guided site directed mutagenesis of conserved residues present in the active site, which is formed in the cleft with amino acids belonging to both domains, and a complete biochemical characterization of the mutants validated it (71, 72). Later, the reported crystal structures of the *A. tumefaciens* GS (73) and, more recently, that of the archaeal GS from *Pyrococcus abyssi* (74), agreed with the previous results, although the structures obtained by crystallography were in the “open” inactive state which complicated the identification of the studied amino acids. For a number of enzymes of the GT-B superfamily it has been proposed that the inactive state undergoes a substrate-triggered closing of the inter-domain cleft, thereby bringing together the catalytic residues that make up a competent active center.

In higher plant tissues, in contrast to the bacterial GS, more than one SS gene has been reported (75-78). Two groups of starch synthases have been described: one is that bound to the starch granule (granule bound starch synthase: GBSS) which can only be solubilized by  $\alpha$ -amylase digestion of the granule; the second one is found in the soluble fraction of the plant extracts and, therefore called, soluble starch synthases (SSS) (7).

Work on several plant systems has shown that multiple forms of SSS and GBSS are present, and the subject has been extensively reviewed by Sivak and Preiss (7). For example, two major types of SSS (SSS I and SSS II) have been described and

characterized in maize (79). These two forms seemed to be distinct on the basis of their physical, kinetic and immunological properties, and thus proposed to be products of two different genes with different roles in starch granule formation. To date, SSS mutant or SSS-knockout plants have been reported to display structural modifications of amylopectin associated with increases in amylose content, suggesting a role for SSS in amylopectin biosynthesis (80-83). In addition, two types of GBSS have been described in maize: GBSS I and GBSS II (77). RNA antisense experiments in potato led to the suggestion that these enzymes would be related to the synthesis of amylose fraction of starch (84), although the precise molecular mechanism is still under debate (85). On the other hand, mutation in the gene coding for the GBSS (*waxy*) in *C. reinhardtii* had been suggested to have a role in amylopectin synthesis (86).

Molecular characterization of maize SSS forms have identified important residues for catalysis and ADP-Glc binding (87-89). More recent reports describe the identification of critical residues determining the differential activity of SSS isoforms in rice plants (90). The varied activity of these SSS isoforms induced structural changes in amylopectin determining some of the gelatinization and pasting properties of starch granules (90, 91). The identified residues are close to a conserved motif first recognized as critical for enzymatic activity in the *E. coli* GS by Yep and collaborators (72) and later observed in the *A. tumefaciens* GS crystalline structure (73). Despite these findings, studies aimed to determine the relationship between structure and function of the multiple isoforms of SS still lag behind the other enzymes involved in starch biosynthesis.

### 1.4.3. Branching enzyme

The last step in the biosynthesis of glycogen and amylopectin (Reaction [3]) is catalyzed by the branching enzyme (E.C. 2.4.1.18;  $\alpha$ -1,4- $\alpha$ -D-glucan:  $\alpha$ -1,4- $\alpha$ -D-glucan 6-glycosyl-transferase; BE), and is responsible for the synthesis of the  $\alpha$ -1,6 linkages seen in these polysaccharides. BE and starch BE (SBE) catalyze two reactions in synthesizing  $\alpha$ -1,6 glucosidic linkages: one is the cleavage of an  $\alpha$ -(1-4) glucosidic linkage and the next one is the transfer of the cleaved oligosaccharide to form a new  $\alpha$ -(1-6) branch (92). As observed with the previous enzymes in the pathway, in *E. coli* and other bacteria only one BE and only one gene (*glgB* in *E. coli*) are present (20, 93-99) but more than two genes have been described in higher plants (7).

Bacterial glycogen and amylopectin present different patterns of branching: the first has shorter branches (10-13 glucose residues) and the  $\alpha$ -1,6 linkages represent 10% of the total glucosidic linkages, whereas the second polysaccharide is formed by longer branches (20-24 residues) with only 5% of  $\alpha$ -1,6 linkages (7). It is thus expected that the corresponding enzymes present different specificity with respect to the branch transfer, or that the type of interaction between the BE and the GS in bacteria is different from that between the SS and the SBE.

While amylose synthesis requires only an active GBSS (introduced in the section 1.4.2.), amylopectin is synthesized by a complex group of enzymes involving, among others, several isoforms of SS, SBE, and starch-debranching enzymes [reviewed in (12)].

Potato contains two isoforms of SBE, SBEI and SBEII, and suppression of both isoforms is required to generate starch with amylose content of >50% and a complete lack of normal amylopectin (100). In maize, one SBEI and two forms of SBEII, SBEIIa and SBEIIb, are found. Suppression of SBEIIb results in the "amylose-extender" phenotype, with amylose contents from 50% to 90% (101), whereas suppression of SBEI or SBEIIa has no impact on endosperm amylose content (102, 103).

The relationship in amino acid sequences between that of BE and amylolytic enzymes in the " $\alpha$ -amylase family" of enzymes have been compared (7, 104). This family includes  $\alpha$ -amylase (E.C. 3.2.1.1), pullulanase (E.C. 3.2.1.41), isoamylase (E.C. 3.2.1.68), neopullulanase (E.C. 3.2.1.135), cyclodextrin glucanotransferase (E.C. 2.4.1.19). Among these enzymes four specific regions within putative catalytic sites are conserved, and they are also recognized in the SBE and the BE known sequences. The role of the amino acids located in these regions (such as Asp<sup>405</sup>, Glu<sup>458</sup>, and Asp<sup>526</sup> in the *E. coli* BE numbering) has been probed by site-specific mutagenesis (98, 105).

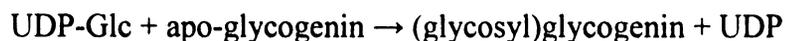
In addition, a structural conservation in the  $\alpha$ -amylase family has been proposed by the prediction the ( $\beta$ - $\alpha$ )8-barrel structural domain in the various groups of enzymes in the family (106). The recently reported three-dimensional structure of the *E. coli* BE (107) supports this view. However, despite the conservation of this central ( $\beta$ - $\alpha$ )8-barrel, the C-terminus and N-terminus portions of BE from various bacteria as well as from higher plants are dissimilar in sequence and in size (7). Chimeric constructs made from different isoforms of the maize endosperm SBE with different specificity and pattern of

branching of the substrate assigned a role to the N-terminal region in determining the size of oligosaccharide chain transferred (108). Indeed, a progressive shortening of the N-terminus led to a gradual increase in the length of the transferred chains, suggesting that this domain provides a support for the glucan substrate during the processes of cleavage and transfer of the  $\alpha$ -(1–4) glucan chains (109-111).

To date, several types of SBE have been reported to be useful for applications such as changing the solubility of starch (112-114) or producing large cyclic glucans (99, 115). Therefore, understanding how BE and SBE work through their molecular identification and characterization has also been important for the starch-processing industry.

## **1.5. Alpha-1,4-POLYSACCHARIDES IN NON-PHOTOSYNTHETIC EUKARYOTES**

Several studies performed in mammalian cells determined that the glucosyl donor obtained in the first reaction of the pathway (Reaction [1]) is UDP-Glc (116) (whereas in bacterial and photosynthetic eukaryotes is ADP-Glc). The reaction is catalyzed by the UDP-Glc PPase (E.C. 2.7.7.9). Also different is that *de novo* synthesis of glycogen is initiated by glycogenin (E.C. 2.4.1.186) (117, 118). This protein is self-glycosylated in a reaction as follows:



Following this reaction, a covalently bound oligosaccharide is formed by addition of glucose monomers linked by  $\alpha$ -1,4-linkages that serves as primer for the elongation catalyzed by the eukaryotic GS (E.C. 2.4.1.11) (Reaction [2]).

In contrast to what occurs in bacteria, the synthesis of mammalian glycogen utilizes UDP-Glc and the production of this sugar nucleotide is not a regulated step. In the mammalian pathway, the GS reaction is rate limiting (1, 2, 116). The mammalian GS is regulated by allosterism and by phosphorylation/ dephosphorylation events, catalyzed by protein kinases and phosphatases. On the other hand, the glycogen phosphorylase (E.C. 2.4.4.1) is also regulated by opposite phosphorylation/ dephosphorylation cycles, to guarantee that they would not be active at the same time as the glycogen synthases, thus avoiding futile cycles (1, 2). The balance between glycogen synthesis and degradation is intimately regulated by insulin and glucagon hormone levels. The hormone epinephrine has the same effect as glucagon but acts in muscular tissue rather than in liver, as glucagon does (1, 2).

## **1.6. ADP-GLUCOSE PYROPHOSPHORYLASE: THE REGULATORY ENZYME IN BACTERIAL AND PLANT $\alpha$ -1,4-POLYSACCHARIDES BIOSYNTHESIS**

### **1.6.1. Quaternary structure**

ADP-Glc PPase catalyzes the reaction converting Glc1P and ATP into ADP-Glc and PP<sub>i</sub> (41) in presence of a divalent cation (physiologically, Mg<sup>2+</sup>) (44). The native

enzyme consists of four subunits, being, depending on the source, a homotetramer ( $\alpha_4$ ) or a heterotetramer ( $\alpha_2\beta_2$ ).

Most of the bacterial enzymes described to date, except for those from the *Bacillus* genus, are homotetramers of 200 kDa and products of the expression of a single gene (Table 1.1) (3, 4, 7-9, 20, 47, 119-122).

In contrast, higher plants and green algae ADP-Glc PPases described to date are heterotetramers, composed by two types of subunits that differ in their molecular masses and amino acid sequence (3, 4, 6-9, 20, 123, 124). Although in some cases the difference in molecular masses between them is not more than 1 kDa (123, 125), these subunits were called “small” ( $\alpha$  subunit, 50-54 kDa) and “large” ( $\beta$  subunit, 51-60 kDa) subunits for convenience. The small subunit of higher plants is highly conserved (85-90% identity) and has a catalytic role, whereas the large subunit seemed to have diverged more (they present 50-60% identity) and have a modulatory role without catalytic function (126-128). Moreover, the most frequent situation in higher plants is the existence of one functional small subunit gene and several large subunit genes that are differentially expressed (7, 128-136). As reported in *Arabidopsis* (136, 137), the several large subunits are differentially expressed in the plant and their particular association with the one catalytic small subunit creates four types of heterotetramers with differential kinetic and regulatory properties according to the necessities for starch synthesis. The similarity between small and large subunits (~50% identity) suggests a common ancestry (128), and it was recently shown that both subunits have most likely derived from a common

catalytic ancestor, with the large subunit maintaining the architecture of the active site, including the substrates binding sites but exhibiting inability to perform catalysis due to mutations of few essential residues (127).

### **1.6.2. Regulation**

Most of the studied ADP-Glc PPases are allosterically regulated enzymes (8, 9). Allosteric modulators of the enzyme so far characterized differ from source to source but they share the common feature of being the key intermediates of the major pathway for carbon assimilation/utilization in that given cell. The activators are markers or signals of carbon and energy excess and the inhibitors represent indicators of low metabolic energy levels. Therefore, these regulatory characteristics, in addition to the fact that ATP is one of the substrates, agree with ADP-Glc PPase having maximal activity for storage polysaccharide synthesis when cellular energy and carbon are in excess.

Specificity for metabolites behaving as allosteric regulators served as the basis for the classification of the ADP-Glc PPases in different groups or “classes”, which has been updated as novel enzymes were being characterized (3, 6-9, 16, 138). Table 1.1. shows the most recent classification reported (8, 9) including the principal pathway of carbon assimilation in the respective cells, allosteric activators and inhibitors, and the quaternary structure of the enzyme.

**Table 1.1. Carbon metabolism and structural and regulatory properties of ADP-Glc PPases from different organisms (8, 9)**

Organism	Carbon metabolism	ADP-Glc PPase			Quaternary structure
		Class	Allosteric regulators		
			Activator(s)	Inhibitor(s)	
<b>Accumulating glycogen</b>					
<b>Prokaryotes</b>					
<i>Escherichia coli</i> , <i>Salmonella enterica</i> serovar Typhimurium, <i>Enterobacter aerogenes</i>	<i>Heterotrophic</i> Utilizing the Embden- Meyerhof pathway (glycolysis)	I	FBP	AMP	Homotrimer ( $\alpha_3$ )
<i>Aeromonas formicans</i> , <i>Micrococcus luteus</i> , <i>Mycobacterium smegmatis</i>	Utilizing the Embden- Meyerhof pathway (glycolysis)	II	FBP, F6P	AMP, ADP	
<i>Serratia Marcescens</i> , <i>Enterobacter hafniae</i> , <i>Clostridium pasteurianum</i>	Utilizing the Embden- Meyerhof pathway (glycolysis)	III	None	AMP	Homotrimer ( $\alpha_3$ )
<i>Agrobacterium tumefaciens</i> , <i>Athrobacter viscosus</i>	Utilizing the Entner-Doudoroff pathway	IV	Pyruvate, F6P	AMP, ADP, P <sub>i</sub>	Homotrimer ( $\alpha_3$ )
<i>Chromatium vinosum</i> , <i>Rhodobacter capsulata</i> , <i>Rhodomicrobium vannielii</i>	<i>Performing anoxygenic photosynthesis</i> Utilizing the Entner-Doudoroff pathway	IV	Pyruvate, F6P	AMP, ADP	
<i>Rhodobacter gelatinosa</i> , <i>Rhodobacter globiformis</i> , <i>Rhodobacter sphaeroides</i> , <i>Rhodocycclus purpureus</i> <i>Rhodospirillum rubrum</i> , <i>Rhodospirillum tenue</i>	Utilizing glycolysis and the Entner- Doudoroff pathways  Utilizing the Tricarboxylic acid and the reductive carboxylic acid cycles	V  VI	Pyruvate, F6P, FBP  Pyruvate	AMP, P <sub>i</sub>  None	Homotrimer ( $\alpha_3$ )

**Table 1.1.** Continued

Organism	Carbon metabolism	Class	ADP-Glc PPase			Quaternary structure
			Allosteric regulators		Inhibitor(s)	
			Activator(s)	Inhibitor(s)		
<i>Bacillus subtilis</i> <i>Bacillus stearothermophilus</i>	Utilizing the Tricarboxylic acid cycle during sporulation	VII	None	None	Heterotetramer ( $\alpha_2\beta_2$ )	
Cyanobacteria	<i>Performing oxygenic Photosynthesis</i>					
<i>Synechococcus</i> PCC 6301 <i>Synechocystis</i> PCC 6803 <i>Anabaena</i> PCC 7120	Fixing CO <sub>2</sub> through the Calvin cycle	VIII	3-PGA	P <sub>i</sub>	Homotetramer ( $\alpha_4$ )	
<b>Accumulating starch</b>						
Eukaryotes						
Green algae <i>Chlorella fusca</i> <i>Chlorella vulgaris</i> <i>Chlamydomonas reinhardtii</i>	Fixing CO <sub>2</sub> through the Calvin cycle	VIII	3-PGA	P <sub>i</sub>	Heterotetramer ( $\alpha_2\beta_2$ )	
Higher plants Photosynthetic tissues Leaves of spinach, wheat <i>Arabidopsis</i> , maize, rice	Fixing CO <sub>2</sub> through the Calvin cycle	VIII	3-PGA	P <sub>i</sub>	Heterotetramer ( $\alpha_2\beta_2$ )	
Non-photosynthetic tissues Potato tubers Endosperm of maize, barley and wheat	Heterotrophic cells Metabolizing sucrose imported from photosynthetic tissues	VIII IX	3-PGA None directly, 3-PGA and F6P reverse inhibitor's effect	P <sub>i</sub> P <sub>i</sub> , ADP, FBP	Heterotetramer ( $\alpha_2\beta_2$ ) Heterotetramer ( $\alpha_2\beta_2$ )	

Enzymes organized in *class I* are those from enteric bacteria, such as *E. coli* and *Salmonella enterica* serovar Thyphimurium. These microorganisms assimilate glucose via glycolysis (the Embden-Meyerhof pathway) which is regulated at the site of FBP (the phosphofructokinase step) (1, 2), and this metabolic intermediate is the main activator of the ADP-Glc PPase (44, 139). These enzymes are inhibited by AMP and the native forms are ~200 kDa homotetramers, products of a single gene (5, 8, 9). ADP-Glc PPases from other microorganisms performing glycolysis have been grouped in *classes II* or *III*, depending on whether they are activated by FBP and F6P, or if they are insensitive to any activator, respectively. Both groups of enzymes are inhibited by AMP and *class II* is also down-regulated by ADP (Table 1.1).

*Class IV* includes those bacteria catabolizing glucose to pyruvate by the Entner-Doudoroff pathway, such as *A. tumefaciens*. The ADP-Glc PPase in these organisms is sensitive to allosteric activation by, mainly, pyruvate and also F6P, and to inhibition by AMP and ADP (Table 1.1.) (19, 122). Interestingly, ADP-Glc PPases from organisms that can assimilate carbon by both the Embden-Meyerhoff and the Entner-Doudoroff pathways are regulated by the three main activators: Pyruvate, FBP and F6P, depending on the physiologic conditions (Table 1.1). They are also inhibited by AMP and P<sub>i</sub>, form a homotetramer which is the product of the expression of a single gene and have been grouped in *class V* (121, 140).

ADP-Glc PPases in *class VI* are from anaerobic bacteria, like *Rhodospirillum rubrum* and *R. tenue*, which are not able to catabolize glucose but grow on pyruvate, lactate or CO<sub>2</sub>. Pyruvate is the product of CO<sub>2</sub> fixation and it is the sole activator of the *R. rubrum* and *R. tenue* ADP-Glc PPases (141, 142). Grouped as *class VII* are the ADP-Glc PPases from sporulating *bacilli* (Table 1.1). These microorganisms synthesize glycogen during early stage of sporulation, a process for survival to hostile environments (120). Under the latter conditions, the tricarboxylic acids cycle (TCA) is the main pathway for carbon utilization that fully metabolizes by-products of glycolysis (143). The different recombinant enzymes from *B. stearothermophilus* (99) and *B. subtilis* (20) were insensitive to regulation by different metabolites typically affecting the activity of other bacterial ADPGlc PPases. It has been shown that in these organisms the native ADP-Glc PPase exhibits a heterotetrameric structure of the type  $\alpha_2\delta_2$  (Table 1.1) product of the expression of two different genes.

*Class VIII* and *IX* group ADP-Glc PPases from cyanobacteria, alga and higher plants (Table 1.1). These organisms photoassimilate atmospheric CO<sub>2</sub> utilizing the reductive pentose phosphate pathway that gives 3-PGA as the first intermediate product. Under light conditions, P<sub>i</sub> is used to regenerate ATP through photophosphorylation (7-9, 144). ADP-Glc PPases of *class VIII* (from cyanobacteria and photosynthetic tissues in higher plants and algae) are allosterically regulated by the ratio of 3-PGA (activating) and P<sub>i</sub> (inhibiting) (3, 145, 146). In non-photosynthetic tissues there are two types of ADP-Glc PPases. One is activated by 3-PGA and inhibited by P<sub>i</sub> (147, 148). The prokaryotic enzymes in this group are coded by a single gene, giving rise to a homotetramer, whereas

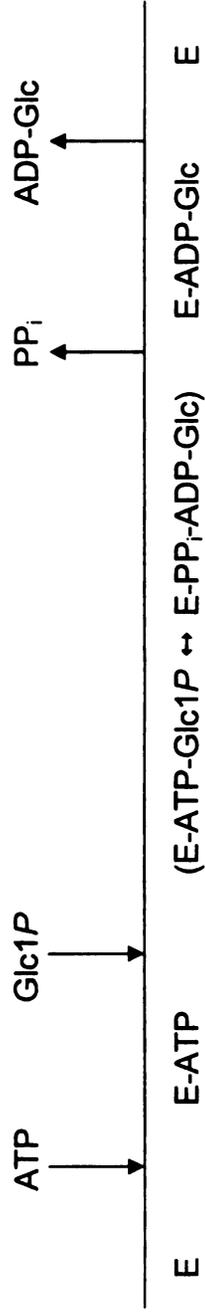
those from algae and higher plants, for instance the potato tuber enzyme, (from *classes VIII and IX*) are the expression products of two genes that originate a native heterotetramer (Table 1.1). A second type of ADP-Glc PPase present in non-photosynthetic tissues shows distinctive regulatory properties and for this reason they have been grouped in a different class (*IX*, Table 1.1). These enzymes revealed lower sensitivity to activators (*133, 149-153*). The activity of the *class IX* enzyme from wheat cannot be increased by allosteric activators, but 3-PGA and F6P can reverse the inhibition caused by  $P_i$ , ADP or FBP (*149*).

In addition, the enzymes from potato tuber and *Arabidopsis* leaves were shown to be regulated by a redox mechanism *in vitro* and *in vivo* (*147, 154-156*). It consists on the formation of an inhibitory disulfide bond between the two small subunits, involving an N-terminal cysteine (*154, 157*), which is not conserved in bacterial enzymes or in other plant ADP-Glc PPases.

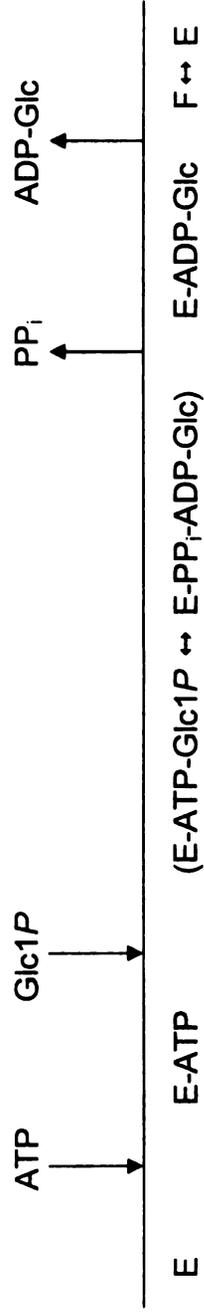
### **1.6.3. Kinetic mechanism of action**

The chemical reaction catalyzed by ADP-Glc PPase proceeds following an ordered sequential kinetic mechanism. It was shown first with the enzyme from *R. rubrum* enzyme by Paule and Preiss in the early 1970s (*158*). Later binding studies with the *E. coli* enzyme (*159*), supported the model showing that ATP and ADP-Glc bind to free forms of the enzyme (Fig. 1.2.A). These studies also pointed out the cooperative properties of this enzyme and the heterotropic interactions between substrates. For

A.



B.



**Figure 1.3. Schemes of the sequential ordered Bi Bi kinetic mechanism proposed for ADP-Glc PPases. A) In *R. rubrum* (158) and in *E. coli* (159); B) in barely leaves (160). Letters *F* and *E* refer to the free forms of the enzyme.**

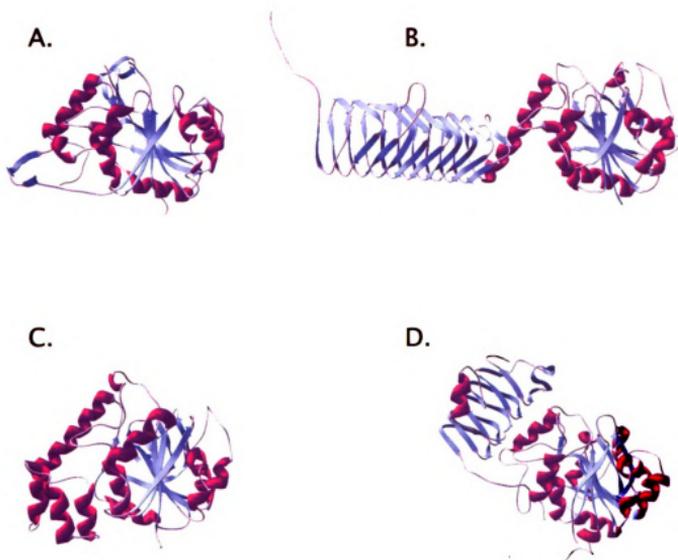
instance, they observed that ATP displayed half-sites occupancy in the absence of Glc1P, but full occupancy when this second substrate was present. Also, Glc1P was not able to bind the enzyme unless ATP and  $Mg^{2+}$  were present (159). Later, initial velocity kinetic studies performed on barley leaf ADP-Glc PPase (160) also agreed with the sequential ordered mechanism, but with ATP and ADP-Glc binding to different forms of free enzyme (Fig. 1.2.B). This isomerization phenomenon ( $F \leftrightarrow E$  in Fig. 1.2.B) may reflect different conformations and/or aggregation states of the heterotetrameric plant protein, not observed in the *R. rubrum* ADP-Glc PPase probably due to the less complex single-subunit-type structure of the bacterial enzyme when compared to the two-subunit-type structure of a higher plant enzyme (158).

The proposed sequential ordered Bi Bi type of mechanism also agrees with the one postulated for other nucleotidyl-diphosphosugar pyrophosphorylases (NDP-sugar PPases), such as the UDP-Glc PPase (161, 162), the UDP-*N*-acetylglucosamine pyrophosphorylase (GlmU) (163), the dTDP-Glc PPase (RmlA) (164, 165) and the CDP-Glc PPase (166). Very recent data obtained by saturation transfer difference (STD)-NMR studies on the UDP-Glc PPase from *Leishmania major* provided the first direct proof for the sequential ordered bi-bi mechanism suggested in earlier studies (161). All these reports point out to a common mechanism of action by which the reactions catalyzed by NDP-sugar PPases proceed. Therefore, it is possible that these enzymes share a common ancestor and, if so, it would be expected that they also share common structural features, like key catalytic residues or even similar structural fold.

#### **1.6.4. Structure-function relationship: Background and motivation of this thesis work**

Research on the molecular and biochemical characterization of the ADP-Glc PPases has radically changed over the last five to ten years. Before, structure-function relationship studies provided information in one dimension, that is, in the context of the known primary structure of the enzyme. For instance, determination of the quaternary organization of the ADP-Glc PPases from different sources has been based on the sequence (nucleotide or amino acid) analysis of the expressed gene products. Functional residues, important for catalysis and regulation, have been identified by a combination of chemical modification experiments, analysis of amino acidic sequence conservation and site directed mutagenesis approaches. But studies aimed to investigate the relationship between structure and function reached a new level of complexity and understanding when new approaches designed to elucidate the organization of a protein structure in two and three dimensions became available. First, the local secondary structures possibly adopted by the enzyme polypeptide could be predicted (8, 9, 167). Later, the evidence of specific interactions between certain regions in the ADP-Glc PPase allowed for drawing hypothesis on the structural organization of the protein but in another dimension (168). That line of work set the basis and inspired the investigation of the domain organization of this enzyme and the closer analysis of their functional roles which are presented in *Chapter 2* and *3* of this thesis.

Efforts directed to obtain the three-dimensional structure of the ADP-Glc PPase from *E. coli* by X-ray crystallization have been unsuccessful. During the last five years the three-dimensional structures of related NDP-sugar PPases (GlmU (163), RmlA (169) and CDP-Glc PPase (165) have been solved (Fig. 1.2.A, B, and C). This was certainly beneficial since new and more efficient computational methods to predict protein folds (like threading) and three-dimensional structures (such as the various approaches in modeling: by homology, *ab initio*, etc) were being developed in parallel. Therefore, the availability of structures of related enzymes and those predictive computational methods became valuable to draw new hypotheses and design biochemical experiments to test them. Indeed, the very recent elucidation of the first crystal structure of a plant ADP-Glc PPase small subunit (Fig. 1.3.D), reported in 2005, was of great advantage for the research in this field. Interesting to note when comparing the solved crystalline structures of the PPases, is that the ADP-Glc PPase is a larger polypeptide and presents a C-terminus, not present in the CDP-Glc PPase and the RmlA (Fig. 1.3.A and C, respectively). The first ADP-Glc PPase crystalline structure provided us with a very useful tool for understanding the function of those regions or residues that had been already reported as important for activity or regulation. The opposite, however, did not necessarily apply since the ability to visualize the structural organization of a protein still left several questions unanswered, especially those regarding the functionality of specific residues, regions or even whole domains. The research presented in *Chapter 4* of this thesis was motivated by the need of assigning a functional role, if any, to specific residues located in the active site of the enzyme, specifically, at the Glc1P site.



**Figure 1.4. Three-dimensional structures of NDP-sugar PPases monomers.** CDP-Glc PPase (169) (A), and GlmU (163) (B) form trimers in the native enzyme. RlmA (164) (C), and ADP-Glc PPase (170) (D), form tetramers. The pyrophosphorylase domain presents a Rossmann-like- $\beta$ - $\alpha$ - $\beta$  fold, which is conserved in the NDP-sugar PPases family. The difference is mainly in the C-terminus.

The work presented in this thesis has been performed on the *E. coli* ADP-Glc PPase, which has proven to be a very valuable model system to give insight into the understanding of such a key enzyme in the biosynthesis of  $\alpha$ -1,4-polysaccharides.

#### **1.6.4.1. Important residues and domains**

The role of key single amino acids in ADP-Glc PPases was studied by site directed mutagenesis after being identified by chemical modification experiments or suggested by the conservation throughout this enzyme's family. In the *E. coli* enzyme, Tyr<sup>114</sup> was modified by the photoaffinity analogues of ATP and ADP-Glc, 8-azido-ATP and 8-azido-ADP-Glc, respectively (171, 172). Also, site directed mutagenesis showed that this residue is involved in affinity for ATP, but might be also close to the Glc1P and the activator FBP site (173). Following the same approach, it was shown that Lys<sup>195</sup> in the *E. coli* ADP-Glc PPase was involved in Glc1P binding (174-176). Site directed mutagenesis of the homologous residues in the small and large subunit of the potato tuber enzyme, Lys<sup>198</sup> and Lys<sup>213</sup>, respectively, confirmed a similar role for the homologous residue in the small subunit, but not in the large (177), which agreed with the fact that the large subunit has a modulatory role and is not involved in catalysis (147). More recently, Asp<sup>142</sup> in the *E. coli* enzyme, which is absolutely conserved among the ADP-Glc PPase family, was characterized by site directed mutagenesis (167). The specific 10,000-fold decreased  $k_{cat}$  displayed by the D142A and D142N mutants clearly proved its key role in catalysis. Similar results were obtained with the homologous Asp<sup>145</sup> in the potato tuber

small subunit ADP-Glc PPase, but not with Asp<sup>160</sup> in the large subunit, as expected due to the lack of catalytic role of this subunit (126).

Likewise, the role of several residues in allosteric regulation was determined [reviewed extensively in (8, 9)]. Interestingly, lysines and arginines involved in allosteric regulation are located mainly close to the C-terminus in the plant [potato tuber (157) and spinach (178, 179)] and cyanobacterial [*Anabaena* (3, 180-183)] ADP-Glc PPases, whereas those characterized in the bacterial enzymes [*E. coli* (175, 176, 184) and *A. tumefaciens* (185)] are located in the N-terminus. These observations led to the suggestion that the regulatory domains in the enzymes from enteric bacterial on one side, and cyanobacteria and plants on the other, were located at different sites.

By the early 2000, information on the structural organization of the ADP-Glc PPases was available only as a model of predicted secondary structures that fitted all the enzymes from the family sequenced to date (7-9). Residues ~40-300 in the ADP-Glc PPases (of a total of ~430-450 amino acids) aligned well with the pyrophosphorylase domain of two related enzymes already crystallized: the GlnU and the RmlA, despite the low sequence similarity between these enzymes (~14%). Therefore, it was possible to postulate a common fold for the pyrophosphorylase domain but yet, no information was available on the ADP-Glc PPase's C-terminus. In 2002, Ballicora and collaborators (168) performed a study with chimeric ADP-Glc PPases from *E. coli* and *A. tumefaciens* determining that, even though important individual residues involved in activator binding had been identified in their respective N-terminus (Lys<sup>39</sup> in

*E. coli* and Arg<sup>33</sup> and Arg<sup>45</sup> in *A. tumefaciens*) the C-terminus was critical for activator affinity and specificity. This became the first evidence of a combined role of the N- and C-terminus in allosteric regulation (168). Whether the C-terminus was an integral part of a unique domain of the ADP-Glc PPase monomer or a separate domain in intimate interaction with the N-terminus, and whether it had a mere role in regulation or was critical for proper enzyme activity were the questions that motivated the work presented in *Chapter 2*.

On the other hand, the N-terminal tail (~ 40 N-terminal amino acids) had also been involved in regulation. Apart from the amino acids characterized as involved in binding the activators in the bacterial enzymes discussed above, up-regulated ADP-Glc PPase mutants presenting deletions in this short stretch of the enzyme were reported (147, 186-188). Important also to consider is the fact that the reductive activation mechanism regulating plant enzymes involves the formation of a disulfide-bond between Cys<sup>12</sup> of the two small subunits present in the plant heterotetramer (189). Reduction of the disulfide bond increases the affinity of the enzyme for substrates and the activator 3-PGA (154). Work presented in *Chapter 3* was performed addressing questions related to the structural and functional nature of the N-terminus role in regulation.

#### **1.6.4.2. Three-dimensional structure**

The very recent availability of the first ADP-Glc PPase crystal structure in 2005, that from the potato tuber small subunit homotetrameric form (170), shed light on the

monomeric structural organization of this enzyme's family and its hypothetical quaternary arrangement. In agreement with some of the previously obtained results, presented in this thesis in *Chapter 2*, the ADP-Glc PPase monomer is organized in two distinctive domains in close interaction. The N-terminal ~12- 300 residues adopt a similar fold as the pyrophosphorylase domain of the already reported crystal structures of other NDP-sugar PPases: GlnU (163), the RmlA (169) and the CDP-Glc PPase (165). Observation of the potato tuber small subunit crystal structure validated all the previous predictive studies performed on the ADP-Glc PPases (8, 9, 167), and allowed for a visualization of the position of those individual residues already characterized. In the potato tuber ADP-Glc PPase crystallization report (170), a chemical mechanism of action was also postulated based on the orientation and interaction of side chains in the active site with the complexed ligands (ATP or ADP-Glc) and by comparison with the other NDP-Glc PPases, whose mechanisms had been also proposed (163-165). Despite the large amount of structural information provided by the reported structure, several questions concerning the functionality of specific residues in the ADP-Glc PPase remained unanswered. For example, only Asp<sup>145</sup> in potato tuber small subunit (and its homologous Asp<sup>142</sup> in *E. coli*) has been characterized and its essential role in catalysis has been biochemically probed (126, 167). However, other residues postulated to be involved in the chemical mechanism, or with other roles in the active site, have not been studied to date. For instance, Asp<sup>280</sup> (Asp<sup>276</sup> in *E. coli*) and Glu<sup>197</sup> (Glu<sup>194</sup> in *E. coli*) were proposed to have a key role in chelating the Mg<sup>2+</sup> required in the chemical reaction and in binding Glc1P, respectively. These residues have been assigned functional roles based on proximity to the ligands and conservation among the NDP-sugar PPases, but no

biochemical characterization had been performed to date. The need for a detailed analysis of the functional role of these and other residues located in the active site formed the basis that inspired the studies on the *E. coli* ADP-Glc PPase reported in *Chapter 4* of the present thesis.

In summary, the results introduced in the present thesis offer an integrated point of view including information on the biochemistry of ADP-Glc PPases, including structural information at the three-dimensional level together with biochemical data regarding ligand binding and functional roles of the protein domains. This work has aimed to contribute to a better understanding of the mechanism of action and regulation of this enzyme that is fundamental in the bacterial and plant reserve polysaccharide metabolism.

## CHAPTER 2

### **Domain organization of the ADP-glucose pyrophosphorylase from *Escherichia coli*<sup>1</sup>**

---

<sup>1</sup>These results have been published in Bejar, C.M., Ballicora, M.A., Gómez-Casati, D.F., Iglesias, A.A., Preiss, J. (2004) The ADP-glucose pyrophosphorylase from *Escherichia coli* comprises two tightly bound distinct domains. *FEBS lett.*, **573**, 99-104.

## 2.1. ABSTRACT

Computational analysis of ADP-glucose pyrophosphorylases predicts a fold with two domains. Co-expression of two polypeptides comprising residues 1-323 and 328-431 from the *Escherichia coli* ADP-glucose pyrophosphorylase yielded an enzyme form as active as the wild type. The only difference with the wild type was a slightly modified affinity for allosteric effectors. The two polypeptides could not be separated by chromatographic procedures. Separate expression of these polypeptides produced inactive unstable forms. All these results indicated that the ADP-glucose pyrophosphorylase comprises two domains with a strong interaction between them. That interaction is important for allosteric properties and structural stability.

## 2.2. INTRODUCTION

The ADP-Glc PPases from enterobacteria are homotetramers ( $\alpha_4$ ) composed by subunits of ~50 KDa (~440 residues) encoded by the *glgC* gene (8). The three dimensional structure of these enzymes is not available, but structural and computational analyses predict that the folds of ADP-Glc PPases have two domains (3, 8, 9, 16, 167). Despite the low homology (~20% identity), the predicted structure of the central part of the ADP-Glc PPases (residues ~20-300) resembles other NDP-sugar PPases and is postulated to be the catalytic core of the enzyme (3, 8, 9, 16, 167). On the other hand, the C-terminus (~100 residues) was proposed to be important for regulation and specificity for activators (168). A similar conclusion has been reached for plant ADP-Glc PPases (9). Both domains are postulated to interact to regulate the activity because there are residues near the N-terminus that participate in the regulatory site (175, 176, 184, 185). In addition, the presence of an N-terminal tail (~10-20 residues) strongly determines the allosteric properties of the enzyme (175, 176, 184, 185, 187, 188).

Partial proteolysis has been a classical procedure to probe domain boundaries (190, 191). In this chapter, molecular biology techniques were used to test the 2-domain hypothesis. The enzyme was divided into two polypeptides corresponding to the putative domains, they were co-expressed, and the resulting enzyme form was purified and its biochemical properties characterized.

## **2.3. MATERIALS AND METHODS**

### **2.3.1. Materials**

Synthesis and purification of oligonucleotides and DNA sequencing were performed by the Genomics Technology Support Facility at Michigan State University. Perfect Protein™ marker was obtained from Novagen, Inc. The Mono-Q HR 10/10, 5/5 and Phenyl-Superose columns were purchased from Amersham Pharmacia Biotech. [<sup>32</sup>P]PPi was purchased from Perkin Elmer Life Sciences and [<sup>14</sup>C]Glc1P from Amersham Pharmacia Biotech. All other reagents were purchased at the highest quality available.

### **2.3.2. Plasmids and expression vectors**

The construction of pETEC, a pET24a derivative with the *E. coli* ADP-Glc PPase, has been previously described (168). Plasmids pMAB5 and pMAB6 are modified versions of the respective compatible expression vectors pMON17335 and pMON17336 (148). They have an *Nde*I rather than *Nco*I site for cloning.

### **2.3.3. Linker-scanning mutagenesis**

Random introduction of a single 15-bp insertion per plasmid (pETEC) was performed by the technique of Hallet et al. (192) with the commercial GPS™ linker-

scanning mutagenesis kit from New England Biolabs. From 130 insertions mutants obtained, 20 colonies had a single 15 bp inserted in the ADP-Glc PPase coding region. The plasmid with the insertion mutant 117 (pETEC-ins117) was selected for further studies and subcloned in pMAB5 (pMAB5-Ec-ins117) after digestion with *NdeI*-*SacI* for expression in AC70R1-504 cells.

#### **2.3.4. Truncated protein constructs**

To obtain an enzyme with a C-terminal truncation of 108 residues ( $Ec_{1-323}$ ), DNA from the 3' coding region of the *E. coli* ADP-Glc PPase gene was removed as follows. A PCR fragment was amplified using pETEC-ins117 as a template, T7 promoter primer (5'-TAATACGACTCACTATAGGG-3') and the oligonucleotide 5'-ACCGGAGAGCTCTGTTTAAACACG-3', which introduced a *SacI* site 3 bp after the previously inserted stop codon. The PCR fragment was subcloned into an *NdeI*-*SacI* digested pMAB5 to obtain pMAB5- $Ec_{1-323}$ .

#### **2.3.5. Construction of a plasmid for expression of the C-terminus alone ( $Ec_{328-431}$ )**

The DNA fragment encoding the C-terminal portion ( $Ec_{328-431}$ ) was amplified from pETEC with the T7 terminator primer (5'-GCTAGTTATTGCTCAGCGG-3') and the oligonucleotide 5'-GGTAGCCACCATATGACCCTTAACT-3', which introduced an *NdeI* site. The fragment was subcloned into pMAB6 using the sites *NdeI* and *SacI* to

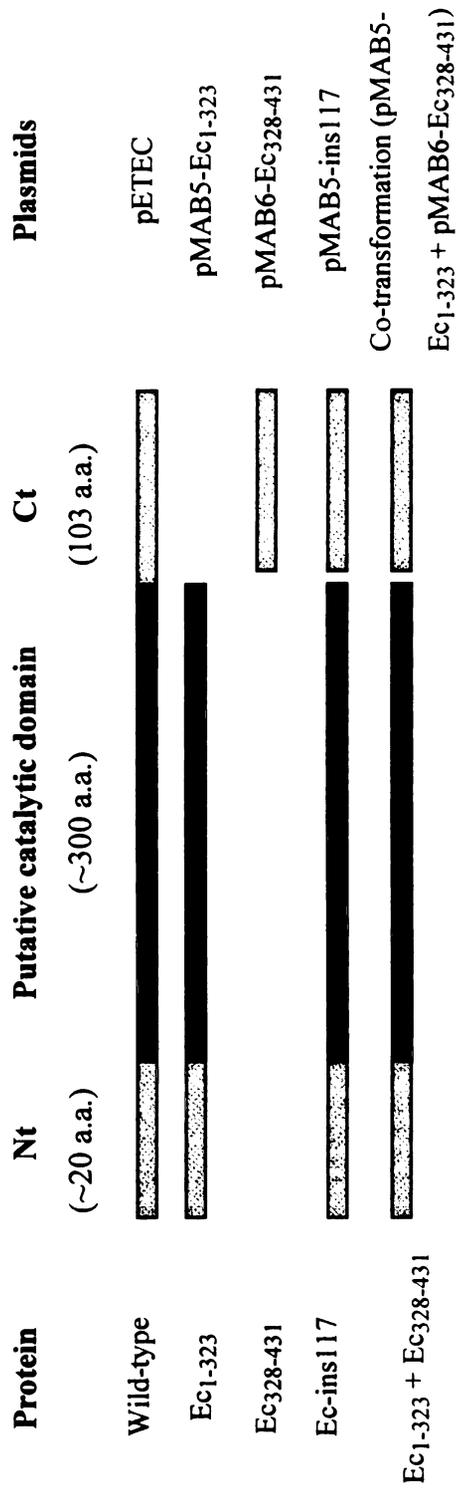
obtain pMAB6-Ec<sub>328-431</sub>. The coding regions of all final products were verified by DNA sequencing and the various constructs used in this work are illustrated in Fig. 2.1.

### **2.3.6. Bacterial strains and expression of the recombinant enzymes**

*E. coli* AC70R1-504 cells without endogenous ADP-Glc PPase activity were used for expression of the enzymes subcloned in pMAB5 and pMAB6 (148). The expression of pMAB5-Ec-ins117 and pMAB5-Ec<sub>1-323</sub> was performed as described previously for pML10 (147). Expression of pMAB6-Ec<sub>328-431</sub> was performed as described previously for pMON17336 (148). Co-expression of pMAB5-Ec<sub>1-323</sub> and pMAB6-Ec<sub>328-431</sub> was performed as was described previously for plasmids pML10 and pMON17336 (147).

### **2.3.7. Protein methods**

Protein assay, electrophoresis (SDS-PAGE) and immunoblotting were performed as described previously (167). N-terminal peptide sequence determination of blotted samples onto PVDF membranes (Bio-Rad) and soluble pure protein samples were performed at the Macromolecular Structure, Sequencing and Synthesis Facility at Michigan State University. Desalting was performed on Bio-Rad 10 DG chromatography columns. Samples were concentrated with Centricon-30 devices (Amicon Inc.).



**Figure 2.1. *E. coli* ADP-Glc PPase constructs.** Ec-ins117 was obtained by linker-scanning mutagenesis. Ec1-323 and Ec328-431 were engineered using PCR-based approaches.

### 2.3.8. Enzymatic assays

*Assay A: Pyrophosphorolysis direction* – Formation of [<sup>32</sup>P]ATP from [<sup>32</sup>P]PP<sub>i</sub> was determined by the method of Morell et al. (123). The reaction was carried out for 10 min at 37 °C in a mixture that contained 50 mM Hepes (pH 8.0), 7 mM MgCl<sub>2</sub>, 1.5 mM [<sup>32</sup>P]PP<sub>i</sub> (1500-2500 cpm/nmol), 2 mM ADP-Glc, 4 mM NaF, and 0.05 mg/ml bovine serum albumin, plus enzyme in a total volume of 0.250 ml. Unless other activators were assayed or specifically stated, 2 mM FBP was added in the reaction mixture.

*Assay B: Synthesis direction* – Formation of ADP-[<sup>14</sup>C]Glc from [<sup>14</sup>C]Glc1P was determined at by the method of Yep et al. (193). The reaction was carried out for 10 minutes at 37 °C in a mixture that contained 50 mM Hepes (pH 8.0), 7 mM MgCl<sub>2</sub>, 0.5 mM [<sup>14</sup>C]Glc1P (~1000 dpm/nmol), 1.5 mM ATP, 1.5 units/ml pyrophosphatase, and 0.2 mg/ml bovine serum albumin, plus enzyme in a total volume of 0.20 ml, unless specifically stated. Unless other activators were tested, 2 mM FBP was added in the reaction mixture. One unit of enzymatic activity is one μmol of product, either [<sup>32</sup>P]ATP or ADP-[<sup>14</sup>C]Glc, formed per min at 37 °C.

### 2.3.9. Purification of Ec-ins117

After induction, twenty liters of AC70R1-504 cells transformed with pMAB5-Ec-ins117 were harvested and crude extracts obtained as described previously (147). The crude extract was precipitated with a 30-60% ammonium sulfate cut. After centrifugation, the precipitate was redissolved in *buffer A* (50 mM Hepes pH 8.0, 5 mM

MgCl<sub>2</sub>, 0.1 mM EDTA, 10% sucrose) and desalted. The sample was applied onto a DEAE-Fractogel column (EMD Chemicals), and eluted with a linear NaCl gradient (0-0.5 M). The active fractions were pooled, desalted, applied to a Green A (Amicon Corp., Lexington, MA) affinity chromatography column (1 ml bed volume) and eluted with a 20-ml linear gradient of NaCl (0-2 M). Purest fractions were pooled, desalted, concentrated, and applied onto a Mono Q HR 5/5 (FPLC, Pharmacia) column equilibrated with *buffer A* and eluted with a linear NaCl gradient (0-0.5 M). The post-Mono Q fractions were pooled, concentrated, resuspended in *buffer B* (*buffer A* plus 1.2 M ammonium sulfate), and applied on a Phenyl-Superose (FPLC, Pharmacia) column equilibrated with *buffer B*. The sample was eluted with a decreasing linear gradient of ammonium sulfate (1.2-0.001 M). The purest fractions were pooled, concentrated, and applied to a Pharmacia Superdex 200 HR 10/30 column equilibrated with *buffer A*. The gel filtration chromatography was run at a flow rate of 0.25 ml/min and fractions with activity were pooled. After this step, Ec-ins117 was >95% pure, with a specific activity of 180 U/mg in the pyrophosphorolysis direction assay (Table 2.1).

### **2.3.10. Purification of Ec<sub>1-323</sub>**

The protein was monitored by immunoblot throughout the purification steps. AC70R1-504 cells harboring pMAB5-Ec<sub>1-323</sub> were grown and induced in 4 liters of LB medium. Crude extracts were precipitated with 0-30% ammonium sulfate and the pellet was resuspended in *buffer A* and desalted. Samples were further purified with a Mono Q HR 10/10 and a Green A affinity column as indicated above (Table 2.1). The fractions

**Table 2.1. Specific activities of mutant and wild-type *E. coli* ADP-Glc PPases**

Concentration of FBP for the mutants was raised to 4 mM. For the sample

Ec<sub>1-323</sub>+Ec<sub>328-341</sub>, concentration of ATP (synthesis direction) was raised to 5 mM.

Sample	Specific Activity		Purity <sup>a</sup> (%)
	Pyrophosphorolysis (U/mg)	Synthesis (U/mg)	
Wild-type	131 ± 3	54 ± 1	90
Ec-ins117	180 ± 9 <sup>b</sup>	ND <sup>c</sup>	>95
Ec <sub>1-323</sub>	< 0.001	<0.0001	50-60
Ec <sub>1-323</sub> + Ec <sub>328-341</sub>	132 ± 5	43 ± 1	>95

<sup>a</sup> Purity of the samples was estimated from SDS-PAGE gels.

<sup>b</sup> Another independent purification produced a pure sample with an activity of 140 U/mg.

<sup>c</sup> not determined

containing the enzyme, ~50-60% pure as detected by immunoblot analysis, were pooled, concentrated, and stored at -80 °C.

### **2.3.11. Purification of co-expressed Ec<sub>1-323</sub> + Ec<sub>328-431</sub>**

AC70R1-504 cells co-transformed with pMAB5-Ec<sub>1-323</sub> + pMAB6-Ec<sub>328-431</sub> were grown and induced in a 1-liter culture. The resulting crude extracts were applied onto a DEAE-Fractogel column (EMD Chemicals Inc.), and eluted with a linear NaCl gradient (0-0.5 M). The active fractions were pooled and precipitated with 30-60% ammonium sulfate. After centrifugation, the pellet was resuspended in *buffer A* and desalted. The sample was applied to a Mono Q HR 10/10 (FPLC, Pharmacia) column equilibrated with *buffer A* and eluted with a linear NaCl gradient (0-0.5 M). Purest fractions, assessed by SDS-PAGE, were pooled, concentrated and applied to a Green A affinity chromatography column. The sample was eluted with a linear gradient of NaCl (0-2 M). After this step both co-expression products, Ec<sub>1-323</sub> + Ec<sub>328-431</sub> accounted for >95% of the protein (Table 2.1).

### **2.3.12. Molecular mass determination**

Purified wild type *E. coli* ADP-Glc PPase and the mutants Ec-ins117, Ec<sub>1-323</sub> + Ec<sub>328-431</sub>, and Ec<sub>1-323</sub> were applied to a column Pharmacia Superdex 200 HR 10/30 equilibrated with *buffer A* and run at 0.25 ml/min. Peaks were followed by absorbance at 280 nm, tested for enzymatic activity or confirmed by immunoblot (Ec<sub>1-323</sub>). For

calibration, ferritin (440 kDa), *E. coli* ADP-Glc PPase (200 KDa), aldolase (158 kDa), and Hemoglobin (67 KDa) were used as standard proteins. The calibration line and the interval confidence (90%) were obtained after plotting the data as log (molecular mass) vs. elution volume in the program Origin<sup>®</sup> 5.0. Molecular mass was expressed with a confidence interval of 90%.

### **2.3.13. Kinetic characterization**

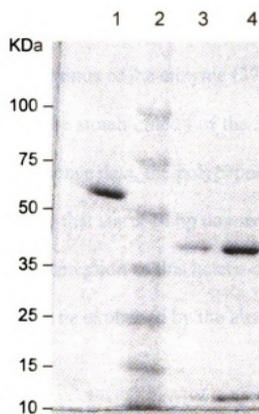
Kinetic data were plotted as specific activity ( $\mu\text{mol min}^{-1} \text{mg}^{-1}$ ) versus substrate or effector concentration. Kinetic constants were acquired by fitting the data to the Hill equation with a nonlinear least square formula using the program Origin<sup>™</sup> 5.0. Hill plots were used to calculate the Hill coefficient  $n_H$  and the kinetic constants that correspond to the activator, substrate or inhibitor concentrations giving 50% of the maximal activation ( $A_{0.5}$ ), velocity ( $S_{0.5}$ ), and inhibition ( $I_{0.5}$ ).

## **2.4. RESULTS AND DISCUSSION**

Random insertions of a single 15-bp fragment into a recombinant *E. coli glgC* gene were produced to study structure-function relationships on the ADP-Glc PPase (M. A. Ballicora and J. Preiss, unpublished results). One of the insertion mutants generated, Ec-ins117, had a stop codon after residue 323 in the translated polypeptide of the *E. coli glgC* gene. Surprisingly, the polypeptide still possessed ADP-Glc PPase activity although the stop codon would truncate 108 residues of the C-terminus. Colonies of BL21(DE3) cells expressing Ec-ins117 or the wild type enzyme, but not a control plasmid (pET24a), stained brown in presence of iodine vapors (unpublished results). This indicated high levels of glycogen in the cell and suggested the presence of ADP-Glc PPase activity (194). For this reason, the properties of this putatively truncated form were characterized.

### **2.4.1. Purification and characterization of Ec-ins 117**

Despite the lower expression level of Ec-ins117 as observed by immunoblot and SDS-PAGE (not shown), the pyrophosphorolysis activity of this mutant was 0.31 U/mg in the crude extracts. That was two orders of magnitude lower than the over-expression of the wild-type ADP-Glc PPase but two orders of magnitude higher than pET24a, the plasmid control without insert (not shown). Attempts to isolate this ~37 kDa truncated protein resulted in the co-purification of a 12 kDa peptide even after five different chromatographic columns (Fig. 2.2). Despite the differences in molecular size, both polypeptides comigrated in a high-resolution gel filtration column (Superdex 200 HR).



**Figure 2.2. SDS-PAGE of *E. coli* wild-type ADP-Glc PPase and Ec-ins117 mutant.** Lane 1, *E. coli* wild-type ADP-Glc PPase; lane 2, pre-stained molecular mass markers; lane 3, Ec-ins117 mutant (0.5  $\mu$ g); and lane 4, Ec-ins117 mutant (1.5  $\mu$ g)

The N-terminal sequence of the 12 kDa band, obtained after SDS-PAGE and blotting, was MTLNSLVSGG, which indicated that the polypeptide corresponds to the C-terminus of the *E. coli* ADP-Glc PPase with residue 328 as the N-terminal methionine (103 amino acids). In addition, N-terminal sequencing of the purified Ec-ins117 in solution confirmed the presence of two polypeptides, suggesting that both the C-terminal fragment together with the N-terminus of the enzyme (37 kDa band) were being expressed. It also indicated that the stoichiometry of the 37 and 12 kDa polypeptides was 1:1 (Fig. 2.3). Based on the sequence data, the polypeptide of 12 kDa appeared to be the product of an open reading frame that starts 24 bp downstream of the inserted stop codon, at Met<sup>328</sup> (Fig. 2.4). The lower expression of this hetero-oligomer form, compared to the wild type expression level, could be explained by the absence of a proper ribosome-binding site before Met<sup>328</sup>.

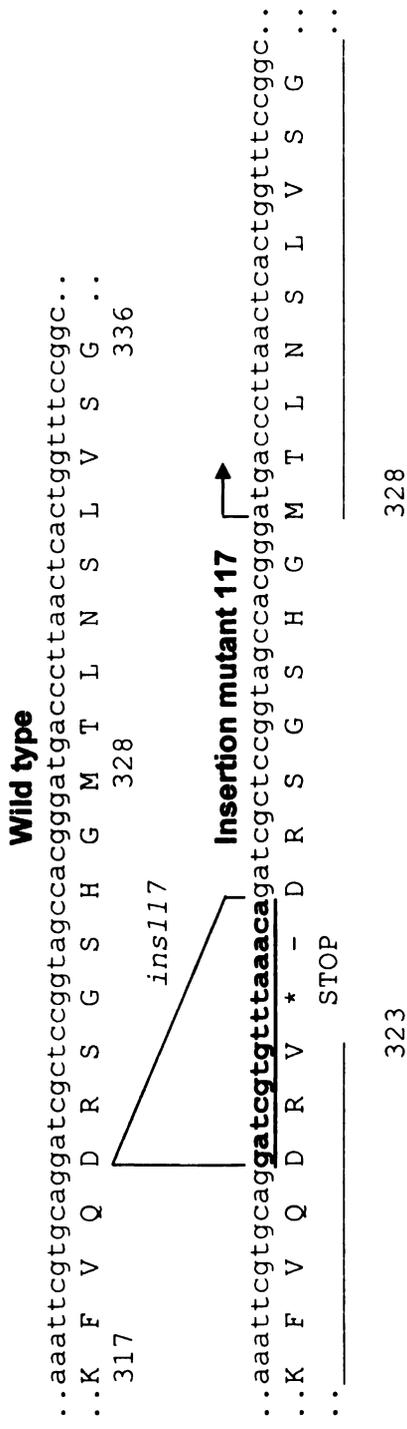
Kinetic properties of the Ec-ins117 mutant purest sample were determined in the pyrophosphorolysis direction (Table 2.2). Apparent affinities for the substrates (ADP-Glc, PP<sub>i</sub>) and cofactor (Mg<sup>2+</sup>), were similar to those of the wild type enzyme. The main difference was the apparent affinity for the activators. The  $A_{0.5}$  for FBP was almost 6-fold higher for Ec-ins117 compared to the wild type.

Amino Acid	Residue Number						
	1	2	3	4	5	6	7
	pmol						
D	3.11	2.35	2.27	2.88	2.21	2.66	<b>10.19</b>
N	0.72	0.51	0.87	<b>9.09</b>	1.44	<b>8.08</b>	2.78
S	3.24	<b>11.10</b>	2.04	1.96	<b>8.85</b>	1.53	2.32
T	1.41	<b>11.40</b>	0.49	0.57	0.35	0.22	0.77
E	0.49	0.50	0.63	<b>9.98</b>	2.59	1.33	1.24
M	<b>14.60</b>	0.33	0.20	0.23	0.27	0.36	0.43
V	<b>14.90</b>	1.89	1.03	1.03	1.21	1.78	<b>11.90</b>
K	0.49	0.47	ND	0.66	<b>13.39</b>	3.24	1.47
L	0.42	1.39	<b>29.18</b>	3.31	ND	<b>13.09</b>	2.98

Sequences deduced							
1	V	S	L	E	K	N	D
2	M	T	L	N	S	L	V

**Figure 2.3. N-terminal sequence analysis of the Ec-ins117 mutant.** Automated sequencing of the Ec-ins117 protein was performed directly from the purest fraction. The values are the picomoles of amino acid derivative found in each cycle of degradation. In each cycle of cleavage, major peaks are in bold. Residues Q, G, H, A, R, Y, P, W, F, I. were negligible in all these cycles and were omitted for simplicity. ND: Not detectable.



**Figure 2.4. Sequence of the insertion 117 in the *E. coli* ADP-Glc PPase.** The upper sequence corresponds to the wild-type ADP-Glc PPase and the lower to the Ec-ins117, in which a 15-bp linker was inserted by linker-scanning mutagenesis as described in section 2.3. Underlined is the 15-bp insertion. A stop codon was introduced at residue 324, generating the transcription and translation of two polypeptides: the ~37 kDa (comprising residues 1-323), and the 12 kDa peptide starting at Met<sup>328</sup>.

**Table 2.2. Apparent affinity for the different effectors of the mutant Ec-ins117 and wild-type ADP-Glc PPase.** The assays were performed in the pyrophosphorolysis direction

Effectors	Enzyme		
	Wild type	Ec-ins117	Ratio <sup>a</sup>
Substrates	$S_{0.5}$ ( $\mu\text{M}$ )		
ADP-Glc	166 $\pm$ 26	241 $\pm$ 25	1.45
PP <sub>i</sub>	81 $\pm$ 14	45 $\pm$ 13	0.55
Cofactor	$S_{0.5}$ ( $\mu\text{M}$ )		
Mg <sup>2+</sup>	1420 $\pm$ 310	2147 $\pm$ 86	1.51
Activators	$A_{0.5}$ ( $\mu\text{M}$ )		
FBP	37.8 $\pm$ 5.2	214 $\pm$ 53	5.7
PLP	0.045 $\pm$ 0.004	0.156 $\pm$ 0.001	3.5
PEP	282 $\pm$ 15	> 2000 <sup>b</sup>	> 7.1
Inhibitor	$I_{0.5}$ ( $\mu\text{M}$ )		
AMP	41 $\pm$ 4 <sup>c</sup>	21 $\pm$ 8 <sup>d</sup>	0.51

<sup>a</sup> This is the value of the kinetic constant of Ec-ins117 divided the wild type.

<sup>b</sup> Saturation was not reached.

<sup>c</sup> To observe AMP inhibition, activator should be added. In this case, FBP was 40  $\mu\text{M}$  to avoid saturating concentrations of activator and allowed the competitive inhibition by AMP. That concentration is about the  $A_{0.5}$  value.

<sup>d</sup> FBP was 300  $\mu\text{M}$ , which is about the  $A_{0.5}$  value.

#### **2.4.2. Purification and characterization of C-terminal truncated forms**

It was important to determine whether the truncated enzyme was active by itself or if the co-expression with the C-terminal 12 kDa polypeptide contributed to the reconstitution of the activity. To ensure the expression of a single polypeptide comprising residues 1-323 ( $Ec_{1-323}$ ), a *SacI* restriction site was introduced immediately downstream of the stop codon by site directed mutagenesis. Only the  $Ec_{1-323}$  coding fragment was subcloned in pMAB5, expressed, and purified as described in section 2.3. With this approach, most of the expressed enzyme was insoluble, whereas the soluble fraction had a peptide of about 37 kDa and some lower bands recognized by immunoblotting, which indicated that it was highly proteolyzed (data not shown).  $Ec_{1-323}$  (~37 kDa) was purified to ~50 % purity where the main contaminants were lower bands still recognized by immunoblot. The activity was negligible (Table 2.1) and did not increase significantly with higher concentrations of FBP, up to 10 mM (data not shown). Gel filtration showed that the sample was highly aggregated (data not shown).

The total loss of activity after the removal of the C-terminus indicated that this domain is involved in more than regulation. In the inactive  $Ec_{1-323}$  mutant, the C-terminal deletion yielded a form highly susceptible to proteolysis. In an attempt to study which specific regions of this domain were important, deletions of 20 to 90 residues (every 10) from the C-terminus were evaluated. The truncated enzymes were expressed in small-scale systems and were visualized by SDS-PAGE and western blot at their expected molecular sizes (data not shown). The recombinant products were found mainly in the

insoluble fraction of the sample and the pyrophosphorolysis ADP-Glc PPase activity was negligible in all crude extracts ( $<0.0001$  U/mg). These results suggest most of, or maybe all of the  $\sim 100$  residues of the C-terminus are necessary for maintaining a functional *E. coli* ADP-Glc PPase. This may be due to the role of specific residues or to the overall structure of the domain that is necessary for the integrity of the enzyme.

### **2.4.3. Characterization of Ec<sub>1-323</sub> + Ec<sub>328-431</sub>**

The results obtained with Ec<sub>1-323</sub> strongly suggested that the properties of Ec-ins117 are due to the presence of both the 37 kDa and the 12 kDa polypeptides. To verify this possibility, the *E. coli glgC* fragment encoding the 103 C-terminal residues of ADP-Glc PPase (Ec<sub>328-431</sub>) was subcloned into a compatible expression vector (pMAB6). Co-expression of Ec<sub>1-323</sub> and Ec<sub>328-431</sub> was performed in AC70R1-504 cells, yielding extracts with activity (39 U/mg) comparable to the wild type (24 U/mg) (168). That level of expression was about 100-fold higher than the Ec-ins117. After the same purification procedure of the wild-type enzyme, Ec<sub>1-323</sub> and Ec<sub>328-431</sub> co-purified, indicating a very strong interaction. The 37-kDa fragment did not appear as proteolyzed as when it was expressed alone (data not shown). In addition, the specific activity values were comparable to that of the wild-type enzyme (Table 2.1). The molecular mass determined by gel filtration in native conditions of the co-expressed Ec<sub>1-323</sub> + Ec<sub>328-431</sub> ( $227 \pm 32$  kDa) and Ec-ins117 mutant ( $213 \pm 30$  kDa) were not distinguishable from the wild-type enzyme (200 kDa). These results, together with the 1:1 stoichiometry (Fig. 2.3), indicate that the structure of both Ec<sub>1-323</sub> + Ec<sub>328-431</sub> and Ec-ins117 are  $\alpha'_4\omega_4$ , rather than  $\alpha_4$  as the

wild type ( $\alpha'$  and  $\omega$  being the polypeptides  $Ec_{1-323}$  and  $Ec_{328-431}$ , respectively). The apparent affinities for the substrate ATP (4.7-fold higher  $S_{0.5}$ ) and for the activator FBP (3.5-fold higher  $A_{0.5}$  in the synthesis direction) were altered compared to those of the wild type (Table 2.3). The inhibition by AMP was not significantly changed and the apparent affinity for  $Glc1P$  was identical (Table 2.3).

**Table 2.3. Kinetic parameters in the synthesis direction of *E. coli* ADP-Glc PPase wild type and Ec1-323 + Ec328-431 form.**

Except for the ATP curves, the concentration of ATP for the Ec1-323 + Ec328-431 was 5 mM to ensure saturation. For the ATP and Glc1P curves, the concentration of FBP was 2 mM for the wild type and 4 mM for Ec1-323 + Ec328-431.

Enzyme	ATP		Glc1P		FBP		AMP			
	$S_{0.5}$ ( $\mu\text{M}$ )	$n_H$	$S_{0.5}$ ( $\mu\text{M}$ )	$n_H$	$A_{0.5}$ ( $\mu\text{M}$ )	$n_H$	$I_{0.5}^a$ ( $\mu\text{M}$ )	$n_H$	$I_{0.5}^b$ ( $\mu\text{M}$ )	$n_H$
Wild type	301 ± 14	1.9	15.0 ± 1.3	1.0	92 ± 1	2.1	3.8 ± 0.4	1.1	92 ± 4	2.7
Ec1-323 + Ec328-431	1410 ± 51	2.9	13.6 ± 1.2	1.2	327 ± 15	2.0	10 ± 2	1.2	41 ± 3	0.9

<sup>a</sup> Inhibition curve performed in presence of 30  $\mu\text{M}$  FBP

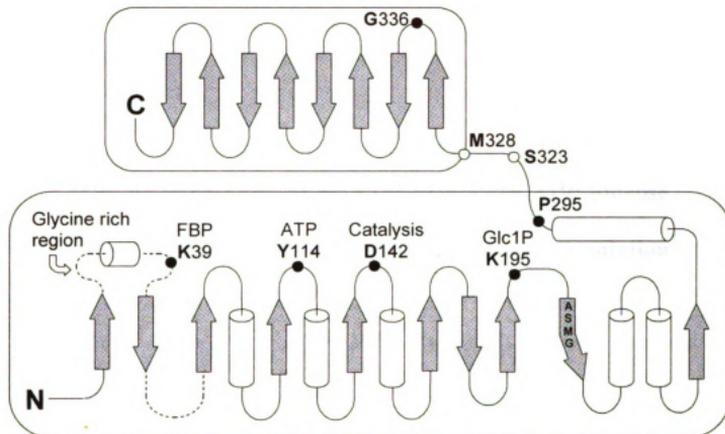
<sup>b</sup> Inhibition curve performed in presence of 130  $\mu\text{M}$  FBP

## 2.5. CONCLUSIONS

Important structural information can be inferred from these results: (i) the stop codon inserted in the Ec-ins117 mutant, which generated a “nick” in the protein sequence, was in a non-critical position for the 3D structure of the protein. (ii) These data also support structure predictions and the hypothesis that the *E. coli* ADP-Glc PPase is organized in at least two distinct domains (8). According to the proposed model of domain organization (8, 9, 167), the stop codon was inserted in a loop that separates the two putative domains (Fig. 2.5). The data also agree with the hypothesis that the C-terminal domain is ~100 residues long and linked to the catalytic domain by a long loop (~30 residues) (8). Removal of this domain yielded an inactive enzyme rather than a shorter unregulated version as other NDP-sugar PPases of 30-40 KDa (195-198). Even the removal of 20 residues at the C-terminus rendered a form without activity. It is possible that the absence of even a small part of the C-terminal domain leaves the enzyme in a misfolded or insoluble form. On the other hand, a “nick” separating the C-terminus and the catalytic domain, obtained by the co-expression of both the N- and the C-terminal polypeptides, produced an enzyme with very similar properties to the wild type. Only the apparent affinity for the activator was slightly modified. Possibly, the lack of a covalent bond between the N- and C-domains yields a more relaxed enzyme that favors less the activated conformation. A slightly higher concentration of activator overcame this problem, 3.5- and 5.7-fold in the synthesis and pyrophosphorolysis direction, respectively. The possibility that the residues removed in the “nick” (S<sup>323</sup>GSHG<sup>327</sup>) are the reason for the decrease in activator affinity cannot be ruled out. However, in other

ADP-Glc PPases, those residues are not conserved or they are deleted, together with Met<sup>328</sup> (8). Most likely, they are part of an area that connects the two major domains and does not play a specific role (Fig. 2.5).

Partial proteolysis has been a classical procedure to detect domain boundaries (190, 191, 199). In this work, molecular biology techniques were used to probe whether the *E. coli* ADP-Glc PPase comprises two domains. After co-expression of the C- and the N-terminal domains, they remained non-covalently bound, with a tight interaction that cannot be disturbed even after passage through several chromatography columns. Presence of the C-terminal domain stabilizes the enzyme and prevents proteolytic degradation of the N-terminal domain. These results provide a structural insight for the differential roles of these regions of the ADP-Glc PPase. The hypothesis is that they are two strongly interacting domains that have to be expressed together to obtain a fully functional enzyme. According to previous analysis (8, 167), the catalytic domain is located in the 37 kDa fragment (similar to most NDP-sugar PPases) and the responsibility for the regulation is shared by both the 12 and the 37 kDa fragments.



**Figure 2.5. Domain organization of the ADP-Glc PPase from *E. coli*.** The secondary structure is based on the prediction from references (8) and (9). Arrows and cylinders represent  $\beta$ -sheets and  $\alpha$ -helices, respectively. Dark circles are important residues that interact with effectors or participate in catalysis. When Gly<sup>336</sup> and Pro<sup>295</sup> are replaced, allosteric mutants are generated (8). After insertion of a stop codon, residues between Ser<sup>323</sup> and Met<sup>328</sup> were “nicked” (see Fig. 2.4).

## CHAPTER 3

### **The structural role of the *Escherichia coli* ADP-glucose pyrophosphorylase's N-terminus in allosteric regulation<sup>2</sup>**

---

<sup>2</sup>This work presented in this chapter has been published: Bejar, C.M., Ballicora M.A., Iglesias, A.A., Preiss, J. (2006) ADP-glucose pyrophosphorylase's N-terminus: structural role in allosteric regulation. *Biochem. Biophys. Res. Commun.*, **343**: 216-221.

### 3.1. ABSTRACT

The functional role of the *Escherichia coli* ADP-glucose pyrophosphorylase's N-terminus in allosteric regulation, and the particular effects caused by its length, were studied and presented in this chapter. Small truncated mutants were designed, and those lacking up to 15-residues were active and highly purified for further kinetic analyses. N $\Delta$ 3 and N $\Delta$ 7 did not change the kinetics parameters respect to the wild-type. N $\Delta$ 11 and N $\Delta$ 15 enzymes were insensitive to allosteric regulation and highly active in the absence of the activator. Co-expression of two polypeptides corresponding to the N- and C-termini generated an enzyme with lower activation properties than the wild type (see *Chapter 2*). In this section, the characterization of a N $\Delta$ 15 co-expression mutant, in which the allosteric regulation was restored to wild-type levels, is presented. Unusual allosteric effects caused by either an N-terminal truncation or co-expression of individual domains may respond to structural changes favoring an up-regulated or a down-regulated conformation rather than to the disruption of specific activator or inhibitor sites.

### 3.2. INTRODUCTION

The relevance of starch as a natural polymer and raw material for food and non-food industries makes ADP-Glc PPase an enzyme with particular interest in biotechnology (7). It has been clearly established that manipulation of the starch content in plants requires managing levels of activity of the enzyme within plastids in specific tissues (7, 60). For these technological purposes, the understanding of the enzyme's regulatory mechanisms is a key issue.

Characterization of amino acids involved in binding of activator for the ADP-Glc PPase from *E. coli* and *Agrobacterium tumefaciens* performed by chemical modification and site-directed mutagenesis allowed for pointing out the relevance of the protein's N-terminus (N-t) in allosteric regulation (175, 176, 184, 185). More recent studies performed with chimeric enzymes from these bacteria have indicated the involvement of the N-t together with the C-terminus (C-t) domain (and the interaction between them) in determining the specificity and affinity for the activator (168). Also, the role of the C-t domain in regulation and its tight interaction with the catalytic region of the protein has been evidenced for the *E. coli* enzyme (reported in Chapter 2).

The importance of the N-terminal region for the allosteric activation and inhibition of the *E. coli* ADP-Glc PPase was evidenced by characterization of truncated enzymes (187, 188). The enzyme lacking 10 to 13 amino acids from the N-t and 2 amino acids from the C-t after treatment with *proteinase K* is almost independent of the need of

FBP for maximal activity and is insensitive to inhibition by AMP (187). These results were confirmed by engineering a recombinant truncated enzyme with 11 amino acids deleted from the N-t and 2 from the C-t (188). The unusual regulatory properties of this truncated ADP-Glc PPase justified a more detailed characterization of the actual role the N-t extension has in enzyme activation and its interaction with the C-t of the protein. The motivation for the work presented in this chapter was to analyze the structural bases determining the enzyme's regulatory properties and how the length of the N-t affects the protein ability to arrange specific 3D conformations. Another question addressed here is how the N-t communicates with other domains involved in regulation. As a result, the present work provides us with a more insights into the relationship between structure, function and regulation of the ADP-Glc PPases.

### 3.3. MATERIALS AND METHODS

#### 3.3.1. Construction of N-terminal truncated enzymes

*E. coli* N-terminal truncated forms (EcN $\Delta$ 3, EcN $\Delta$ 7, EcN $\Delta$ 11, EcN $\Delta$ 15, EcN $\Delta$ 19 and EcN $\Delta$ 22) were encoded by pETEC (168) derivatives and were obtained using the following forward primers, which introduced an *Nde*I site (in *italics*):

EcN $\Delta$ 3: 5' ATG GTT *CAT ATG* GAG AAG AAC GAT CAC T 3'

EcN $\Delta$ 7: 5' GAG AAG *CAT ATG* CAC TTA ATG TTG GCG C 3'

EcN $\Delta$ 11: 5' CAC TTA *CAT ATG* GCG CGC CAG CTG CCA 3'

EcN $\Delta$ 15: 5' GCG CGC *CAT ATG* CCA TTG AAA TCT GTT 3'

EcN $\Delta$ 19: 5' CCA TTG *CAT ATG* GTT GCC CTG ATA CTG G 3'

EcN $\Delta$ 22: 5' TCT GTT *CAT ATG* ATA CTG GCG GGA GGA C 3'

The T7 terminator was used as reverse primer. The fragments were sub-cloned in pET24a vector between *Nde*I-*Sac*I sites to form pETEC-N $\Delta$ 3, pETEC-N $\Delta$ 7, pETEC-N $\Delta$ 11, pETEC-N $\Delta$ 15, pETEC-N $\Delta$ 19, and pETEC-N $\Delta$ 22.

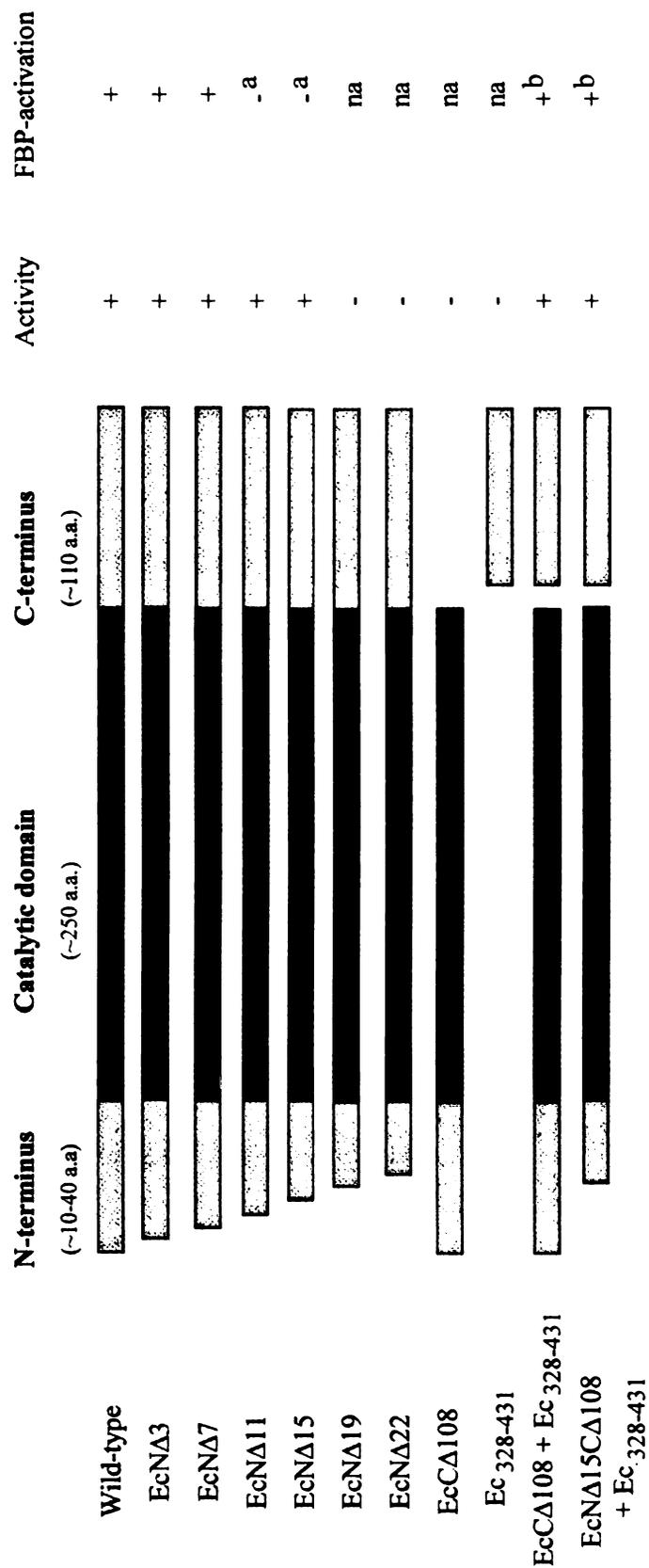
EcN $\Delta$ 15-C $\Delta$ 108 coding DNA was amplified by PCR from pETEC-N $\Delta$ 15 with the downstream primer used to construct EcC $\Delta$ 108 (see *Chapter 2*), which introduced a *Sac*I site right after the codon corresponding to amino acid 323. The amplified coding region was then cloned as *Nde*I-*Sac*I fragment in pMAB5. The plasmid coding for *E. coli* ADP-Glc PPase C-terminal residues 328-431, pMAB6-Ec<sub>328-431</sub>, was previously obtained

(see *Chapter 2*). All constructs used (Fig. 3.1) were verified by DNA sequencing.

### 3.3.2. Expression of the recombinant enzymes

N-terminal truncated enzymes encoded by pETEC derivative vectors, were expressed in *E. coli* BL21(DE3) cells as previously described (168). EcN $\Delta$ 15-C $\Delta$ 108 + Ec<sub>328-431</sub> were co-expressed in an ADP-Glc PPase deficient *E. coli* B strain, AC70R1-504 (148), as described for EcC $\Delta$ 108 + Ec<sub>328-431</sub> (*Chapter 2*). In small-scale expression assays, single colonies of the pETEC-derivatives transformants were grown in 50 ml Luria broth medium at 37 °C with 50  $\mu$ g/ml kanamycin up to OD<sub>600</sub> = 0.8. The expression was induced at room temperature (23-25 °C) for 4 h by addition of 1 mM final concentration of IPTG. For the co-expression of the pMAB-derivative plasmids, cells were grown in the same conditions but adding 50  $\mu$ g/ml kanamycin plus 70  $\mu$ g/ml spectinomycin and carrying out induction with 1 mM final concentration of IPTG and 5  $\mu$ g/ml nalidixic acid for 16 h. In both cases, after induction, cells were chilled on ice and harvested by centrifugation. Cell pastes were resuspended in 3 ml *buffer A* (50 mM HEPES pH 8.0, 5 mM MgCl<sub>2</sub>, 0.1 mM EDTA, and 10% w/v sucrose). All subsequent protein purification steps were conducted at 0-4 °C. Cells were disrupted by sonication, centrifuged for 15 min at 15,000  $\times$ g and the supernatants (crude extracts) were stored at –80 °C. Dilutions of the samples (2-fold and 10-fold) were tested for activity in the pyrophosphorolysis direction as described below.

***E. coli* ADPGlc PPase mutants employed**



**Figure 3.1. Summary of the *E. coli* ADP-Glc PPase constructs studied and their activation properties. (a)** The mutant enzymes were fully active in the absence or at low concentration of the activator FBP. (b) Affinity for the activator was lower than the wild type enzyme. (na) non-applicable.

### 3.3.3. Purification of N-truncated ADPGlc PPase mutants

Recombinant *E. coli* BL21(DE3) were grown in 1-2 L and induced as described above. The cell pastes were resuspended in 15-30 ml of *buffer A*, sonicated, and the lysates were cleared by centrifugation. The resulting crude extracts were applied to a DEAE-Fractogel column (EMD Chemicals INC), and eluted with a linear NaCl gradient (0-0.5 M). The active fractions were pooled and precipitated with a 30-60% ammonium sulfate cut. After centrifugation, the pellet was resuspended in *buffer A* and desalted on Bio-Rad 10 DG chromatography columns equilibrated with the same buffer. The desalted samples were individually applied to a Mono Q HR 10/10 (FPLC, Pharmacia) column equilibrated with *buffer A* and eluted with a linear NaCl gradient (0-0.5 M). The purest fractions (as assessed by SDS-PAGE) were pooled and concentrated using Centricon-30 devices (Amicon Inc.). After this step, EcNΔ3, EcNΔ7, and EcNΔ15 were >95% pure. EcNΔ11 MonoQ active fractions were pooled, concentrated by 80% ammonium sulfate precipitation, applied to a phenyl-Superose (FPLC, Pharmacia) column, and eluted with a decreasing gradient of ammonium sulfate (1.2-0.0 M). The active fractions were pooled and concentrated with Centricon-30 devices.

### 3.3.4. Purification of co-expressed EcNΔ15– CΔ108+ Ec<sub>328-431</sub>

Over-expression and purification were carried out as described for EcCΔ108+Ec<sub>328-431</sub> (see *Chapter 2*).

### **3.3.5. Protein concentration assays**

Protein concentration in the crude extracts and in the subsequent purification steps was measured using the bicinchoninic acid reagent (200) (Pierce Chemical Co.), with BSA as the standard. Protein concentration of the purified enzymes was determined by UV absorbance at 280 nm using an extinction coefficient of  $1.0 \text{ ml mg}^{-1} \text{ cm}^{-1}$  (201).

### **3.3.6. Protein Electrophoresis and Immunoblotting**

Purification of the recombinant proteins was monitored by SDS-PAGE as described by Laemmli (202), using 4-15% Tris-HCl pre-cast gradient polyacrylamide gels (Bio-Rad). Perfect<sup>TM</sup> protein markers were used as molecular mass standards. Following electrophoresis, protein bands were either visualized by staining with Coomassie brilliant blue R-250 or electroblotted onto a Protran<sup>TM</sup> nitrocellulose membrane. The membrane was then treated with affinity purified anti-*E. coli* B strain AC70R1 ADP-Glc PPase IgG (201). The resulting antigen-antibody complex was visualized by alkaline phosphatase-linked goat anti-rabbit IgG, and then stained with BM purple AP-substrate precipitating reagent (Roche Molecular Biochemicals).

### **3.3.7. Enzymatic activity assays**

*Assay A: Pyrophosphorolysis* – Formation of [<sup>32</sup>P]ATP from [<sup>32</sup>P]PP<sub>i</sub> in the direction of pyrophosphorolysis at 37 °C was determined by the method previously

described (123). Unless otherwise indicated, the reaction was carried out for 10 min in a mixture that contained 50 mM HEPES (pH 8.0), 7 mM MgCl<sub>2</sub>, 1.5 mM [<sup>32</sup>P]PP<sub>i</sub> (1500-2500 cpm/nmol), 2 mM ADPGlc, 2 mM FBP, 4 mM NaF, and 0.05 mg/ml BSA, plus enzyme in a total volume of 250 μl.

*Assay B: Synthesis* – Formation of [<sup>14</sup>C]ADP-glucose from [<sup>14</sup>C]Glc1P in the synthesis direction was determined at 37 °C by the method of Yep and coworkers (193). The reaction was carried out for 10 min in a mixture that contained of 50 mM HEPES (pH 8.0), 7 mM MgCl<sub>2</sub>, 0.5 mM [<sup>14</sup>C]Glc1P (~1000 dpm/nmol), 1.5 mM ATP, 2 mM FBP, 0.0015 units/μl pyrophosphatase, and 0.2 mg/ml BSA, plus enzyme in a total volume of 200 μl, unless specifically stated variations.

One unit of enzymatic activity is equal to 1 μmol of product, either [<sup>32</sup>P]ATP or [<sup>14</sup>C]ADP-Glc, formed per min at 37 °C.

### **3.3.8. Kinetic characterization**

Kinetic data were plotted as specific activity (nmol min<sup>-1</sup> mg<sup>-1</sup>) versus substrate or effector concentration. Kinetic constants were acquired by fitting the data to the Hill equation with a nonlinear least square method using the program Origin™ 5.0. Hill plots were used to calculate the Hill coefficient  $n_H$  and the kinetic constants that correspond to the activator, substrate or inhibitor concentrations giving 50% of the maximal activation ( $A_{0.5}$ ), velocity ( $S_{0.5}$ ), and inhibition ( $I_{0.5}$ ).

### **3.4. RESULTS AND DISCUSSION**

#### **3.4.1. Purification and characterization of *E. coli* ADPGlc PPase N-terminal truncated forms**

To evaluate in detail the structural role of the N-t extension in allosteric regulation, small deletion mutants of the ADP-Glc PPase from *E. coli* were designed (Fig. 3.1) and tested. EcN $\Delta$ 3, EcN $\Delta$ 7, EcN $\Delta$ 11, EcN $\Delta$ 15, EcN $\Delta$ 19 and EcN $\Delta$ 22 truncated proteins were over-expressed in *E. coli* BL-21(DE3) and purified in small-scale. EcN $\Delta$ 3, EcN $\Delta$ 7 and EcN $\Delta$ 15 were enriched to >95%, and EcN $\Delta$ 11 to 60-70 % purity (Table 3.1). As shown in Table 3.1, the four mutants had specific activities comparable to that of the purified wild-type enzyme. Thus, the N-t comprising at least 15 amino acid residues of the *E. coli* ADP-Glc PPase is not essential for catalytic activity. Conversely, deletions of 19- and 22-residues from the N-t significantly decreased the activity of ADP-Glc PPase in the crude extracts. In fact, enzymatic activity of EcN $\Delta$ 19 measured by the pyrophosphorolysis assay was two orders of magnitude lower than that of the wild-type (also in the crude extract), whereas activity of EcN $\Delta$ 22 was negligible (Table 3.1).

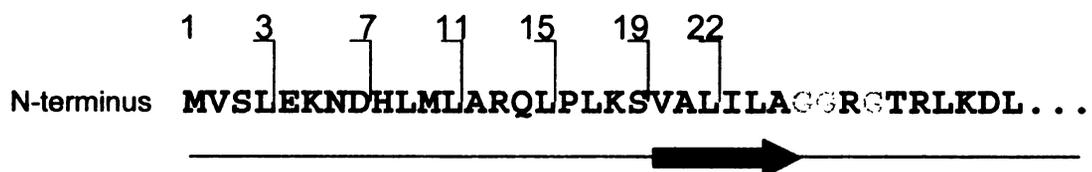
Results in Table 3.1 were consistent with the predicted secondary structures of ADP-Glc PPase (167) and with the 3D model previously proposed (8). This model was also validated with the data derived from the crystal structure of an inhibited form of the homotetrameric (small subunit) enzyme from potato tuber (170). In the structural model the N-terminal tail is a loop preceding a  $\beta$  strand starting at Ser<sup>19</sup> (Fig. 3.2).

**Table 3.1. Specific activities of wild-type and N-terminal truncated *E. coli* ADP-Glc PPases. Specific activities and FBP-activation kinetics were determined by the pyrophosphorolysis assay. Purity of the samples was estimated from SDS-PAGE gels.**

Sample	Specific Activity (U/mg)	Purity (%)	FBP activation kinetics		Activation
			$A_{0.5}$ ( $\mu$ M)	$n_H$	
Ec wild-type	131	90	36.4 $\pm$ 1.7	2.1	29.5
EcN $\Delta$ 3	94	>95	52.3 $\pm$ 5.6	1.8	23.5
EcN $\Delta$ 7	65	>95	88.7 $\pm$ 3.9	2.9	22.0
EcN $\Delta$ 11	98	60-70	-	-	1.0
EcN $\Delta$ 15	105	>95	-	-	1.2
Ec wild-type <sup>a</sup>	24	20-30	-	-	-
EcN $\Delta$ 19 <sup>b</sup>	0.5	15-20	-	-	-
EcN $\Delta$ 22 <sup>b</sup>	< 0.005	15-20	-	-	-

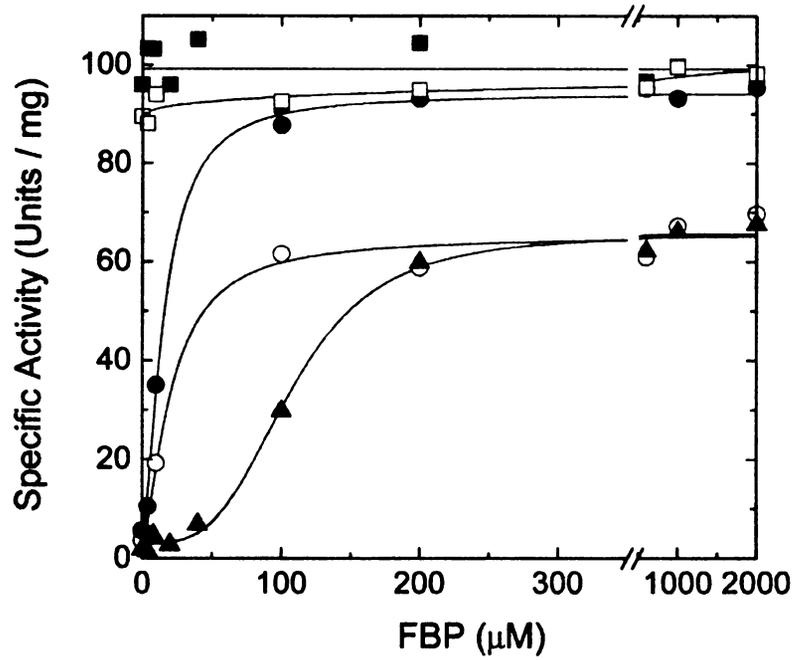
<sup>a</sup> Determined in crude extracts (168)

<sup>b</sup> Determination in crude extracts.

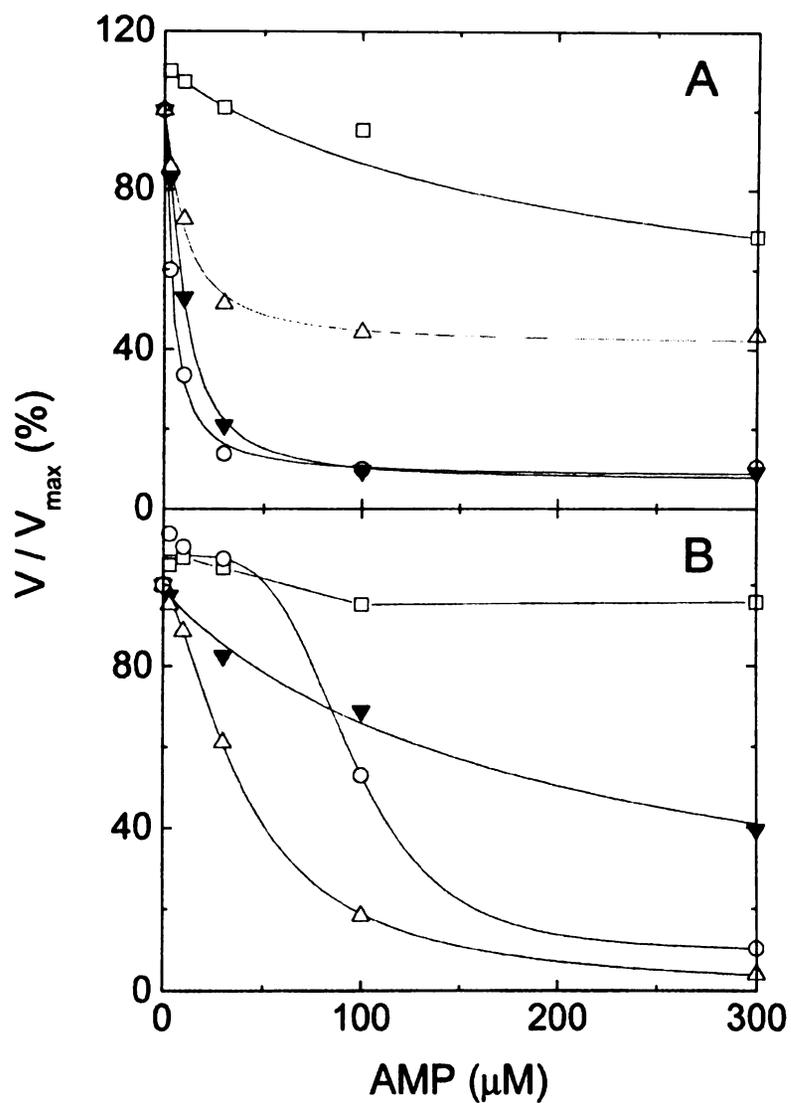


**Figure 3.2. N-terminal *E. coli* ADP-Glc PPase sequence and predicted secondary structures.** Residues 1 to 18 are predicted to be a loop. Ser<sup>19</sup> is predicted to be at the N-terminal end of a  $\beta$  strand in which Leu<sup>22</sup> resides. The Gly-rich loop (in light grey) is similar to P-loop-like motif present in protein kinases and nucleotide-binding sites.

Following that  $\beta$ -strand is a glycine-rich loop similar to the P-loop motif present in protein kinases and nucleotide binding sites. The 19- and 22-residue N-terminal deletions may cause destabilization of the local secondary structure (including the glycine-rich loop) that propagates to other important regions in the overall protein structure. Kinetic analysis of EcN $\Delta$ 3, EcN $\Delta$ 7, EcN $\Delta$ 11 and EcN $\Delta$ 15 showed that removal of up to 7-residues from the N-terminus does not significantly alter the FBP-activation properties of the enzyme. EcN $\Delta$ 3 and EcN $\Delta$ 7 mutants behaved in ways similar to the wild-type enzyme, which is activated 20- to 30-fold by the specific effector (Fig. 3.3 and Table 3.1). On the other hand, EcN $\Delta$ 11 behaved as reported previously (187, 188) and as EcN $\Delta$ 15. Both mutant enzymes were highly active even in the absence of FBP, showing almost no activation (Fig. 3.3 and Table 3.1). Additionally, EcN $\Delta$ 15 was tested for its ability to be inhibited by AMP (Fig. 3.4). Because it has been demonstrated that inhibition of the *E. coli* ADP-Glc PPase by AMP requires the presence of at least traces of FBP in the medium (44), inhibition kinetics were performed at two different activator concentrations: 30 and 130  $\mu$ M (Fig. 3.4.A and .B, respectively). A similar lack of sensitivity of EcN $\Delta$ 15 for this inhibitor was also exhibited by the EcN $\Delta$ 11 mutant (187, 188). Results suggested that deletion of 11- or 15- residues from the N-t induces spatial arrangements favoring a more active conformation of the enzyme, and possibly disrupting the AMP site.



**Figure 3.3. FBP-activation of wild-type and the N-terminal truncated *E. coli* ADP-Glc PPases.** (○) Wild type, (●) NΔ3, (▲) NΔ7, (■) NΔ11, (□) NΔ15. Saturation plots were determined by the pyrophosphorolysis reaction.



**Figure 3.4. AMP inhibition kinetics of wild-type and mutants *E. coli* ADP-Glc PPases.** All determinations were obtained by synthesis assay in presence of (A) 30 or (B) 130  $\mu\text{M}$  FBP. (o) Wild type, ( $\square$ )  $\text{N}\Delta 15$ , ( $\Delta$ )  $\text{EcC}\Delta 108 + \text{Ec}_{328-431}$ , ( $\diamond$ )  $\text{EcN}\Delta 15\text{C}\Delta 108 + \text{Ec}_{328-431}$ .

### 3.4.2. Characterization of EcN $\Delta$ 15-C $\Delta$ 108 + Ec<sub>328-431</sub>

To evaluate the opposite effects on FBP activation caused by an N-terminal deletion (187, 188) and by a nick separating the catalytic domain and the C-t as described in Chapter 2, EcN $\Delta$ 15C- $\Delta$ 108 + Ec<sub>328-431</sub> were co-overexpressed and purified as described in section 3.2. As illustrated in Table 3.2, after the *Green A* affinity chromatography the co-expressed proteins were purified to at least 90%. The EcN $\Delta$ 15C- $\Delta$ 108+Ec<sub>328-431</sub> co-expression product showed a specific activity comparable to that of the EcC $\Delta$ 108 + Ec<sub>328-341</sub> (see Chapter 2), the EcN $\Delta$ 15, and the wild-type enzymes (Table 3.2). Two protein bands of the expected molecular sizes (*i.e.*, ~35 kDa and ~12 kDa) were observed in SDS-PAGE and western blot (data not shown) of purified EcN $\Delta$ 15C- $\Delta$ 108 + Ec<sub>328-431</sub>. Interestingly, the kinetic properties of this co-expressed protein were comparable to those of the wild-type enzyme regarding affinity for the substrate ATP, affinity for the activator FBP and inhibition by AMP (Table 3.3).

Saturation kinetics for the FBP activation (Fig. 3.5) and AMP inhibition (Fig. 3.4) indicated that the EcN $\Delta$ 15C- $\Delta$ 108 + Ec<sub>328-431</sub> co-expressed enzyme recovered the sensitivity for the allosteric effectors that was respectively lost or decreased in the EcN $\Delta$ 15 or EcC- $\Delta$ 108 + Ec<sub>328-431</sub> mutant enzymes. Thus, the opposite effects in FBP-activation caused by an N-terminal deletion and by a nick separating the C-t and the catalytic domain seemed to compensate when these structural modifications were combined in the same protein. Assuming that an allosteric enzyme in solution is present as a mixture of conformations with different levels of activities, a 15 N-terminal residue

**Table 3.2. Specific activities of wild-type and mutants *E. coli* ADP-Glc PPases. Specific activities were determined in purified samples. Purity of the samples was estimated from SDS-PAGE gels.**

Sample	Specific Activity (U/mg)		Purity (%)
	Pyrophosphorolysis	Synthesis	
Wild-type <sup>a</sup>	131	54	90
EcCA108 <sup>a</sup>	< 0.001	<0.0001	50-60
EcCA108+ Ec328-431 <sup>a</sup>	132	43	>95
EcNA15	105	62	>95
EcNA15CA108+ Ec328-431	86	30	90

<sup>a</sup> Data obtained from *Chapter 2*, Table 2.1.1. (EcCA108 is same as Ec1-323)

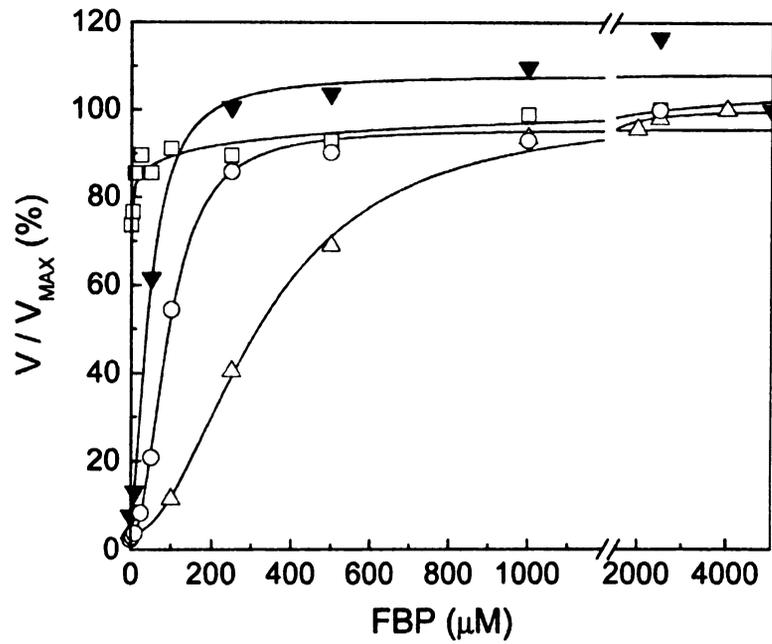
**Table 3.3. Kinetic parameters of wild-type and mutants *E. coli* ADP-Glc PPases.** All determinations were obtained by synthesis assay.

Enzyme	ATP		FBP activation		AMP inhibition			
	$S_{0.5}$ ( $\mu\text{M}$ )	$n_H$	$A_{0.5}$ ( $\mu\text{M}$ )	$n_H$	30 $\mu\text{M}$ FBP		130 $\mu\text{M}$ FBP	
					$I_{0.5}$ ( $\mu\text{M}$ )	$n_H$	$I_{0.5}$ ( $\mu\text{M}$ )	$n_H$
Ec wild type <sup>a</sup>	301.1 $\pm$ 14.1	1.9	92.1 $\pm$ 0.5	2.1	3.8 $\pm$ 0.4	1.1	94.0 $\pm$ 9.9	2.7
EcCA108 + Ec <sub>328-431</sub> <sup>a</sup>	1410.2 $\pm$ 51.4	2.9	327.4 $\pm$ 15.5	2.0	9.7 $\pm$ 1.9	1.2	40.7 $\pm$ 2.6	1.6
EcNA15CA108 + Ec <sub>328-431</sub>	494.6 $\pm$ 33.4	1.6	45.9 $\pm$ 8.1	1.6	9.4 $\pm$ 0.7	1.6	206 $\pm$ 24	0.9
EcNA15	nd	nd	6.9 $\pm$ 0.6	2.0	ni.	-	ni	-

nd: not determined

ni: no inhibition was observed when assayed up to 300  $\mu\text{M}$  concentration of AMP.

<sup>a</sup>: data obtained from *Chapter 2*, Table 2.2 (EcCA108 is same as Ec<sub>1-323</sub>)



**Figure 3.5. FBP activation kinetics of wild-type and mutants *E. coli* ADP-Glc PPases.** All determinations were obtained by synthesis assay. (○) Wild type, (□) N $\Delta$ 15, (△) EcC $\Delta$ 108 + Ec<sub>328-431</sub>, (□) EcN $\Delta$ 15C $\Delta$ 108 + Ec<sub>328-431</sub>.

truncation may induce the same structural change as the allosteric activator, which drives the equilibrium between enzyme species towards the more active one. In contrast, the structure arrangement generated when the catalytic domain and the C-t are co-expressed as individual polypeptides may favor a less active conformation. Combination of these two modifications in EcN $\Delta$ 15-C $\Delta$ 108 + Ec<sub>328-431</sub> might have compensated their opposite kinetic effects, making this mutant enzyme more similar to the wild-type enzyme. It is noteworthy that this mutant was also sensitive to AMP inhibition, eliminating the above suggestion of a possible AMP site disruption caused by deletion of 15 N-terminal residues. It is more likely that EcN $\Delta$ 15 is locked in a more active conformation that can not be reversed by the inhibitor.

Integration of these biochemical results involving the N-terminal extension with actual structural data remains to be done since the only ADPGlc PPase crystal structures available lack a solved diffraction pattern corresponding to the first 11 N-terminal amino acids. The structural data points to the flexible character of this region of the enzyme which agrees with its suggested regulatory role as “allosteric switch”. It would be very valuable to obtain direct structural information of all the mutants analyzed in the present work. They would provide with a series of snapshots of the various conformations achieved by the enzyme in the different activated states, allowing for a better understanding of the allosteric regulatory mechanism.

## CHAPTER 4

### Molecular architecture of the glucose 1-phosphate site in

### ADP-glucose pyrophosphorylases<sup>3</sup>

---

<sup>3</sup>This work has been accepted for publication: C. M. Bejar, X. Jin, M. A. Ballicora, and J. Preiss (2006)

Molecular architecture of the glucose 1-phosphate site in ADP-glucose pyrophosphorylases. *J. Biol. Chem.*  
*In Press.*

#### 4.1. ABSTRACT

The work presented in this chapter was intended to answer questions regarding the three-dimensional organization of the *Escherichia coli* ADP-Glc PPase and to use that information as a tool for the identification of potentially important residues in the active site. Thus, a homology modeling approach to generate the monomeric structure of this enzyme complexed with ADP-Glc and the results from a detailed study of the substrate (glucose 1-phosphate) binding site are presented here. A set of amino acids in the model has been identified to be in close proximity to the glucose moiety of the ADP-Glc ligand and found to be conserved within the ADP-Glc PPase family. The role of these amino acids, Glu<sup>194</sup>, Ser<sup>212</sup>, Tyr<sup>216</sup>, Asp<sup>239</sup>, Phe<sup>240</sup>, Trp<sup>274</sup>, and Asp<sup>276</sup>, was studied by site directed mutagenesis, through the characterization of the kinetic properties and thermal stability of the designed mutants. All purified alanine mutants had one or two orders of magnitude lower apparent affinity for Glc1P, compared to the wild type, indicating that the selected set of amino acids plays an important role in their interaction with the substrate. Those amino acids were subjected to conservative and non-conservative mutations to investigate the effect of size, hydrophobicity, polarity, aromaticity, or charge on the affinity for Glc1P. In this study, the architecture of the Glc1P binding site is characterized. The homology model overlaps with the Glc1P site of other PPases, such as the *Pseudomonas aeruginosa* dTDP-Glc PPase and the *Salmonella typhi* CDP-Glc PPase. Therefore, the data reported here may have implications for other members of the nucleotide-diphospho-glucose PPases family.

## 4.2. INTRODUCTION

The first ADP-Glc PPase crystal structure available was that of the homotetrameric *Solanum tuberosum* (potato tuber) small subunit in its allosterically inhibited form at a resolution of 2.1 Å (170), reported in 2005. The same study also reported the structural determination of the enzyme complexed with either ATP or ADP-Glc at 2.6 and 2.2 Å, respectively. Attempts to obtain information on the *E. coli* enzyme structure through X-ray crystallography were unsuccessful. The potato tuber small subunit has only about 33% sequence identity with the *E. coli* enzyme, but the similar predicted secondary structure profiles, together with available biochemical data, suggest that they share a common three-dimensional fold (8).

Previous chemical modification (172) and site directed mutagenesis studies (174) on the *E. coli* ADPGlc PPase enzyme identified Lys<sup>195</sup> as an important residue for Glc1P interaction. Replacement by other amino acids generated 100-10,000-fold increases in the  $S_{0.5}$  for this substrate while showing all other kinetic constants at wild type levels. Later, Fu *et al.* reported similar results analyzing the homologous Lys<sup>198</sup> in the potato tuber catalytic subunit (177). The proposed role of this amino acid is to form an ionic bond between the  $\epsilon$ -amino with the negatively charged phosphate of the Glc1P. Results with hexose 1-phosphate analogues, differing from the Glc1P in their hydroxyl groups, suggested that other residues in the active site participate in substrate binding (177).

The objective in this work has been to obtain structural information on the *E. coli* ADP-Glc PPase by building a homology model and to probe a set of highly conserved residues in the N-terminus domain possibly involved in Glc1P binding. We studied the role of Glu<sup>194</sup>, Ser<sup>212</sup>, Tyr<sup>216</sup>, Asp<sup>239</sup>, Phe<sup>240</sup>, Trp<sup>274</sup>, and Asp<sup>276</sup> by means of site directed mutagenesis, kinetic characterization of the mutant enzymes and their thermal stability. All residues were replaced by alanine and other amino acids to evaluate the importance of size, charge or hydrophobicity on the effects observed in substrate interaction.

Because these residues are highly conserved among ADP-Glc PPases, it was of interest to address the question of whether they are present in other pyrophosphorylases that use Glc1P as substrate. The observations made by comparison of the putative Glc1P site from our *E. coli* ADP-Glc PPase model and the reported crystal structures of two pyrophosphorylases, the *P. aeruginosa* dTDP-Glc PPase RmlA (164) and the *S. typhi* CDP-Glc PPase (169), have led us to propose that the results presented here have implications beyond the family of the ADP-Glc PPases.

### 4.3. MATERIALS AND METHODS

#### 4.3.1. Materials

Oligonucleotides were synthesized and purified by the Macromolecular Facility at Michigan State University. [<sup>32</sup>P]PPi was purchased from Perkin Elmer Life Sciences and [<sup>14</sup>C]Glc1P from ICN Pharmaceuticals Inc. NaPPi, ATP, ADP-Glc, AMP, and inorganic pyrophosphatase were purchased from Sigma Chemical Co. *Pfu* DNA polymerase was purchased from Stratagene (La Jolla, CA, USA). All other reagents were of the highest quality available.

#### 4.3.2. Homology modeling

Comparative (homology) modeling of the *E. coli* ADP-Glc PPase (residues 12-431) was carried out with the program Modeller6v1 (203-205) using the atomic coordinates of the *S. tuberosum* ADP-Glc PPase small subunit chain B complexed with ADP-Glc (Protein Data Bank code 1YP2, (170)) as template. Sequence alignment was performed manually to match functionally conserved residues, predicted secondary structures, and hydrophobicity profiles. Secondary structures were predicted using the PHD (<http://www1.embl-heidelberg.de/Services/sander/predictprotein/>) and the PSI-PRED programs (<http://bioinf.cs.ucl.ac.uk/psipred/>). The models were assessed by the VERIFY\_3D program (206, 207) ([http://shannon.mbi.ucla.edu/DOE/Services/Verify\\_3D/](http://shannon.mbi.ucla.edu/DOE/Services/Verify_3D/)).

### 4.3.3. Multiple sequence alignment

A multiple sequence alignment was generated using the server *ClustalW* (<http://www.ebi.ac.uk/clustalw>) with representative ADP-Glc PPases belonging to different bacterial and plant taxa. The sequences used were from *Escherichia coli B* (NCBI accession number P0A6V1) (208), *Agrobacterium tumefaciens* (P39669) (209), *Synechococcus sp. WH 8102* (NP\_897211) (210), *Thermotoga maritima* (Q9WY82) (211), *Streptococcus pneumoniae* (Q97QS7) (212), *Vibrio cholerae* (Q9KLP4) (213), *Clostridium cellulolyticum* (Q9L385) (direct submission), *Geobacillus stearothermophilus* (O08326) (120), *Mycobacterium tuberculosis* (O05314) (214), *Deinococcus radiodurans* (Q9RTR7) (215), *Anabaena sp PCC7120* (P30521) (145), *Arabidopsis thaliana* -APS1 small subunit- (P55228) (216), *Solanum tuberosum* -small subunit- (P23509) (147), *Zea mays* -small subunit, endosperm- (AAK69627) (217), *Chlamydomonas reinhardtii* -small subunit- (AAF75832) (218).

### 4.3.4. Site directed mutagenesis

Site directed mutagenesis was done by overlap extension PCR (219). The template was plasmid pMAB3, with the *E. coli* ADP-Glc PPase gene between *NdeI*-*SacI* sites, previously obtained in our lab (Ballicora *et al.*, unpublished). The flanking primers annealing with the T7 promoter and the *SacI* site (underlined) were 5'-TAATACGACTCACTATAGGG-3' and 5'-GATATCTGAATTCGAGCTC-3', respectively. Table 4.2 shows the overlapping primers for each mutant. The final PCR

products were gel purified, digested with *NdeI* and *SacI*, and subcloned to obtain the different pMAB3-single mutant plasmids. Plasmid pETEC-N $\Delta$ 15-D276N was obtained using pETEC-N $\Delta$ 15 (see *Chapter 3*), as template, with the T7 promoter and T7 terminator as flanking primers and the same mutated overlapping primers used for pMAB3-D276N (Table 4.1). All plasmids were sequenced at the MSU Genomics Facility to confirm incorporation of only the desired mutation.

#### **4.3.5. Bacterial strains and expression of the recombinant ADPGlc PPases**

*E. coli* AC70R1-504 cells that lack endogenous ADP-Glc PPase activity were used for expression of the wild type and pMAB3-mutant enzymes as described previously for pML10 (147). EcN $\Delta$ 15-D276N was expressed as EcN $\Delta$ 15 (see *Chapter 3*).

#### **4.3.6. Purification of pMAB3-single mutants**

One-liter cultures of AC70R1-504 cells transformed with the pMAB3-single mutant plasmids, or BL21(DE3) cells transformed with pETEC-N $\Delta$ 15-D276N, were grown in 25  $\mu$ g/mL kanamycin-Luria broth (1 liter) at 37 °C up to an  $A_{600}$  of 0.8. Induction was initiated by the addition of IPTG (1 mM final concentration), with subsequent incubation at 25 °C for 16 h. After induction, cells were harvested, and crude extracts were obtained as described previously (147). After centrifugation, the precipitate was resuspended in *buffer A* (50 mM Hepes pH 8.0, 5 mM MgCl<sub>2</sub>, 0.1 mM EDTA, 10% sucrose). The samples were individually applied onto a DEAE-Fractogel column (EMD

Chemicals), and eluted with a linear NaCl gradient (0-0.5 M). The active fractions were pooled and desalted. After this step, samples were 60-70% pure and suitable for performing kinetic analysis. Mutants E194A/Q/D, D276A/N, W274A, Y216F and D239N were resuspended in *buffer B* (*buffer A* plus 1.2 M ammonium sulfate), applied to a Phenyl-Superose (FPLC, Pharmacia) column equilibrated with *buffer B* and eluted with a 1.2-0.001 M linear gradient of ammonium sulfate. Further purification of the rest of the mutants and the wild type was performed by applying the DEAE pool samples to a Matrex™ gel *Green A* affinity chromatography column (Amicon Corporation, USA) and eluting with a linear gradient of 0-2 M NaCl. The purest fractions of each enzyme were pooled, desalted and concentrated, and after these steps the proteins were >95% pure, as assessed by SDS-PAGE (not shown).

#### **4.3.7. Protein methods**

Protein assay, electrophoresis (SDS-PAGE) and immunoblotting were performed using protocols previously described (167). Samples were desalted and concentrated with Centricon-30 devices (Amicon Inc.).

**Table 4.1. Complementary oligonucleotides used to introduce single mutations.** Site directed mutagenesis were performed by the overlap extension polymerase chain reaction. Triplets coding for the replacing amino acid are shown in bold.

<b>Mutation</b>	<b>Primer</b>	<b>5'-3' nucleotide sequence</b>	<b>Plasmid obtained</b>
E194A	<i>Forward</i>	CGAATTCGTTGCCAAA CCTGC	pMAB3-E194A
	<i>Reverse</i>	GCAGGTTTGGCAACGAATTCG	
E194D	<i>Forward</i>	CGAATTCGTTGACAAA CCTGC	pMAB3-E194D
	<i>Reverse</i>	GCAGGTTTGTCAACGAATTCG	
E194Q	<i>Forward</i>	CGAATTCGTTCAAAAA CCTGC	pMAB3-E194Q
	<i>Reverse</i>	GCAGGTTTTTTGAACGAATTCG	
S121A	<i>Forward</i>	TCTCTGGCGGCGATGGGTATC	pMAB3-S121A
	<i>Reverse</i>	GATACCCATCGCCGCCAGAGA	
S121V	<i>Forward</i>	TCTCTGGCGGTGATGGGTATC	pMAB3-S121V
	<i>Reverse</i>	GATACCCATCACCGCCAGAGA	
S121T	<i>Forward</i>	TCTCTGGCGACCATGGGTATC	pMAB3-S121T
	<i>Reverse</i>	GATACCCATTGGCGCCAGAGA	
S121Y	<i>Forward</i>	TCTCTGGCGTATATGGGTATC	pMAB3-S121Y
	<i>Reverse</i>	GATACCCATATACGCCAGAGA	
Y216A	<i>Forward</i>	ATGGGTATCGCCGCTTTGACGCC	pMAB3-Y216A
	<i>Reverse</i>	GGCGTCAAAGACGGCGATACCCAT	
Y216F	<i>Forward</i>	ATGGGTATCTTCGTCTTTGACGCC	pMAB3-Y216F
	<i>Reverse</i>	GGCGTCAAAGACGAAGATACCCAT	
D239A	<i>Forward</i>	TCCAGCCACGCCTTTGGCAAAG	pMAB3-D239A
	<i>Reverse</i>	CTTTGCCAAAGGCGTGGCTGGA	
D239N and EcNΔ15-D276N	<i>Forward</i>	TCCAGCCACAACCTTTGGCAAAG	pMAB3-D239N and pETEC- EcNΔ15- D276N
	<i>Reverse</i>	CTTTGCCAAAGTTGTGGCTGGA	
D239E	<i>Forward</i>	TCCAGCCACGAATTTGGCAAAG	pMAB3-D239E
	<i>Reverse</i>	CTTTGCCAAATTCGTGGCTGGA	
F240A	<i>Forward</i>	TCCAGCCACGACGCGGGCAAAGA	pMAB3-F240A
	<i>Reverse</i>	TCTTTGCCCGCGTCGTGGCTGGA	
F240M	<i>Forward</i>	AGCCACGACATGGGCAAAGAT	pMAB3-F240M
	<i>Reverse</i>	ATCTTTGCCCATGTCGTGGCT	
W274A	<i>Forward</i>	GAGCCGTACGCGCGCGATGTG	pMAB3-W274A
	<i>Reverse</i>	CACATCGCGCGCGTACGGCTC	
W274F	<i>Forward</i>	GAGCCGTACTTTTCGCGATGTG	pMAB3-W274F
	<i>Reverse</i>	CACATCGCGAAAGTACGGCTC	
W274L	<i>Forward</i>	GAGCCGTACTTTCGCGATGTG	pMAB3-W274L
	<i>Reverse</i>	CACATCGCGAACGTACGGCTC	
D276A	<i>Forward</i>	TACTGGCGCGCGGTGGGTACC	pMAB3-D276A
	<i>Reverse</i>	GGTACCCACCGCGCGCCAGTA	
D276N	<i>Forward</i>	TACTGGCGCAATGTGGGTACC	pMAB3-D276N
	<i>Reverse</i>	GGTACCCACATTGCGCCAGTA	
D276E	<i>Forward</i>	TACTGGCGCGAGGTGGGTACC	pMAB3-D276E
	<i>Reverse</i>	GGTACCCACCTCGCGCCAGTA	

#### 4.3.8. Enzymatic assays

*Assay A: Pyrophosphorolysis direction* - Formation of [<sup>32</sup>P]ATP from [<sup>32</sup>P]PPi was determined by the method of Morell *et al.*(123). The reaction was carried out for 10 min at 37 °C in a mixture that contained 50 mM Hepes pH 8.0, 10 mM MgCl<sub>2</sub>, 1.5 mM [<sup>32</sup>P]PPi (1500-2500 dpm/nmol), 4 mM ADP-Glc, 4 mM NaF, 2 mM FBP and 0.05 mg/mL BSA, plus enzyme in a total volume of 0.25 mL.

*Assay B: Synthesis direction* – Formation of ADP-[<sup>14</sup>C]Glc from [<sup>14</sup>C]Glc1P was determined by the method of Yep *et al.* (193). The reaction was carried out for 10 min at 37 °C in a mixture that contained [<sup>14</sup>C]Glc1P (~400 dpm/nmol), ATP, MgCl<sub>2</sub> and FBP in varying concentrations according to the mutant enzyme assay, 50 mM Hepes pH 8.0, 1.5 units/mL pyrophosphatase, and 0.2 mg/mL BSA, plus enzyme in a total volume of 0.20 mL. One unit of enzymatic activity is one μmol of product, either [<sup>32</sup>P]ATP or ADP-[<sup>14</sup>C]Glc, formed per min at 37 °C.

#### 4.3.9. Kinetic characterization

Kinetic data were plotted as specific activity (U mg<sup>-1</sup>) vs. substrate or effector concentration. Kinetic constants were acquired by fitting the data to the Hill equation with a nonlinear least square formula using the program Origin™ 5.0. Hill plots were used to calculate the Hill coefficient  $n_H$  and the kinetic constants that correspond to the substrate or activator concentrations giving 50% of the maximal velocity ( $S_{0.5}$ ) or activation ( $A_{0.5}$ ).

#### **4.3.10. Thermal stability**

Enzyme samples were in buffer *A* supplemented with BSA to 1 mg/mL in a final volume of 100  $\mu$ L. Half of the sample (50  $\mu$ L) was incubated in a water bath equilibrated at 60°C for 5 min and placed on ice immediately after. The remaining 50  $\mu$ L were kept on ice as control. The enzyme activity for both, the heat-treated and control samples, were determined in the ADP-Glc synthesis direction, as described in *Assay B*.

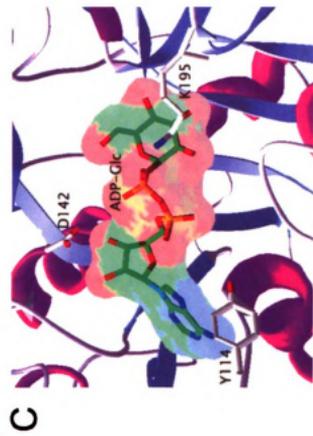
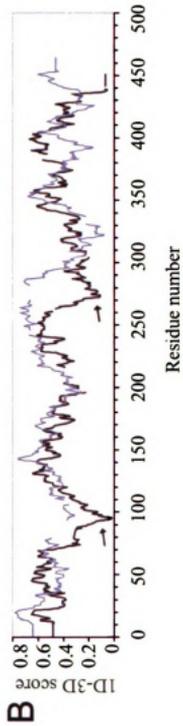
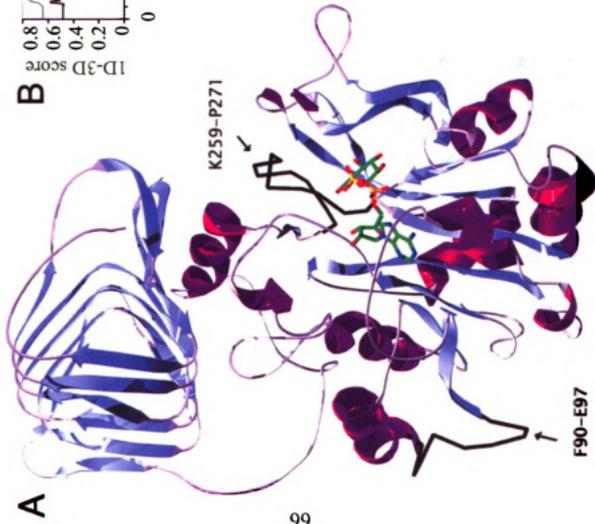
## 4.4. RESULTS

### 4.4.1. Homology modeling

We obtained a three-dimensional model of the *E. coli* ADP-Glc PPase by comparative modeling, using the coordinates of the recently solved crystal structure of the potato tuber small subunit ADP-Glc PPase (PDB code 1YP2) as template, as described in *Experimental Procedures* (Fig. 4.1.A). While modeling is generally guaranteed to be successful if residue identity is >40%, for lower percentages errors can be reduced employing an accurate sequence alignment (220-222). Our two enzymes shared only 33% residue identity, therefore the alignment was manually edited incorporating information such as conservation of functional residues and prediction of secondary structures.

Using Modeller6v1, we generated 143 models after several iterative refinements of the alignment to accommodate gaps, deletions and insertions of the query sequence respect to the template in the best possible way. We assessed the models with the program VERIFY\_3D (206, 207), as described in *Experimental Procedures*, which evaluates the compatibility of a given residue (1D) in a certain environment (3D). A score below zero for a given residue means that the conformation adopted by that residue in the model is not compatible with its surrounding environment. In our study, we considered only those models with all 1D-3D averaged scores above zero and, among them, we chose the one with most similar profile to that generated by the template's

**Figure 4.1. Structural model of *E. coli* ADP-Glc PPase.** **A.** Cartoon representation of the monomer. The N-terminus presents a Rossmann-like fold and holds the ADP-Glc molecule in the active site (carbons are green, all other atoms are colored by type). The C-terminus adopts a  $\beta$ -helix fold and is connected to the N-terminus by a long loop. Loops of low structural reliability, comprising residues F90-E97 and K259-P271, are colored in black and pointed by arrows. **B.** Verify\_3D profile obtained from the assessment of the *E. coli* ADP-Glc PPase structural model. The profile in black corresponds to our model and the one in grey to the template crystal structure. Gaps in the template profile correspond to gaps in the sequence alignment with the *E. coli* enzyme and to stretches of amino acids not solved in the crystal structure. The two big depressions in the *E. coli* profile pointed by the arrows are regions of low structural reliability and correspond to the F90-E97 and K259-P271 loops. **C.** Close-up view of the modeled active site, with a bound ADP-Glc molecule (carbons in green). Asp<sup>142</sup>, Tyr<sup>114</sup> and Lys<sup>195</sup> (white carbons), which are involved in catalysis (167), binding of ATP (173) and binding of Glc1P (174), respectively, are in the active site and close to the ligand.



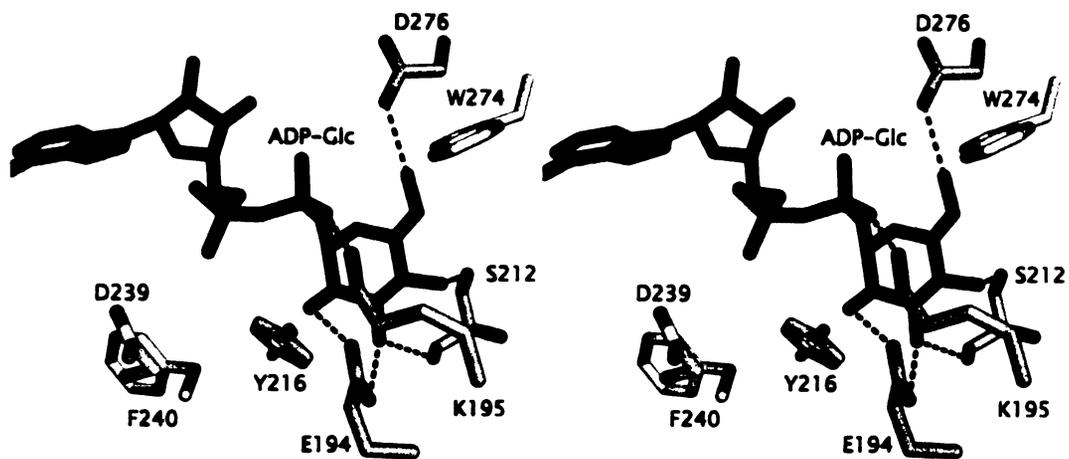
crystal structure (Fig. 4.1.B). The two profiles followed the same general trend except for two specific regions, both corresponding to residues located in or adjacent to loops that are not present in the template structure (indicated by arrows in Fig. 4.1.A). The first, encompassing residues F90-E97 in the *E. coli* enzyme, aligns with a region in the potato tuber enzyme that is disordered in the crystal structure. The second loop, containing amino acids K259-P271, is an insertion in the bacterial enzyme. Therefore, the final conformation of these two loops in the model, which might also affect immediately adjacent secondary structures, accounted for the differences with the template structure's profile. According to the model, these loops are not part of the active site nor do they contain important conserved residues.

In agreement with the biochemical results obtained before and described in *Chapter 2*, the modeled monomer shows a two domain structural organization (Fig. 4.1.A). The N-terminus of ~300 residues, presents a  $\beta$ - $\alpha$ - $\beta$  motif arranged in an open twisted  $\beta$ -sheet, surrounded by  $\alpha$ -helices. It resembles the Rossmann fold typically present in nucleotide-binding domains (223). Residues important for catalysis, Asp<sup>142</sup> (167), and for substrates binding, Tyr<sup>114</sup> for ATP (173), and Lys<sup>195</sup> for Glc1P (174), are located in the active site pocket in close proximity to the ADP-Glc molecule (Fig. 4.1.C), observations that further validate the quality of our model. The C-terminus is a separate domain folded as a  $\beta$ -helix and linked to the N-terminus by a long loop. The two domains are in intimate interaction through extensive hydrophobic contacts, which supports the requirement of a full-length polypeptide to obtain normal enzymatic activity and regulation (see *Chapter 2*).

#### 4.4.2. Selection of residues for analysis

The three-dimensional model complexed with ADP-Glc shows the ligand placed in a well-defined pocket in the active site (Fig. 4.1.A) and several direct interactions between the ligand and the enzyme are evident (Fig. 4.2). Lys<sup>195</sup> makes a salt bridge with the glucose phosphate interaction that has been biochemically probed by Hill *et al.* (174) in the *E. coli* ADP-Glc PPase and by Fu *et al.* (177) analyzing the homologous residue (Lys<sup>198</sup>) in the potato tuber enzyme. Additionally, the hydroxyl groups of the glucosyl moiety of the ligand are involved in a complex net of hydrogen bonds with the enzyme. Side chains of Glu<sup>194</sup>, Asp<sup>276</sup> and Ser<sup>212</sup>, and the backbone of the latter, participate in such interactions.

We performed a multiple sequence alignment using the catalytic subunits of fifteen ADP-Glc PPases from several sources, each of them representative of a different taxonomic group. Figure 4.3 depicts part of the aligned sequences, comprising residues located in and around the putative Glc1P binding domain in the N-terminus of the protein. The residues that, in the model, appear interacting through hydrogen bonds with the glucosyl moiety of the ligand are absolutely conserved among all ADP-Glc PPases analyzed, suggesting that they are involved in a conserved role, such as substrate binding. According to our structural model, other conserved residues in this region are also located in the substrate-binding pocket. Based on our observations, we selected Tyr<sup>216</sup>, Asp<sup>239</sup>, Phe<sup>240</sup> and Trp<sup>274</sup> to be characterized together with Glu<sup>194</sup>, Ser<sup>212</sup> and Asp<sup>276</sup>.



**Figure 4.2. *E. coli* ADP-Glc PPase- substrate interaction.** Stereo representation of the putative Glc1P binding site showing the residues studied in this work (white carbon) and their proposed hydrogen bonding interactions (dashed blue lines) with the bound ADP-Glc molecule (green carbons).

**Figure 4.3. Sequence alignment of the *E. coli* ADP-Glc PPase and its homologues.** The primary structures of representative bacterial ADP-Glc PPases and catalytic subunits from photosynthetic organisms were aligned. The region showed here encompasses residues located at and near the putative Glc 1P binding site, according to our homology model. Sources abbreviations are: Eco, *Escherichia coli* B; Atum, *Agrobacterium tumefaciens*; Synech, *Synechococcus* sp. WH 8102; Tmar, *Thermotoga maritima*; Spneu, *Streptococcus pneumoniae*; Vcho, *Vibrio cholerae*; Ccell, *Clostridium cellulolyticum*; Bstear, *Geobacillus stearothermophilus*; Mtub, *Mycobacterium tuberculosis*; Drad, *Deinococcus radiodurans*; Ana, *Anabaena* sp. PCC7120; Atha, *Arabidopsis thaliana*; Stub, *Solanum tuberosum* -small subunit-; Zmay, *Zea mays* -small subunit, endosperm-; Crein, *Chlamydomonas reinhardtii* -small subunit-. NCBI accession numbers and references are detailed in MATERIALS AND METHODS. Prediction of secondary structure is depicted on top (E:  $\beta$ -strand; C: coil; H:  $\alpha$ -helix). A consensus sequence at the bottom has also been included. Residues studied here are highlighted in grey.

<b>Bacterial</b>	EEEEEEC-----CCCCCCCCCCEEEEEEEEEECHHHHHHHHH
<i>Eco</i>	190 IEFVEKP-----ANPPSMPNDPSKSLASMGIVVFDADYLYELLE 229
<i>Atum</i>	183 IDFI EKP-----ADPPGIPGNEGALASMGIVVFHTKFLMEAVRR 222
<i>Synech</i>	174 KEFREKPKGDSLLEMAVDTSRFGLSANSAKERPYLASMGIVVFSRDTLFDLL-- 225
<i>Tmar</i>	174 VDFE EKP-----AKP-----RSNLASLGIYVFNIEFLKVKVLE 206
<i>Spneu</i>	175 VEFEEKP-----AQP-----KSTKASMGIVIFDWQRLRNMLVA 207
<i>Vcho</i>	173 TCFVEKP-----ADPPCIPNRPDHSLASMGIVIFNMDVLKALTE 212
<i>Ccell</i>	176 YEFEEKP-----KNP-----KSTLASMGIVIFTWSTLREYLIK 208
<i>Bstear</i>	175 VEFEEKP-----AEP-----KSNLASMGIVIFNWPLLKQYLQI 207
<i>Mtub</i>	155 RSFVEKP-----LEPPGTPDDPDFTFVSMGNVIFTTKVLIDAIRA 194
<i>Drad</i>	178 TEFHEKV-----PDPPTIPGQADLSLTSMGNVIFSRRALEELLEA 217
<i>Ana</i>	173 IDFSEKPKGEALTKMRVDTTVLGLTPEQAASQPIASMGIVVFKKDVLIKLLK- 225
<b>Plant</b>	
<i>Atha</i>	262 IEFAEKPKGEHLKAMKVDTTILGLDDQRAKEMPFIASMGIVVSRDVMLDLLR- 314
<i>Stub</i>	193 IEFAEKPKGEQLQAMKVDTTILGLDDKRAKEMPFIASMGIVVISKDVMLNLLR- 245
<i>Zmay</i>	217 IEFAEKPKGEQLKAMVDTTILGLDDVRAKEMPFIASMGIVVFSKDVMLQLLR- 269
<i>Crein</i>	257 IEFAEKPKGEALTKMRVDTGILGVDPATAAAKPIASMGIVVMSAKALRELL- 309
CONSENSUS	* ** . : * * * . : . :

<b>Bacterial</b>	CCCCCCCCCHHHHHHHHHHCC--CEEEEECEEEEEEECC--EEEECCCH
<i>Eco</i>	230 DDRDENSSHDFGKDLIPKITEAG--LAYAHFPLSCVQSDPDAE-PYWRDVGTL 280
<i>Atum</i>	223 DAADPTSSRDFGKDIIPYIVEHG--KAVAHRFADSCVRSDFEHE-PYWRDVGTI 273
<i>Synech</i>	226 --DSNPGYKDFGKEVIPEALKRGD-KLKSIVFDD-----YWEDIGTI 264
<i>Tmar</i>	207 DENDPNSSHDFGKDVIPRILRENLSLYAFRFDG-----YWRDVGTL 248
<i>Spneu</i>	208 AEKSKVGMSDFGKNVIPNYLESGE-SVYAYEFSG-----YWKDVGTI 248
<i>Vcho</i>	213 DAEIEQSSHDFGKDVIPKLIATG--SVFAYSFCSG--KGRVARD-CYWRDVGTI 261
<i>Ccell</i>	209 DNECSDSVNDFGKNIPAMLDGDK-SMWAYQYSG-----YWRDVGTI 249
<i>Bstear</i>	208 DNANPHSSHDFGKDVIPMLLREKK-RPFAYPFEF-----YWKDVGTV 248
<i>Mtub</i>	195 DADDHSDHDMGGDIVPRLVADG--MAAVYDFSDNEVPGATDRDRAYWRDVGTL 266
<i>Drad</i>	218 SISQETGYDFGHNVIPRALSDGY-HVQAYDFHKNPIPGQ-ERPNTYWRDVGTL 269
<i>Ana</i>	226 --EALERT-DFGKEIIP-DAAKDH-NVQAYLFDD-----YWEDIGTI 262
<b>Plant</b>	
<i>Atha</i>	315 --NQFPGANDFGSEVIPGATSLGL-RVQAYLYDG-----YWEDIGTI 253
<i>Stub</i>	246 --DKFPGANDFGSEVIPGATSLGM-RVQAYLYDG-----YWEDIGTI 284
<i>Zmay</i>	270 --EQFPEANDFGSEVIPGATSIGK-RVQAYLYDG-----YWEDIGTI 308
<i>Crein</i>	310 --NRMPGANDFGNEVIPGAKDAGF-KVQAFAFDG-----YWEDIGTV 348
CONSENSUS	* * : : * . : ** * : * :

#### 4.4.3. Expression and purification of pMAB3-single mutants

All selected amino acids were mutated to alanine to analyze their potential role in Glc1P interaction. We performed additional mutations to investigate if the observed effect on the affinity for Glc1P was due to their shape, size, charge or aromaticity. *E. coli* wild type and mutant ADP-Glc PPases were expressed and purified as described in *Experimental Procedures*. They had the expected molecular weights in SDS-PAGE and they were recognized by the anti AC70R1 *E. coli* ADP-Glc PPase antibody in immunoblots (not shown). Y216A either failed to be expressed or rendered the protein completely susceptible to proteolysis since no band  $\geq 10$  kDa was detected by immunoblotting in either the soluble or the insoluble fractions of the expression cell lysates. After the first chromatographic step, all enzymes were 60-70% pure and suitable for kinetic characterization assays. An additional chromatographic step yielded >95% pure enzymes, which allowed for the proper determination of their specific activities.

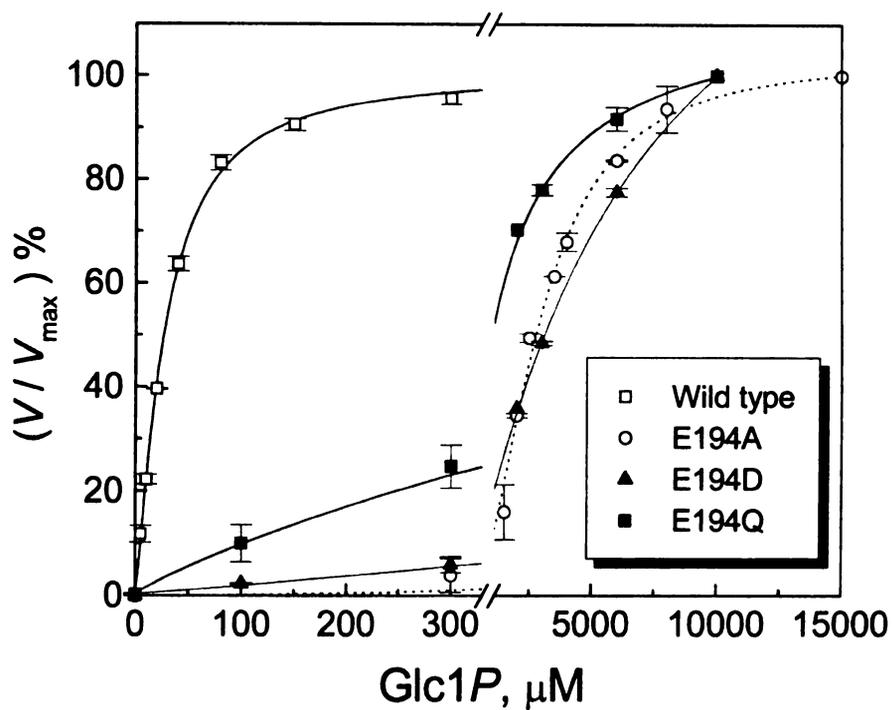
#### 4.4.4. Kinetic characterization

The kinetic characteristics of the mutant enzymes were compared to those of the wild type. All alanine mutations decreased the apparent affinity of the enzyme for the substrate Glc1P, as  $S_{0.5}$  values for all the mutants were one or two orders of magnitude larger than that of the wild type (Table 4.2). The most important increments in this kinetic parameter were observed with mutations in mutations on Glu<sup>194</sup> and Ser<sup>212</sup>. Glc1P saturation curves obtained for the Glu194 mutants are shown in Fig.4 as an example to

**Table 4.2. Comparison of specific activities and apparent affinity for the substrate Glc1P of the *E. coli* wild type and mutant enzymes. Determinations were obtained with pure enzymes in the synthesis direction of the reaction by the Yep *et al.* method (193), as described in *MATERIALS AND METHODS*.**

Enzyme	$k_{\text{cat}}$		-fold		$S_{0.5}$ ( $\mu\text{M}$ )		-fold		$k_{\text{cat}}/K_m$ ( $\text{s}^{-1} \cdot \text{mM}^{-1}$ )
	( $\text{s}^{-1}$ )								
Wild type	370.0	$\pm 14.4$	1.0		17	$\pm 2$	1		21765
E194A	15.43	$\pm 0.07$	24.0		2812	$\pm 127$	165		6
E194D	92.7	$\pm 11.7$	4.0		6587	$\pm 1160$	388		14
E194Q	80.7	$\pm 2.0$	4.6		1441	$\pm 369$	85		56
S212A	371.2	$\pm 4.1$	1.1		241	$\pm 34$	14		1440
S212V	22.4	$\pm 1.9$	16.5		6416	$\pm 886$	377		4
S212T	179.0	$\pm 0.7$	2.1		4659	$\pm 274$	274		38
S212Y	1.6	$\pm 0.1$	231.2		90	$\pm 3$	5		18
Y216F	29.0	$\pm 1.0$	12.8		785	$\pm 39$	46		37
D239A	32.7	$\pm 0.2$	11.3		524	$\pm 126$	31		62
D239E	347.7	$\pm 15.9$	1.1		169	$\pm 27$	10		2146
D239N	169.0	$\pm 23.7$	2.2		264	$\pm 38$	16		640
F240A	171.7	$\pm 59.7$	2.2		204	$\pm 10$	12		842
F240M	487.0	$\pm 28.3$	0.8		122	$\pm 4$	7		3992
W274A	384.0	$\pm 23.4$	1.0		367	$\pm 16$	22		1046
W274F	266.0	$\pm 8.0$	1.4		50	$\pm 3$	3		5320
W274L	247.5	$\pm 0.6$	1.5		525	$\pm 32$	31		463
D276A	0.36	$\pm 0.01$	1027.8		1706	$\pm 127$	100		0.2
D276N	0.37	$\pm 0.03$	1000.0		1447	$\pm 151$	85		0.3
D276E	112.7	$\pm 4.5$	3.3		416	$\pm 6$	24		275
K195Q <sup>b</sup>	193.3	$\pm 50$	1.9		16700	$\pm 380$	982		12

<sup>b</sup>Data from Ref. (174)



**Figure 4.4. Steady-state kinetic measurement for Glc1P dependence for wild type and E194A, E194D and E194Q mutant enzymes.** Initial velocities were determined in the ADP-Glc synthesis direction using Assay B. For wild type ( $\square$ ), E194A ( $\circ$ ), E194D ( $\blacktriangle$ ) and E194Q ( $\blacksquare$ ),  $V_{\text{max}}$  were 111.0, 4.6, 27.8 and 24.2 U/mg, respectively. Reactions for each enzyme were performed in presence of saturating concentrations of ATP,  $\text{MgCl}_2$  and FBP.

illustrate the shift in the  $S_{0.5}$  between wild type and Glu<sup>194</sup> mutant enzymes. The E194A mutant showed a 165-fold increase (Table 4.2) compared to that of the wild type. Therefore, we made substitutions to aspartic acid and glutamine to evaluate the importance of the charge and the side chain size in such effect. Mutation to glutamine increased the  $S_{0.5}$  for Glc1P 85-fold, pointing out the importance of the negative charge for substrate binding. However, mutation to aspartic acid, which also bears a negative charge, caused a larger negative effect in this kinetic parameter (Table 4.2) highlighting the significance of the side chain size. These two mutations, E194D and E194Q, caused a 4- and 5-fold reduced  $V_{max}$  with respect to the wild type, whereas the E194A mutation decreased it 24-fold. The apparent affinities for ATP, Mg<sup>2+</sup>, and the activator FBP were not significantly affected by any of these mutations on Glu<sup>194</sup> (Table 4.3). Our results validate the hydrogen bonds observed in the structural model (Fig. 4.3) and strongly suggest that Glu<sup>194</sup> is playing a role in Glc1P binding.

Our structural model proposes that Ser<sup>212</sup> binds O3 and O4 of the sugar moiety of the ligand through hydrogen bonds with the side chain and backbone, respectively (Fig. 4.2 and 4.6). Here, we probed the role of the side chain in Glc1P binding. All Ser<sup>212</sup> mutations maintained apparent affinity properties for ATP, Mg<sup>2+</sup> and FBP at wild type levels (Table 4.3). S212A also showed values of  $k_{cat}$  similar to the wild type but it displayed a 14-fold increased  $S_{0.5}$  for Glc1P (Table 4.2). S212V and S212T caused dramatic effects in the apparent affinity for Glc1P, with 377- and 274-fold increased  $S_{0.5}$  (Table 4.2). Mutation to valine decreased the  $k_{cat}$  ~16-fold, whereas mutation to threonine did it ~2-fold, compared to the wild type. Surprisingly, replacement of the Ser<sup>212</sup> to a

tyrosine only increased the apparent affinity for Glc1P 5-fold. The  $k_{\text{cat}}$  in this mutant, however, was 231-lower than the wild type. These results strongly suggest that Ser<sup>212</sup> is located in the Glc1P binding pocket and it contributes to the enzyme's affinity for this substrate.

Asp<sup>276</sup> was replaced by alanine, asparagine and glutamic acid. The three mutations decreased the apparent affinity for Glc1P 100-, 85-, and 24-fold, respectively (Table 4.2). Our results point out the importance of Asp<sup>276</sup> for Glc1P binding and the significance of both the negative charge and the size of its side chain on such effect. The analyses of these mutants suggest an additional role for Asp<sup>276</sup> besides Glc1P interaction given that other kinetic parameters were also affected. D276A and D276N had ~1000-fold lower  $V_{\text{max}}$  than the wild type (Table 4.2), and 3.4 and ~4-fold higher  $S_{0.5}$  for ATP, respectively (Table 4.3). Instead, D276E displayed a 3-fold decreased  $V_{\text{max}}$  respect to the wild type (Table 4.2) but a bigger change on the apparent affinity for ATP, characterized by an 8-fold increased  $S_{0.5}$  for this substrate (Table 4.3). On the other hand, all three mutations decreased the apparent affinity for Mg<sup>2+</sup> ~4- to 6-fold (Table 4.3). These results would correlate with the role of Mg<sup>2+</sup> ion chelator proposed for the homologous residue, Asp<sup>280</sup>, in the *S. tuberosum* enzyme (170).

Furthermore, the three Asp<sup>276</sup> mutants had 5- to 15-fold higher  $A_{0.5}$  for FBP than that of the wild type (Table 4.3). To investigate whether this residue was involved in the activator site, we studied another mutant. Previous reports (see *Chapter 3* and (188)) showed that deletions of 11 and 15 residues from the N-terminus of the *E. coli*

**Table 4.3. Kinetic parameters of the *E. coli* wild type and mutant ADP-Glc PPases.** Reactions were performed in the synthesis direction (assay B) as described in *MATERIALS AND METHODS*. Data represent the average of two or three identical experiments  $\pm$  the average difference of the duplicates or triplicates.

Enzyme	ATP		$Mg^{2+}$		FBP	
	$S_{0.5}$ (mM)	-fold increase	$S_{0.5}$ (mM)	-fold increase	$A_{0.5}$ ( $\mu$ M)	-fold increase
Wild type	0.59 $\pm$ 0.03	1.0	2.6 $\pm$ 0.05	1.0	59.4 $\pm$ 4.7	1.0
E194A	1.20 $\pm$ 0.04	2.0	7.0 $\pm$ 0.6	2.7	321.0 $\pm$ 48.1	5.4
E194D	0.49 $\pm$ 0.01	0.8	5.3 $\pm$ 0.05	2.0	87.0 $\pm$ 17.0	1.5
E194Q	0.17 $\pm$ 0.08	0.3	4.5 $\pm$ 0.2	1.7	20.9 $\pm$ 2.8	0.4
S212A	0.68 $\pm$ 0.07	1.2	5.1 $\pm$ 0.2	2.0	85.3 $\pm$ 1.2	1.4
S212V	0.38 $\pm$ 0.07	0.7	3.1 $\pm$ 0.05	1.2	38.2 $\pm$ 4.4	0.6
S212T	0.41 $\pm$ 0.04	0.7	3.3 $\pm$ 0.2	1.3	37.3 $\pm$ 4.2	0.6
S212Y	0.43 $\pm$ 0.03	0.7	3.7 $\pm$ 0.05	1.4	121.9 $\pm$ 11.0	2.0
Y216F	0.35 $\pm$ 0.02	0.6	7.8 $\pm$ 0.6	3.0	126.0 $\pm$ 7.0	2.1
D239A	0.16 $\pm$ 0.03	0.3	3.7 $\pm$ 0.5	1.4	168.0 $\pm$ 9.0	2.8
D239E	0.96 $\pm$ 0.03	1.6	5.2 $\pm$ 0.1	2.0	118.2 $\pm$ 19.6	2.0
D239N	0.56 $\pm$ 0.05	0.9	5.6 $\pm$ 0.3	2.2	76.6 $\pm$ 2.9	1.3
F240A	1.14 $\pm$ 0.10	1.9	4.2 $\pm$ 0.2	1.6	109.0 $\pm$ 40.1	1.8
F240M	1.98 $\pm$ 0.14	3.4	5.8 $\pm$ 0.3	2.2	72.0 $\pm$ 2.1	1.2
W274A	1.04 $\pm$ 0.04	1.8	6.0 $\pm$ 0.2	2.3	304.0 $\pm$ 11.0	5.1
W274F	0.28 $\pm$ 0.01	0.5	2.4 $\pm$ 0.1	0.9	59.5 $\pm$ 9.2	1.0
W274L	0.48 $\pm$ 0.03	0.8	3.21 $\pm$ 0.01	1.2	226.7 $\pm$ 7.0	3.8
D276A	2.03 $\pm$ 0.02	3.4	11.2 $\pm$ 0.4	4.3	403.0 $\pm$ 36.4	7.0
D276N	2.3 $\pm$ 0.1	3.9	13.5 $\pm$ 2.6	5.2	760.4 $\pm$ 49.0	12.8
D276E	4.77 $\pm$ 0.04	8.1	15.2 $\pm$ 0.2	5.8	281.8 $\pm$ 46.8	5.0
K195Q <sup>b</sup>	0.19 $\pm$ 0.01	0.3	3.4 $\pm$ 0.1	1.3	21 $\pm$ 2	0.4

<sup>b</sup> Data from Ref. (174)

ADP-Glc PPase rendered activated enzymes even in the absence of FBP with all other kinetic parameters similar to those of the wild type. Based on these results, we combined both the N-terminal deletion and the single mutation D276N to create EcN $\Delta$ 15-D276N. The activity of the partially purified double mutant was  $0.027 \pm 0.003$  U/mg, similar to the partially purified D276N single mutant (data not shown), whereas the  $A_{0.5}$  for FBP was  $51 \mu\text{M}$ , similar to that of the wild type (Table 4.3). This strongly suggests that Asp<sup>276</sup> is not directly involved in activator binding but is a pivotal residue for the correct interaction of the substrates with the enzyme influencing the resulting conformational changes upon their binding.

The role of the size and aromaticity of Trp<sup>274</sup> was studied by substituting it with alanine, leucine and phenylalanine. Mutation to alanine was characterized by a 22-fold decrease in the apparent affinity for Glc1P and did not have significant effect in the  $V_{\text{max}}$  of the enzyme (Table 4.2) or in the apparent affinities for ATP, Mg<sup>2+</sup>, and FBP (Table 4.3). We obtained similar results when a leucine was placed in that position. Instead, all parameters remained almost unchanged compared to wild type levels when Trp<sup>274</sup> was replaced by a phenylalanine, indicating that aromaticity is required in that position for proper interaction of the Glc1P with the enzyme.

Tyr<sup>216</sup> is conserved not only among ADP-Glc PPases (Fig. 4.3) but also in the RmlA -Tyr<sup>176</sup>- (Fig. 4.6.A and B) (164). Mutation to phenylalanine allowed us to study the role, if any, of the side chain hydroxyl group in this position. The Y216F mutant displayed a 46-fold lower apparent affinity for Glc1P (Table 4.2) and it showed small

variations in the apparent affinities for ATP,  $Mg^{2+}$ , and FBP, from 1 to 3-fold increase in the respective kinetic constants (Table 4.3). However, this substitution, in which the OH was removed caused a 10-fold decrease in the  $V_{max}$  of the mutant enzyme. Our structural model is not showing any direct interaction between this residue and the bound ADP-Glc (Figure 4.3). Our biochemical data, however, suggest that the side chain OH plays a role in Glc1P interaction, possibly by driving the correct positioning of the substrate in the pocket, which also affects the concomitant catalytic reaction.

Asp<sup>239</sup> and Phe<sup>240</sup> are also conserved residues located in close proximity to the ligand that do not show any evident interaction with it in the three-dimensional model. However, the D239A mutation decreased 31-fold the apparent affinity for Glc1P and 11-fold the  $V_{max}$  without significant change in the other kinetic constants. Likewise, D239N and D239E increased the  $S_{0.5}$  for Glc1P 16- and 10-fold, respectively, compared to that of the wild type. The  $V_{max}$ , instead, was 2-fold lower than the wild type in the D239N mutant and was not affected by the D239E substitution (Tables 4.2 and 4.3). Replacement of Phe<sup>240</sup> to alanine and methionine affected the apparent affinity for Glc1P, displaying  $S_{0.5}$  12- and 7-fold, respectively, higher than that of the wild type. No significant changes in the  $V_{max}$  and all other kinetic parameters analyzed here were observed with F240A and F240M (Tables 4.2 and 4.3). Together, these results suggest that Asp<sup>239</sup> and Phe<sup>240</sup> are important residues for Glc1P interaction. They also point out the significance of Asp<sup>239</sup> negatively charge side chain for proper catalytic activity.

#### 4.4.5. Thermal stability

The enzymes were also studied respect to their thermal stability as explained in *Experimental Procedures*. Wild type enzyme and all Glu<sup>194</sup>, Tyr<sup>216</sup> and Asp<sup>276</sup> mutants, as well as S212A, S212V, F240M and D239E, showed ~80-85% activity after heat treatment (Table 4.4).

It is interesting to note that mutation of Phe<sup>240</sup> to alanine caused the thermal stability of the protein to decrease 50% under the assayed conditions. This result suggests that, in position 240 of the *E. coli* ADP-Glc PPase, not only hydrophobicity but also size of the side chain is important for the enzyme to adopt a proper and heat-stable folding. A similar situation was observed with Trp<sup>274</sup>. Replacement of this residue by alanine and leucine, two hydrophobic but small side chain amino acids, rendered enzymes with less than 1% residual activity after heat treatment, whereas a Phe in that position allowed the mutant enzyme to retain at least 50% of the activity (Table 4.4).

Mutants S212T and S212Y retained 60% and 70% of their initial activities. These mutations not only affected the apparent affinity for Glc1P, but also they affected the  $k_{cat}$ , suggesting a structural distortion of the active site. It is possible that this side chains are also misplacing significant structural determinants or disrupting important stabilizing interactions in the protein. Mutations in Asp<sup>239</sup> rendered enzymes with 2%, 59% and 84% residual activity after heat treatment when replaced to alanine, asparagine and glutamic acid, respectively. A negative charge is necessary in position 239 to guarantee the stability of the enzyme at temperatures higher than the optimum for activity.

**Table 4.4. Thermal stability of the wild type and single mutants.** Enzyme activity was measured after heat treatment (5 min at 60 °C) or control (0 °C) conditions as described in *MATERIALS AND METHODS*.

Enzyme	$k_{\text{cat}} (\text{s}^{-1})$						% initial activity
	Control			60 °C			
Wild type	370.0	±	14.4	315.8	±	12.2	85
E194A	15.43	±	0.07	13.0	±	1.3	84
E194D	92.7	±	11.7	75.0	±	2.3	81
E194Q	80.7	±	2.0	66.3	±	6.3	82
S212A	371.2	±	4.1	335.7	±	25.9	90
S212V	22.4	±	1.9	18.0	±	1.4	81
S212T	179.0	±	0.7	107.0	±	11.3	60
S212Y	1.6	±	0.1	1.1	±	0.1	70
Y216F	29.0	±	1.0	24.7	±	1.3	85
D239A	32.7	±	0.2	0.5	±	0.2	2
D239E	347.7	±	15.9	304.3	±	20.4	84
D239N <sup>a</sup>	152.0	±	11.7	90.3	±	2.7	59
F240A	111.6	±	2.3	56.7	±	9.0	51
F240M <sup>a</sup>	254.4	±	13.4	207.0	±	3.7	86
W274A	384.0	±	23.4	< 0.02			< 0.04
W274F	266.0	±	8.0	133.1	±	5.2	50
W274L	247.5	±	0.6	< 0.003			0.6
D276A	0.36	±	0.01	0.307	±	0.003	85
D276N	0.37	±	0.03	0.30	±	0.03	82
D276E	112.7	±	4.5	89.4	±	0.5	79

<sup>a</sup>: enzymes were ~70-80% pure

## 4.5. DISCUSSION

In this work, we report the first detailed characterization of the sugar-phosphate site and the three-dimensional structure of the *E. coli* ADP-Glc PPase, the Glc1P site. We selected a set of residues implicated in shaping this substrate pocket by examination of the primary sequences of several ADP-Glc PPases and the three-dimensional structural of the *E. coli* enzyme complexed with ADP-Glc obtained by homology modeling. The role of the selected residues in binding Glc1P was probed by means of site directed mutagenesis and steady state kinetics. The kinetic characterization of the individual mutants revealed the importance of the replaced amino acids.

Knowledge on the three-dimensional structure of the *E. coli* ADP-Glc PPase was essential to understand the complex net of interactions established between the protein and the substrate for proper binding. The first published ADP-Glc PPase crystal structure is that of the homotetrameric potato tuber small subunit solved by Jin *et al.* (170), which we used as template to build a model of the *E. coli* enzyme. The sequence identity between these two proteins is 33%, which is close to the lowest range of accepted homology for performing modeling (222). However, the functional similarity between our query and template proteins and a careful inspection of the sequence alignment, which included information on predicted secondary structures and functional conserved residues, increased the probabilities of obtaining a reliable model.

The *E. coli* ADP-Glc PPase has been subject of numerous structure-function relationship studies, including those aimed to elucidate the functional role of individual amino acids. Previously, Lys<sup>195</sup> was studied (174) and showed a very specific effect in Glc1P interaction. The reported mutations in this residue increased 100- to 10,000- fold the  $S_{0.5}$  for Glc1P without affecting other kinetic constants. To illustrate this, data reported for mutant K195Q have been included here in Tables 4.2 and 4.3. Those results are consistent with a very specific role of Lys<sup>195</sup> in the binding of Glc1P, probably by ionic interaction between the positively charged side chain  $\epsilon$ -amino and the negative phosphate group of the Glc1P (174, 177). It is possible that the rest of the amino acids in the substrate pocket, which are the subject of this work, interact with the sugar hydroxyls to increase the affinity of the binding and to provide the correct positioning of the ligand for catalysis.

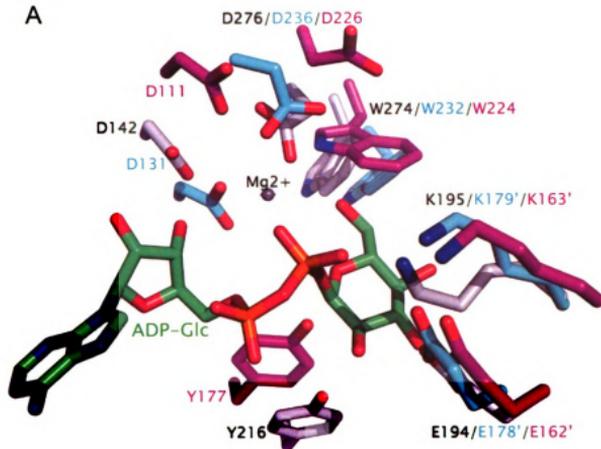
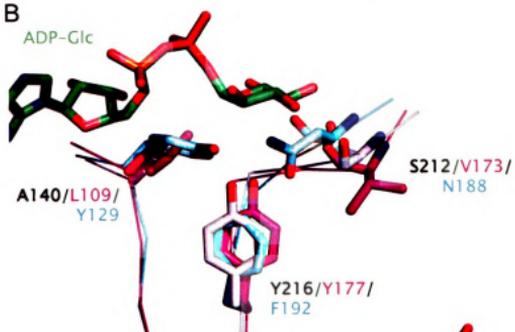
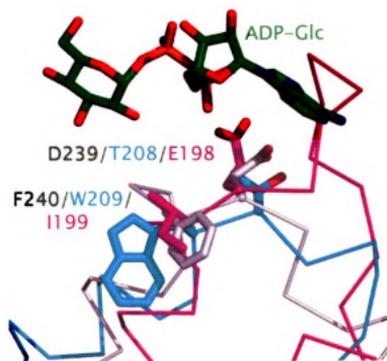
The three-dimensional model of the ADP-Glc PPase that we obtained here allowed us to visualize the spatial arrangement of a set of conserved residues potentially involved in the interaction between the enzyme and the substrate Glc1P. It has been reported that although proteins can bind carbohydrates in many different ways, certain amino acids show high propensity to be in a sugar binding site (224, 225). Some examples of this are the aromatic rings that can pack against the hydrophobic face of the sugar (226), and carboxylates that can form bidentate hydrogen bonds with two adjacent hydroxyls of the saccharide (225). In our model, we identified Trp<sup>274</sup>, Tyr<sup>216</sup> and Phe<sup>240</sup>, as well as Glu<sup>194</sup>, Asp<sup>239</sup> and Asp<sup>276</sup>, some of which were showing direct contacts with the modeled ligand ADP-Glc.

We also performed a close examination of the reported three-dimensional structures of enzymes that catalyze very similar reactions as the ADP-Glc PPase. These enzymes are the *P. aeruginosa* RmlA (164) (PDB code 1G23) and the *S. typhi* CDP-Glc PPase (169) (PDB code 1TZF). We inspected closely their active sites and we identified residues homologous to Glu<sup>194</sup>, Lys<sup>195</sup>, Asp<sup>276</sup>, Trp<sup>274</sup>, as well as the catalytic Asp<sup>142</sup> (167) of the ADP-Glc PPase in their active sites (Fig. 4.5.A). Interestingly, the CDP-Glc PPase is a trimeric enzyme with three active sites formed in the interface of adjacent monomers (169). Most of the residues contributing to the architecture of the Glc1P site belong to one of the subunits except for Glu<sup>178</sup> and Lys<sup>179</sup>, homologous to Glu<sup>194</sup> and Lys<sup>195</sup> in the *E. coli* ADP-Glc PPase, that are provided by the neighboring subunit (169).

The ADP-Glc PPase structural model shows the hexose moiety of ADP-Glc largely engaged in hydrogen bonds to surrounding residues (side chains of Lys<sup>195</sup>, Glu<sup>194</sup>, Ser<sup>212</sup> and Asp<sup>276</sup>, Fig. 4.2) and protein backbone (Ser<sup>212</sup> and Gly<sup>179</sup>, Fig. 4.6). Lys<sup>195</sup> is interacting with the  $\beta$ -phosphate of the ADP-Glc molecule. This observation is validated by the biochemical characterization reported by Hill *et al.* (174) and discussed here.

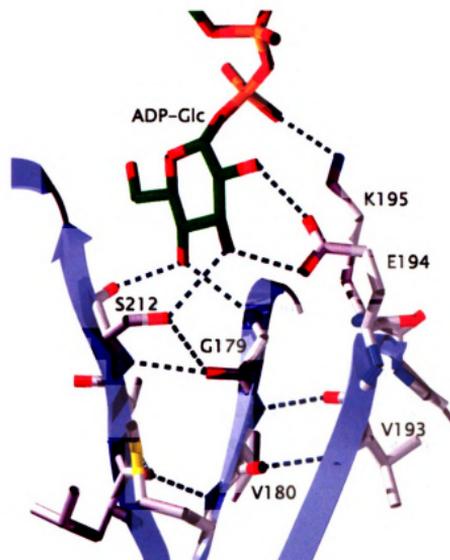
Glu<sup>194</sup> is proposed to interact with O2 and O3 of the sugar ring by a bidentate hydrogen bond. The Glu<sup>194</sup> mutants displayed the biggest changes in Glc1P apparent affinity when substituted by other residues (Table 4.2). Removal of the negative charge,

**Figure 4.5. Superposition of the amino acids in the Glc1P site from three NDP-Glc PPases. A.** Superposition of residues from the *E. coli* ADP-Glc PPase model with carbons in white (this work) and the crystal structures of the *P.aeruginosa* RmlA (PDB code 1G23) with carbons in magenta(164), and the *S. typhi* CDP-Glc PPase (PDB code 1TZF) with carbons in cyan (169). The ADP-Glc PPase model has a RMSD of 1.9 with the RmlA, and of 2.2 Å with the CDP-Glc PPase. We show the CDP-Glc PPase active site as it is in the active enzyme, with residues Asp<sup>131</sup>, Trp<sup>232</sup> and Asp<sup>236</sup> belonging to one subunit and Glu<sup>178'</sup> and Lys<sup>179'</sup> from the neighboring monomer. The ADP-Glc molecule was modeled in the ADP-Glc PPase enzyme, and the Mg<sup>2+</sup> is present in the CDP-Glc PPase crystal structure. **B.** Residues homologues to Ala<sup>140</sup>, Ser<sup>212</sup> and Tyr<sup>216</sup> in *E. coli* ADP-Glc PPase (white carbons) overlay with amino acids Tyr<sup>129</sup>, Asn<sup>188</sup> and Phe<sup>192</sup> in the CDP-Glc PPase (cyan carbons) and with Leu<sup>108</sup>, Val<sup>172</sup> and Tyr<sup>176</sup> in the RmlA (carbons in magenta), respectively. **C.** Amino acids Asp<sup>239</sup> and Phe<sup>240</sup> in the *E. coli* ADP-Glc PPase (white carbons), overlay with homologous residues Thr<sup>208</sup> and Trp<sup>209</sup> in the CDP-Glc PPase (cyan carbons), and Glu<sup>198</sup> and Ile<sup>199</sup> in the RmlA (carbons in magenta).

**A****B****C**

as observed with the glutamine mutant, caused a large decrease in this kinetic parameter (85-fold), suggesting its importance for substrate interaction. Still, the size of the side chain is also essential given that substitution by an aspartic acid decreased the apparent affinity for Glc1P more than 380-fold. Given that the distance between two atoms engaged in a hydrogen bond is crucial for the establishment of such interaction, the effect observed with a shorter side chain in position 194 supports the existence of a hydrogen bond between the ligand and Glu<sup>194</sup>. In addition, the enzyme activity seems to be affected by modifications at this position. It is possible that Glu<sup>194</sup> plays a key role in positioning the substrate in the correct orientation for catalysis, which also agrees with a critical contribution of the size in the functionality of this residue. Our results support the central role of Glu<sup>194</sup> in Glc1P binding and explain the absolute conservation of this amino acid in the ADP-Glc PPase family (Fig. 4.3) and other NDP-glucose pyrophosphorylases, such as the RmlA and the CDP-Glc PPase (Fig. 4.5.A).

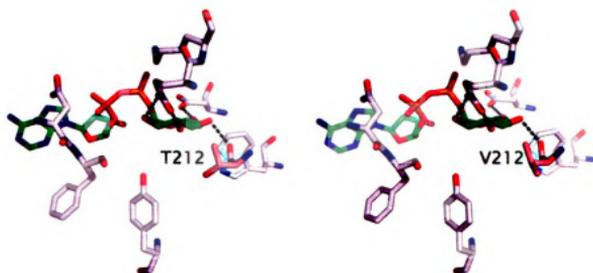
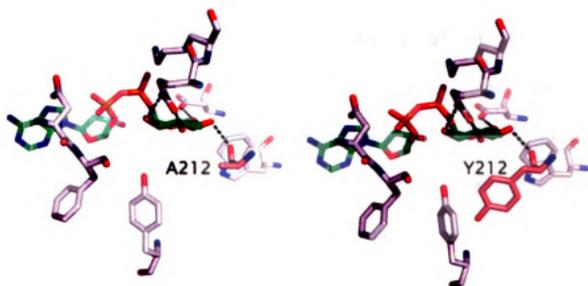
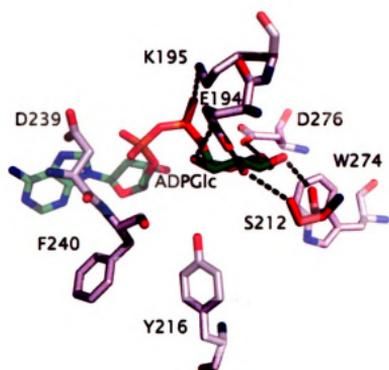
Ser<sup>212</sup> may bind the Glc1P through hydrogen bonds between the side chain and the backbone and O3 and O4 of the sugar ring, respectively (Fig. 4.2 and 4.6). We probed the role of the side chain OH group in this interaction by making conservative and non-conservative mutations. Although in various degrees, all mutants in Ser<sup>212</sup> affected the apparent affinity for Glc1P. Homology modeling of the Ser<sup>212</sup> mutants active site residues complexed with ADP-Glc show that the interaction predicted in the wild type enzyme model between Lys<sup>195</sup> ε-amino and the phosphate of the ligand, is disrupted when Ser<sup>212</sup> is replaced by other amino acid (Fig. 4.7). The 14-fold increase in Glc1P  $S_{0.5}$



**Figure 4.6. Hydrogen bond network involving Ser<sup>212</sup> in the *E. coli* ADP-Glc PPase Glc1P site.** Ser<sup>212</sup> interacts with adjacent secondary structures through a complex network of hydrogen bonds (dashed lines) involving its side chain and backbone. This network of interactions might be important for the correct positioning of other key residues for Glc1P binding, such as Glu<sup>194</sup> and Lys<sup>195</sup>.

**Figure 4.7. Stick representation of the Ser<sup>212</sup> mutants modeled active sites.**

**(A)** Modeled active site of the wild type enzyme, depicting the interactions proposed between Ser<sup>212</sup> and the ADPGlc molecule. The side chain hydrogen bond is not longer predicted in the modeled S212A **(B)**, S212Y **(C)**, S212T **(D)**, and S212V **(E)** mutants. The hydrogen bond between the Lys195  $\epsilon$ -amine is only predicted in the wild type **(A)**.



caused by the S212A mutant might be explained by the disruption of one hydrogen bond between the side chain and O3 of the glucose moiety of the ligand. Surprisingly, the effect of the side chain OH provided by threonine is counteracted by the presence of an additional methyl group, by comparison with serine. A similar situation is observed with valine in position 212. This extra methyl group is largely disrupting the proper conformation of the binding pocket. The model predicts that Ser<sup>212</sup> is spatially close to secondary structures containing Glu<sup>194</sup> and Lys<sup>195</sup> and as previously indicated, important in Glc1P interaction. Ser<sup>212</sup> is also largely engaged in a hydrogen bond network with these structures (Fig. 4.6). These observations might explain that some of the mutations in Ser<sup>212</sup> affect the apparent affinity for Glc1P as mutations in Glu<sup>194</sup> and Lys<sup>195</sup> did. Surprisingly, substitution of Ser<sup>212</sup> by a bulky side chain amino acid, tyrosine, caused a slight change in the apparent affinity for this substrate specifically. It is possible that, as the homology model predicts the preferred orientation for the tyrosine side chain in position 212 is the one directing the phenyl group away to the Glc1P pocket, burying the side chain against other hydrophobic side chains and stabilizing this position by a hydrogen bond between the phenyl OH and an adjacent backbone (not shown). It is possible that the burying of the phenyl group is causing structural arrangements, which probably extend to other parts of the active site affecting specifically an important catalytic residue. This would be explained by the dramatic reduction in the  $k_{cat}$  displayed by the mutant S212Y. Therefore, the side chain of Ser<sup>212</sup> might contribute to the overall affinity for Glc1P by making direct interactions with the O3 of the sugar ring and with adjacent backbones containing important residues for the positioning of this substrate. On the other hand, the model shows Ser<sup>212</sup> peptide carbonyl group binding O4 of the hexose

through a hydrogen bond (Fig. 4.2 and 4.6). This interaction can also be observed in the crystal structures of the other NDP-Glc PPases RmlA (164) and CDP-Glc PPase (165). The peptide carbonyl groups of Val<sup>172</sup> in the RmlA and Asn<sup>188</sup> in the CDP-Glc PPase, homologous to Ser<sup>212</sup> in ADP-Glc PPase (Fig. 4.6), also make hydrogen bonds with the substrate, implying that this interaction is important for the correct geometry of the Glc1P in the binding pocket. Apart from the specific interactions, the size of the side chain is important for the proper architecture of the Glc1P binding site.

Asp<sup>276</sup> is important for the enzyme interaction with Glc1P and it may bind the O6 of the hexose through a hydrogen bond (Fig. 4.2). Substitutions by other residues affected the apparent affinity for this substrate ~25- to 100-fold, supporting this hypothesis. However, Asp<sup>276</sup> might have a broader role rather than exclusively interacting with the Glc1P molecule since the  $V_{\max}$  and the apparent affinity for the other substrates were also affected by the studied mutations (Table 4.2 and 4.3). Asp<sup>276</sup> is spatially close to the catalytic Asp<sup>142</sup> (167), and its homologous Asp<sup>280</sup> in the potato tuber ADP-Glc PPase has been proposed as Mg<sup>2+</sup> chelator (170). These observations would explain why the different substitutions in Asp<sup>276</sup> also affected other kinetic parameters besides the Glc1P apparent affinity. In contrast to the mutations on other residues in the Glc1P site, the activation by FBP was also altered in the Asp<sup>276</sup> mutants (Table 4.2). The results obtained with mutant EcNΔ15-D276N strongly suggest that this amino acid is not directly participating in activator binding.

Asp<sup>276</sup> may be located in a hinge-like region of the active site between the ATP and the Glc1P sub-domains. Apart from interacting with the sugar ring and the Mg<sup>2+</sup>, it may also contact other residues from adjacent secondary structures establishing a network of interactions that drives the conformational changes experienced upon binding of the substrates. Comparison of the potato tuber ADP-Glc PPase crystal structures complexed with ATP or ADP-Glc illustrates such sub-domain movement (170). The observations drawn by Haugen and Preiss (201) would also contribute to explain the negative effects in all the kinetic properties of the enzyme when Asp<sup>276</sup> was mutated. They demonstrated that: (a) ATP alone displays half-site occupancy in the homotetrameric enzyme, (b) Glc1P does not bind to the enzyme unless MgCl<sub>2</sub> and ATP are present, and (c) in the presence of Glc1P, ATP displays full-site occupancy. A synergistic effect in the binding of FBP and ATP was also reported. Thus, the cooperative properties and the heterotropic interactions between substrates and effectors (201) would also explain the broad effect on the kinetic properties of the enzyme when this strategically located residue is modified in its physicochemical properties.

Aromatic residues, typically Trp or Phe, are key components of several saccharide-binding sites (226). Usually, these aromatic rings have been found involved in stacking interactions against the face of a sugar (225). However, in our structural model, none of the three aromatic residues in close proximity to the glucosyl moiety of the ligand is orienting its side chain parallel to the sugar ring. The great conservation of Trp<sup>274</sup> observed among ADP-Glc PPases (Fig. 4.3) and other pyrophosphorylases (Fig 4.6.A.) might be explained by its structural role within the Glc1P site. Substitution by

short aliphatic side chain amino acids, like alanine and leucine, affected not only the apparent affinity for Glc1P (Table 4.2), but also decreased greatly the thermal stability of the enzyme (Table 4.4). These effects were lesser when Trp<sup>274</sup> was mutated to a phenylalanine, suggesting that aromaticity is important in that position. This amino acid might provide the necessary stacking interactions to shape the Glc1P site correctly, while establishing the proper hydrophobic interactions that increase the thermal stability of the protein.

Tyr<sup>216</sup> is also located close to the ligand but no evident interaction is observed between the sugar ring and the side chain OH group. We evaluated the role of such group in Glc1P binding with the Y216F mutation, which lower the  $V_{\max}$  10-fold and the apparent affinity for this substrate 46-fold (Table 4.2). Tyr<sup>216</sup> is conserved in all ADP-Glc PPases studied so far (Fig. 4.3) and is present in the RmlA -Tyr<sup>176</sup>- (Fig. 4.6.A and B). The CDP-Glc PPase, instead, bears a phenylalanine -Phe<sup>192</sup>- in the homologous position but Tyr<sup>129</sup>, located in an adjacent  $\beta$ -strand, orients its side chain so that the OH group overlaps with that of Tyr<sup>216</sup> in the ADP-Glc PPase and of Tyr<sup>176</sup> in the RmlA (Fig. 4.6.B). Given the conservation of the aromatic ring in that position, it is possible that Tyr<sup>216</sup> is playing a structural role in the Glc1P site architecture. On the other hand, the OH group could also be making a hydrogen bond with a water molecule in direct contact with the substrate, as observed with Tyr<sup>176</sup> in the RmlA (164). This interaction might be crucial to drive the correct positioning of the Glc1P molecule for the enzymatic reaction since not only the apparent affinity for this substrate but also the catalytic activity was affected with the removal of the side chain OH group.

We also analyzed the possible roles of Asp<sup>239</sup> and Phe<sup>240</sup> as part of the Glc1P site. Our results with mutant D239E showed that the change in size significantly affected the apparent affinity for Glc1P and that a negative charge in position 239 is necessary to maintain significant enzymatic activity and thermal stability (Tables 4.2 and 4.4). On the other hand, substitutions for asparagine and alanine caused the biggest alteration in apparent affinity for Glc1P, catalytic activity (Table 4.2), and thermal stability (Table 4.4). The structures of RmlA and CDP-Glc PPase show other hydrogen bond donors in the position homologous to Asp<sup>239</sup>, a glutamic acid (Glu<sup>198</sup>) and a threonine (Thr<sup>208</sup>), respectively (Fig. 4.6.C). Moreover, in the RmlA structure Glu<sup>198</sup> interacts with O2 of the dTDP-Glc molecule through a bridging water molecule. Hydrogen bonds and ion pairs with ordered water molecules are considered important interactions that increase the thermal stability of the protein (227) and the binding affinity and specificity for the substrate (225). Another possibility is that Asp<sup>239</sup> indirectly interacts with the Glc1P through a solvent molecule, which is crucial for substrate positioning and enzymatic activity.

The data obtained with Phe<sup>240</sup> mutants demonstrate that a hydrophobic bulky residue is needed to maintain the properties of the enzyme at wild type levels. The role of Phe<sup>240</sup> might be merely structural and the effects on Glc1P apparent affinity may be a consequence of the close proximity to Asp<sup>239</sup>. In the three-dimensional model, Phe<sup>240</sup> is surrounded by a hydrophobic environment, and it is probably necessary to anchor the loop containing Asp<sup>239</sup> in the correct position. Phe<sup>240</sup> is conserved in most of the ADP-Glc PPases (Fig. 4.3), except in those from the *Mycobacterium* sp. taxonomic group,

which bear a methionine in the homologous position. Similarly, the CDP-Glc PPase shows a tryptophan (Trp<sup>209</sup>), whereas the RmlA has a smaller hydrophobic residue (Ile<sup>199</sup>) (Fig. 4.6.C). These observations together with our biochemical results support the role of Phe<sup>240</sup> as an important structural component of the Glc1*P* site.

In this work, we present data supporting that key amino acids in the ADP-Glc PPase have a role in the enzyme's affinity for Glc1*P*. Whether establishing direct hydrogen bonds with the hydroxyls in the sugar ring or solvent molecules, or properly shaping the substrate pocket, they all have an important role in determining the architecture of the Glc1*P* site. This is the first thorough biochemical characterization of the kind performed on the ADP-Glc PPases. The combination of biochemical data and the information from the three-dimensional model allowed us to hypothesize on the structural basis of the substrate binding. Comparison of our model with other NDP-Glc PPases reveals remarkable similarities, suggesting that the architecture of the Glc1*P* site is conserved. Biochemical data involving the examined amino acids have not been reported on other pyrophosphorylases to date. We believe that the results reported in this work can be extended to other members of the NDP-glucose pyrophosphorylase family providing new insights towards the understanding of the evolution of these enzymes.

## **CHAPTER 5**

### **Conclusions**

The first ADP-Glc PPase characterized was that isolated from soybean (41). Afterwards, the homologous enzymes from several prokaryotes and photosynthetic eukaryotes were reported (8, 9, 44, 228, 229). Understanding ADP-Glc PPases has been of central interest due to their key role in the pathway of  $\alpha$ -1,4-polysaccharide biosynthesis in bacteria (glycogen) and plants (starch). Characterization of the enzyme from the different sources pointed out that ADP-Glc PPase is a regulatory enzyme, with its activity being allosterically modulated by key intermediate metabolites of central routes for carbon and energy in the corresponding cell. Thus, the study of the enzyme has also been significantly valuable to understand protein structure-to-function relationships as well as enzymatic regulation by allosterism, a mechanism that is present in every metabolic system. In addition, ADP-Glc PPases proved helpful to address questions concerning enzyme evolution, particularly those integrating structure, function and regulation of the protein and the relationship with metabolic pathways operating in the different organisms.

Of all the isolated ADP-Glc PPases, the one from *E. coli* has been the best characterized and has served as a working model system to give insights into the rest of the members of the ADP-Glc PPases. Despite the advance thus reached to understand the structure and functioning of the enzyme, several queries still remained unsolved at the time that this thesis work started. For instance [1] information on the tertiary structure of the enzyme and how it assembled in the quaternary structure was not yet available. On the other hand, it has been suggested that part of the N-terminal region of ADP-Glc PPases adopt a similar fold than that of other known PPases, referred as the catalytic or

PPase domain (8, 9, 167). However, not only do these related enzymes either have smaller C-terminal regions (164, 165) or their C-termini have a different enzymatic activity (163), but ADP-Glc PPases generally present more extended N-terminal ends as well. Considering that ADP-Glc PPases are allosterically regulated but the other NDP-sugar PPases are not (8, 9, 163), another intriguing question has been [2] whether the role of the N- and C-terminal regions in ADP-Glc PPases would be determining the allosteric properties of these enzymes, which are not present in the other, unregulated and smaller, PPases. Of special interest has been also the investigation of the structural basis of [3] the intimate communication between the N- and C-terminal regions to regulate enzyme activity (168) and [4] the allosteric properties determined by a short stretch of amino acids comprising ~15-17 N-terminal residues (147, 154, 186-188).

ADP-Glc PPases (158-160), and other NDP-sugar PPases (163-165), catalyze their respective chemical reactions following a sequential ordered kinetic mechanism, with binding of the first substrate (a nucleoside-triphosphate) which allows for the binding of the second substrate (a sugar 1-phosphate). Individual amino acids involved in catalysis or interacting with substrates and allosteric modulators have been identified and their respective roles characterized in the ADP-Glc PPase (8, 9). However, [5] a more complete analysis of the structural basis responsible for ATP and Glc1P binding (especially at the three-dimensional level) and determining affinity and specificity for these substrates remained to be investigated in detail.

The work reported in this thesis has been motivated by questions presented, with

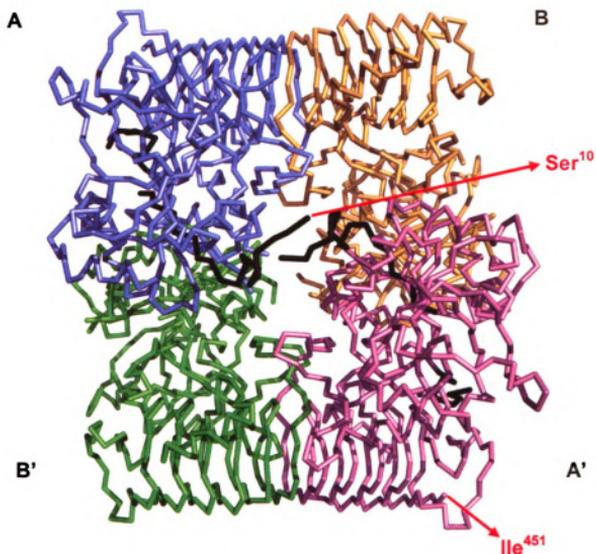
respect to the ADP-Glc PPase family, using the *E. coli* ADP-Glc PPase as a model system. Because of a certain degree of structural similarity between this family of enzymes and other PPases, some of the results presented are potentially applicable to the family of the NDP-sugar PPases.

### **5.1. Structural and functional role of N- and C-termini in allosteric regulation**

The work reported in *Chapter 2* presents solid evidence that envisage the domain organization of the monomer in the enzyme even when a direct atomic visualization, like the one provided by a crystal structure, was not yet available. The ADP-Glc PPase monomer from *E. coli* comprises two independent folding domains which are capable of reconstituting an active enzyme as the wild type even when they are expressed as independent polypeptides. The N-terminal domain, accounting for  $\frac{3}{4}$  of the total length of the whole monomer, has been predicted to include a Rossmann-like fold (167, 181) and contains all the important amino acids with proposed roles in catalysis (167, 181) or substrate binding (167, 173, 174, 181). The type of fold that the C-terminus might adopt has been more difficult to predict since related NDP-sugar PPases of known structure are either smaller polypeptides consisting of only a conserved PPase N-terminal domain (164, 165), or have a C-terminus playing a different function. An example of the latter is the case of the UDP-*N*-acetylglucosamine PPase (GlmU), where the C-terminus contains the active site for *N*-acetyl transferase activity exhibited by this bifunctional enzyme (163).

In addition, results in *Chapter 2* strongly suggest that the integrity of the C-terminus is needed to maintain a functional ADP-Glc PPase. Removal of 10 to ~100 residues from the C-terminal end yielded inactive, unfolded proteins, which were also highly susceptible to proteolysis. One possible explanation for the requirement of a full-length C-terminal domain to maintain a functional ADP-Glc PPase was suggested by comparison with the related enzyme GlmU, in which the C-terminus has a structural role in preserving the quaternary organization of the homotrimeric enzyme (163). This idea was ruled out after the potato tuber small subunit ADP-Glc PPase crystal structure was reported (170). Surprisingly, the C-terminus in the ADP-Glc PPase adopts the same fold as in GlmU, a  $\beta$ -helix, but it has a shorter length (Fig. 1.2). In addition, the oligomerization arrangement of the solved tetrameric ADP-Glc PPase structure was reported to be unique and unprecedented (Fig. 5.1) (170). If the enzyme is viewed as a dimer of dimers with each monomer labeled as A, A', B, and B', A and B make extensive stacking contacts in an end-to-end fashion between their  $\beta$ -helical C-terminus, but not along this structure as observed in the GlmU (163). Also, the big loop connecting the N- and the C-terminus from A and B makes a big interface.

When the two domains were expressed as independent polypeptides, as reported here in *Chapter 2*, the ADP-Glc PPase structure would be viewed as having a gap or nick of five amino acids in the loop linking the two domains. Even though the primary structure was interrupted, the enzyme was still fully active and the molecular mass of the native enzyme determined by size exclusion chromatography was similar to that of the wild type enzyme, indicating that the oligomeric arrangement had not been altered.



**Figure 5.1. Three-dimensional structure of the potato tuber ADP-Glc PPase.** Crystal structure of the tetrameric small subunit ADP-Glc PPase from potato tuber (residues 10-451) (*J70*). Monomers A, A', B and B' are colored in blue, magenta, yellow and green, respectively. Amino acids 10- 40 in monomers A and A' are colored in black.

Instead, the activity and stability of the enzyme was highly compromised with truncations from the C-terminal end, which possibly affected the interface between this and the N-terminal pyrophosphorylase domain. Therefore, inter-subunit interactions other than the end-to-end contacts between  $\beta$ -helices from neighboring monomers and those between their loops connecting both domains contribute to oligomerization. Moreover, the close contacts between the C- and N-terminus are important for enzyme activity, possibly due to the immediacy to the ATP site in the pyrophosphorylase N-terminal domain (170).

In ADPGlc PPases, the C-terminus has been also implicated in allosteric regulation. Several key residues with activator binding roles have been identified in this domain in the enzymes from spinach leaves (Lys<sup>440</sup> in the small subunit) (179), potato tuber (Lys<sup>404</sup> and Lys<sup>441</sup>) (157), and from *Anabaena* (Lys<sup>382</sup> and Lys<sup>419</sup>) (181, 182). Later studies with chimeric enzymes between *A. tumefaciens* and *E. coli* established that the interaction between this domain and the N-terminus is critical to determine affinity and specificity for the activators (168). Therefore, it is possible that the activator site in these enzymes is arranged three-dimensionally by residues belonging to both domains. These bacterial enzymes had been subject of several chemical modification and site directed mutagenesis studies which identified residues involved in activator binding that were located in the N-terminus (Lys<sup>39</sup> in *E. coli* and Arg<sup>32</sup>, Arg<sup>33</sup> and Arg<sup>45</sup> in *A. tumefaciens*). Therefore, it is possible that residues from the C-terminus are positioned in close proximity to those identified in the N-terminus in order to shape the architecture of the

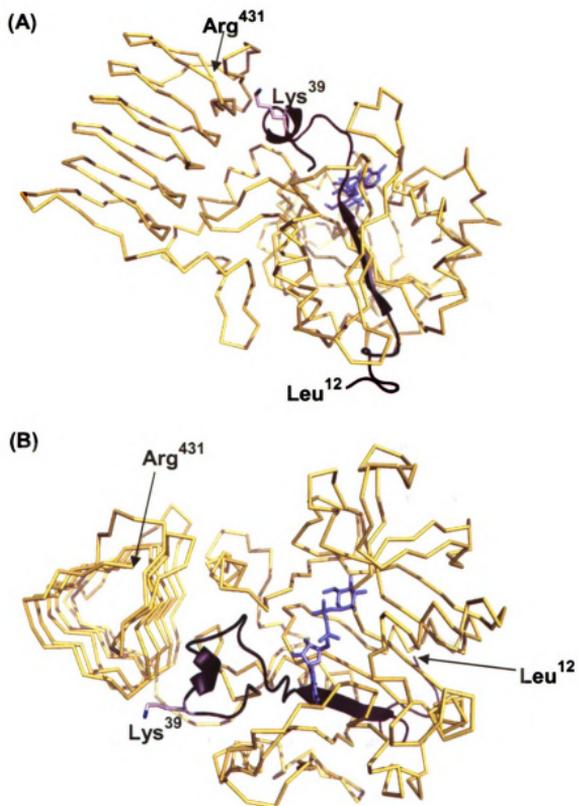
activator site. Amino acids playing a similar role located in this domain have not been proposed in any plant enzyme. But the sequences of several ADP-Glc PPases show that a Lys or an Arg is in a position homologous to the *E. coli* enzyme's Lys<sup>39</sup>, such as Lys<sup>40</sup> in potato tuber small subunit. It would be interesting to investigate whether these residues are also involved in activator binding. Thus, the combination of the N- and C-terminus to shape the activator site would be a common structural feature applying to other members of the ADP-Glc PPase family and the specificity for the different allosteric activators in each organism could be determined by specific N- and C-terminal amino acid side chains present in that site. The structural components of the activator site in ADP-Glc PPases from different organisms could have evolved together with the type of metabolic pathways of carbon assimilation operating in the cell.

The allosteric properties of ADP-Glc PPases are not only established by local structural determinants, such as the amino acids in the putative activator binding site. Studies on the enzymes from potato tuber (186) and *E. coli* (187, 188) and a detailed analysis performed on this bacterial enzyme which is reported in *Chapter 3* of this thesis, have implicated the N-terminal tail (~1-17 amino acids) in allosteric regulation. The enzyme became fully activated and insensitive to allosteric inhibition when those residues were removed. However, their direct role in inhibitor binding was ruled out, as discussed in *Chapter 3*. Instead, this short stretch of amino acids seem to influence the conformations adopted by the enzyme, which range from more activated to less activated forms, as the binding of activators and inhibitors may do.

The potato tuber ADP-Glc PPase small subunit Cys<sup>12</sup>, located in this short region, forms an inter-subunit disulfide bond whose reduction is required so to respond to the allosteric activators (154, 189). Figure 5.1. depicts the way in which the monomers are arranged in the quaternary structure allowing the disulfide bond between monomers A and A' to be formed in the center of the tetramer (170). The Cys is not conserved among the family of the ADP-Glc PPases and the mechanism of reductive activation has been described only in the potato tuber and *Arabidopsis leaves* enzymes.

Even though the N-terminal tails of the ADP-Glc PPases have been predicted as loops, the amino acid sequence in this region is not conserved. All together, the presented evidence points out that in this stretch of residues it is the structure rather than specific amino acids that determines its “allosteric switch” role in regulation. The mechanism is still unclear and direct evidence on the orientation or positioning of this region is yet unavailable since the electron density map in the potato tuber enzyme was disordered (170). However, in the three-dimensional model of the *E. coli* ADP-Glc PPase monomer which comprises residues 12-431 (Fig. 5.2), the N-terminal tail (amino acids 12-19) and the following secondary structures (residues 20-40) span along the monomer, from one side to the other. This stretch of amino acids includes the Gly-rich loop, which is in close proximity to the adenine ring of the modeled ADP-Glc molecule, and that was proposed as part of the ATP binding site (170). Lys<sup>39</sup>, located at the opposite end respect to the N-terminal tail, is structurally close to the C-terminus and the putative activator binding site.

A direct relationship between the N-terminal tail, the ATP binding pocket and the putative activator binding site in the interface between the C- and the N-terminal PPase



**Figure 5.2. Three-dimensional model of *E. coli* ADP-Glc PPase monomer.** (A) Front and (B) top view of the monomer (residues 12-431) complexed with ADP-Glc in blue. N- and C-terminus are two distinctive domains making intimate contacts among each other. Amino acids 10-40 are in black. Lys<sup>39</sup> side chain is shown (carbon atoms in white and nitrogen in blue).

domain is observed. Thus, it is possible that the effect on the flexibility of the N-terminal tail propagate through the continuous secondary structures affecting the active site similarly to the effect produced by the activator binding at the other end of this secondary structures stretch. Also, in view of the positioning of the N-terminal tails in the tetramer (Fig. 5.1) it is likely that they interact with each other, such as in the oxidized state of the potato tuber ADP-Glc PPase, to lock the enzyme in a less flexible structure. These types of interactions may prevent the conformational changes that drive the equilibrium towards a more activated state of the enzyme. Contrarily, an increase in flexibility of the tails by reduction of the disulfide bond may allow the enzyme to undergo conformational changes to a more active state upon the binding of the activators. Complete removal of 15-residues in this N-terminal tail might also promote the conformational changes shifting the equilibrium to a more active state that cannot be reversed by the allosteric inhibitors (*Chapter 3*).

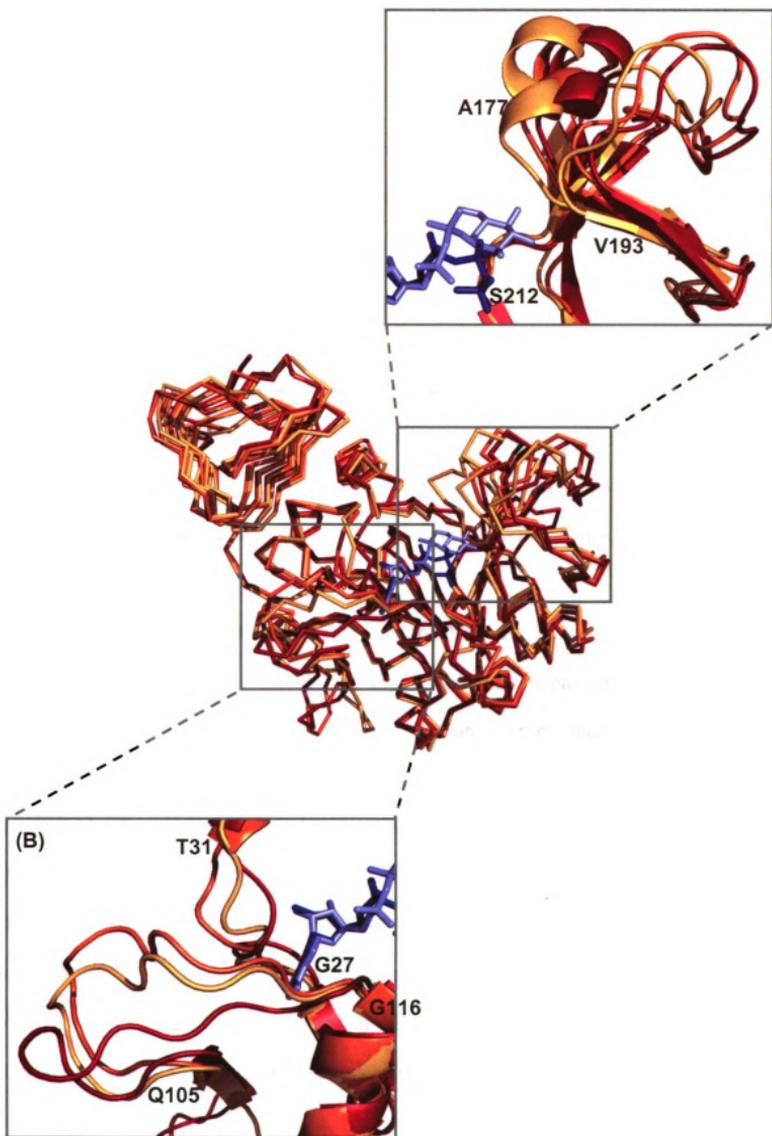
The only available crystallographic structure ADP-Glc PPase reported to date was attained in its allosterically inhibited form. It would be beneficial for the elucidation of the allosteric regulation mechanism to obtain crystalline structures of the enzyme in other states of activation. For this purpose, crystallization of the wild type enzymes from the different sources and in the presence of the various and specific allosteric activators would be the goal. Instead, the allosteric mutants already characterized, such as the N $\Delta$ 15 variant reported in *Chapter 3*, would result in a valuable alternative.

## 5.2. Structural basis of substrate binding

In order to answer questions regarding the three-dimensional organization of the *E. coli* ADP-Glc PPase, and because all previous X-ray crystallography attempts had been unsuccessful with this particular enzyme, a three-dimensional model was predicted using sequence homology-based computational methods and presented in *Chapter 4* of this thesis. The template employed was the small subunit from the potato tuber ADP-Glc PPase because, despite having only about 33% sequence identity, these enzymes have been suggested to adopt a common three-dimensional fold (8).

Three forms of the potato tuber small subunit crystal structure have been reported: the apo form of the enzyme, the one complexed with ATP, and the one complexed with ADP-Glc (PDB codes 1YP2, 1YP3 and 1YP4, respectively) (170). In that work, only two out of four monomers of the ATP-bound form, A and A', bind the substrate, observation that is in agreement with previous substrate binding studies showing that ATP is the first substrate to bind displaying only half occupancy in the absence of the second substrate, Glc1P (159). Instead, monomers A, A' and B are binding ADP-Glc in the 1YP4 structure. Also, some regions in the protein, mainly loops, show great mobility when the structures from the three forms are compared, suggesting conformational changes upon substrate binding. To illustrate this, homology models of the apo and ATP bound *E. coli* enzyme were also created here, using the methodology described in *Chapter 4* for the modeling of the ADP-Glc bound structure, and using the respective potato tuber small subunit crystal structures as templates, 1YP2 and 1YP3. Figure 5.3

**Figure 5.3** Overlap of the three *E. coli* ADP-Glc PPase models: non-ligand-, ATP- and ADP-Glc-bound forms. Apo form of the enzyme is in red, ATP-bound form in orange and ADP-Glc-bound form in yellow. The insets show a close up of the (A) region comprising residues Val<sup>170</sup>-Ser<sup>212</sup>, close to the glucose-moiety of the ADP-Glc., (B) loops in the adenine vicinity undergoing conformational changes upon ATP (blue molecule) or ADP-Glc (cyan) binding, Gly<sup>27</sup>-Thr<sup>31</sup> and Gln<sup>105</sup>-Gly<sup>116</sup>.

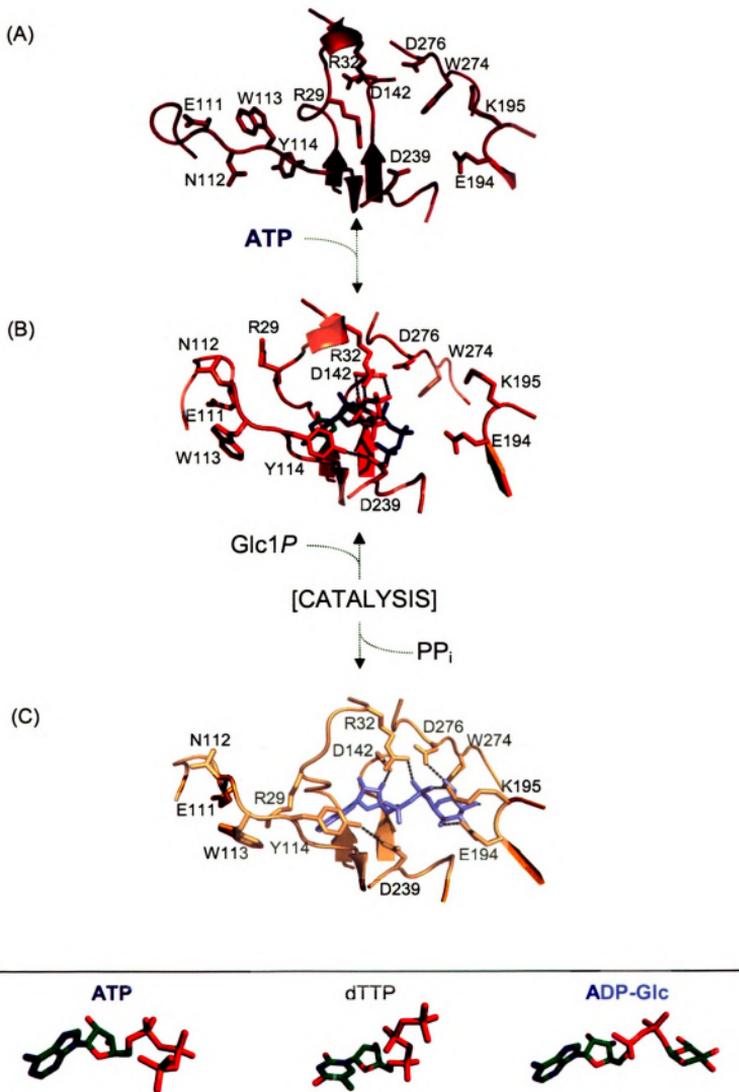


depicts an overlap of the three models and the changes undergone by these loops are evidenced.

The overall structure of the N-terminal PPase domain in ADP-Glc PPase contains a Rossmann-type fold, and presents a core of a mixed seven-stranded  $\beta$ -sheet surrounded by six  $\alpha$ -helices (see Fig. 4.1.A. and C.). The active center lies in a deep pocket and is lined by a number of polar residues that are conserved in related nucleotidyltransferases, some of which have been implicated in either catalysis or substrate binding and affinity. The active site pocket is delimited by two lobes. The first lobe encompasses strands  $\beta$ 1- $\beta$ 4 (Leu<sup>12</sup>-Tyr<sup>145</sup>, *E. coli* ADP-Glc PPase numbering) which interacts primarily with the adenine ring of ADP-Glc, whereas the second lobe includes strands  $\beta$ 5-  $\beta$ 7 (Lys<sup>146</sup>-Gly<sup>278</sup>) shaping a cavity that interacts mainly with the Glc portion of ADP-Glc.

The adenine ring, in both the ATP- and the ADP-Glc- bound forms, makes extensive interactions with two loops: the Gly-rich loop, encompassing G<sup>27</sup>GXGXRL<sup>33</sup>, and a loop formed by amino acids Gln<sup>105</sup>-Gly<sup>116</sup>, numbered according to the *E. coli* enzyme (Fig. 5.3.B). The main chain O carbonyl of Gly<sup>28</sup> and of Arg<sup>115</sup> make hydrogen bonds with N7 and N6 of the adenine ring, respectively, and the side chains of Leu<sup>25</sup> and Ile<sup>72</sup> make extensive hydrophobic interactions. Also, the pentose ring O2 and O3 bind to Lys<sup>43</sup>  $\epsilon$ -NH and to Asp<sup>142</sup> OD1, respectively. There is biochemical evidence suggesting a role for Tyr<sup>114</sup> in ATP-binding (173), and mutations in Trp<sup>113</sup> also implied the positioning of this residue close to the ATP site (unpublished). Comparison of the three models show that these residues all present different orientations whether an adenine ring is bound or

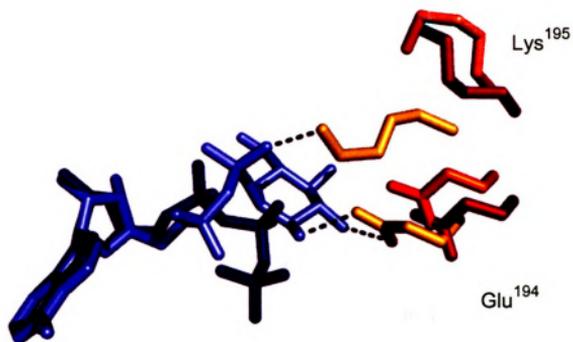
**Figure 5.4. Conformational changes in ADP-Glc PPase active site upon ligand binding. (A)** Apo form of the enzyme, before ligand binding. **(B)** Binding of ATP promotes the side chain orientation change of several amino acids in the active site: Arg<sup>29</sup>, Arg<sup>32</sup>, Glu<sup>111</sup>, Asn<sup>112</sup>, Trp<sup>113</sup>, Tyr<sup>114</sup>, Asp<sup>142</sup>, Glu<sup>194</sup>, Lys<sup>195</sup>, Asp<sup>239</sup>, and Asp<sup>276</sup>, are shown here. The  $\beta$ -phosphate of the modeled ATP -in blue- (modeled according to the potato tuber ADP-Glc PPase crystal structure) partially occupies the Glc-binding site. The molecule of dTTP from the RmlA crystal structure (was docked in the model to illustrate the possible positioning of the  $\beta$ - and  $\gamma$ - phosphates, interacting with Arg<sup>32</sup>). **(C)** Binding of ADP-Glc causes similar changes in the orientation of the side chain as observed with ATP, except for residues close to the Glc-portion of the ligand that come closer to it and interact. The most important changes are experimented by Glu<sup>194</sup> and Lys<sup>195</sup>.



not. The ATP- and in the ADP-Glc-bound forms show similar changes with respect to the apo form of the enzyme in the vicinity of the adenine ring (Fig. 5.4)

The binding of Glc1P in ADP-Glc PPase occurs only when ATP has at least occupied two active sites in the native enzyme (159). Thus, the binding of the ATP might promote conformational changes that favor the binding of the second substrate. Comparison of the three structures (Fig. 5.3.A) depicts the large motion in the carbohydrate-interaction lobe, encompassing residues Val<sup>170</sup>-Ser<sup>212</sup>, close to the glucose-moiety of the ADP-Glc. Specifically important in this region are residues Lys<sup>195</sup>, whose role in Glc1P has been reported (174), and Glu<sup>194</sup>, which has been studied in this work and presented in *Chapter 4*. These two residues are both part of a conserved motif in many NDP-sugar PPases, but only information on the role of Lys<sup>195</sup> in the *E. coli* ADP-Glc PPase (and its homologous Lys<sup>198</sup> in the potato tuber small subunit) was available to date. Mutations in these residues decreased their apparent affinity for Glc1P ~900- and ~300-fold, respectively, suggesting their critical role in binding of this substrate. Overlay of the structures shows them as involved in the biggest motion in the molecule upon substrate binding (Fig. 5.3, 5.4 and 5.5). These residues are shifted out of the binding pocket in the non-ligand-bound form of the enzyme, while they are pulled more inward in the ATP-bound form and they are pulled in significantly in the ADP-Glc-bound form, allowing direct interaction with the substrate.

The positioning of Lys<sup>195</sup> and Glu<sup>194</sup> is critical for determining the affinity for the Glc1P. The structure of the potato tuber ADP-Glc PPase small subunit complexed with



**Figure 5.5. Conformational changes experienced by Glu<sup>194</sup> and Lys<sup>195</sup> upon ligand binding.** The structures of the modeled *E. coli* ADP-Glc PPase's apo (red), ATP-bound (orange), and ADP-Glc-bound forms were overlaid to illustrate the differences in the three-dimensional positioning respect to the Glc1P pocket of Glu<sup>194</sup> and Lys<sup>195</sup>. ATP (blue) and ADP-Glc (cyan) were the ligands modeled.

ADP-Glc shows that, of the three monomers binding ADP-Glc (A, A' and B), the ligand molecule of B is the only one with the glucose moiety oriented and accommodated to a correct distance to properly interact with Lys<sup>195</sup> and Glu<sup>194</sup>. In all other subunits, these two residues are too far out of the binding pocket to make a direct interaction with the substrate, which might have caused the higher mobility of the glucose portion of the ligand and its resulting disordered electron density map. Other residues might be involved in guaranteeing the correct positioning of Lys<sup>195</sup> and Glu<sup>194</sup>. For instance, mutations of Ser<sup>212</sup>, reported in *Chapter 4*, suggested an important role in Glc1P binding and affinity. The homology model with ADP-Glc proposes, on one hand, direct interactions between the backbone and side chain of this amino acid and the glucose moiety (Fig. 4.3). But more importantly, Ser<sup>212</sup> is proposed to be involved in a complex network of hydrogen bonds with neighboring secondary structures (Fig. 4.6), which are those containing Lys<sup>195</sup> and Glu<sup>194</sup>, and experiencing the big conformational changes upon ligand binding.

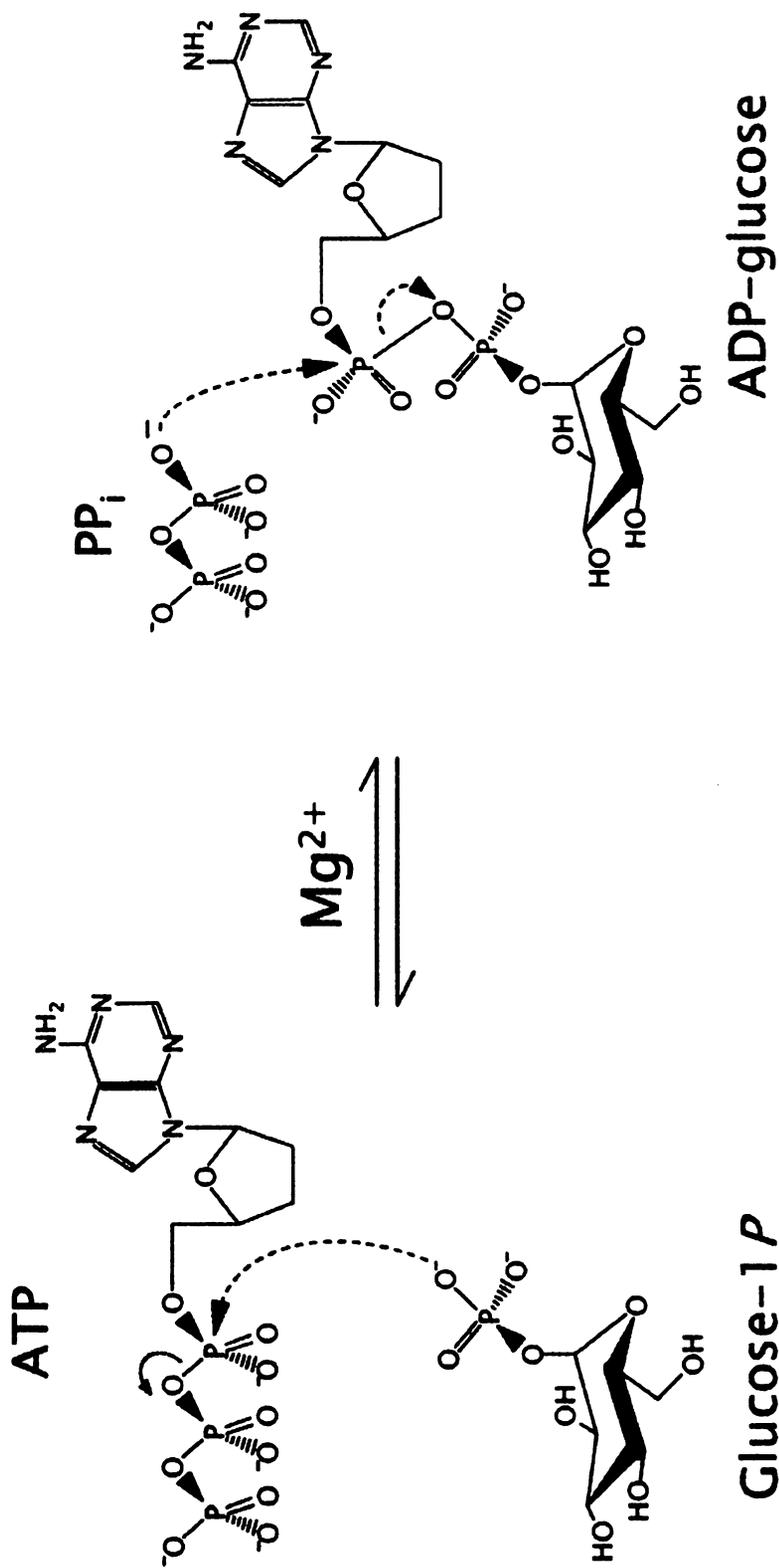
The orientation of the side chain of Asp<sup>239</sup> is also changed upon binding of ATP or ADP-Glc. Results presented in *Chapter 4* suggest that this residue is important in Glc1P binding and it is also important for enzymatic activity and thermal stability. The modeled structures suggest that the binding of ATP and ADP-Glc promotes the re-orientation of Asp<sup>239</sup> and Tyr<sup>114</sup> side chains allowing formation of a hydrogen bond. This interaction may be involved in coupling the conformational changes occurring in the ATP-site to those in the Glc1P site and determining the stability of the enzyme.

The detailed mechanism underlying the conformational changes that couple substrate binding, which are the basis for the cooperative effects observed between substrates, still remains unclear. However, the evidence presented in *Chapter 4* of this thesis made it possible to propose biochemical roles for key residues present in the Glc1P site of the *E. coli* ADP-Glc PPase and highly conserved among other PPases.

### **5.3. Structural basis of the catalytic mechanism**

Close observation of the active site and the residues likely involved in the catalytic mechanism of reaction have been already presented and discussed for other NDP-sugar PPases, such as the GlnU (163), RmlA (164), CDP-Glc PPase (165) and ADP-Glc PPase from potato tuber (170). The similarity in the kinetic mechanism and chemical reaction catalyzed, together with the structural conservation of the amino acids involved, allow to propose a general catalytic mechanism which applies also to the *E. coli* ADP-PPase.

The chemical reaction catalyzed by these enzymes has been proposed to be of the  $S_N2$  type and proceeds by a sequential ordered mechanism (Fig. 1.2 and 5.6). In the ADP-Glc synthesis direction of the reaction, ATP binds to the enzyme, followed by Glc1P, which acts as a nucleophile attacking the  $\alpha$ -phosphate of ATP. The  $\beta$ - and  $\gamma$ -phosphates of ATP are displaced and released as inorganic pyrophosphate. As required for  $S_N2$  type of reactions (230), the leaving group ( $PP_i$ ) and the attacking nucleophile (O1P of Glc1P) reside in opposite sides of the central atom in the active site of the enzyme (Fig



**Figure 5.6. Chemical reaction catalyzed by the ADP-Glc PPase. A P—O bond between a non-esterified oxygen atom of the Glc1P phosphate group and the  $\alpha$ -phosphate of ATP would be formed in this single-step mechanism.**

5.4). The amino acids in the Gly-rich loop, in particular Arg<sup>32</sup>, would interact with the  $\beta$ - and  $\gamma$ -phosphates of ATP, neutralizing the negative charges and guaranteeing the proper orientation of the PP<sub>i</sub> moiety by folding them back over the nucleotide to leave the opposite space free for Glc1P binding (Fig. 5.4.B).

The ATP molecule modeled from the ATP-bound form of the potato tuber ADP-Glc PPase small subunit (Fig. 5.4.B), shows its  $\beta$ -phosphate partially occupying the Glc-binding site and represents a conformation incompetent for catalysis (170). In that structure, a sulfate ion was found occupying the position of the  $\gamma$ -phosphate of dTTP and CTP in the dTTP- and CTP-bound forms of RmlA (164) and CDP-Glc PPase (165), respectively. Here, the molecule of dTTP was modeled and overlaid to that of ATP to illustrate the position that the  $\beta$ - and  $\gamma$ -phosphates of ATP most likely adopt when binding to the enzyme (Fig. 5.4.B).

Therefore, the flexibility of the Gly-rich loop is possibly important in allowing binding of the ATP and departure of PP<sub>i</sub> as well as in enzyme activity. In addition, the particularity of the loop becomes mostly important due to the linear relationship between this region and the allosteric regulation sites, such as the N-terminal tail at one end and the putative allosteric activator site (close to the C-terminus) at the other end. This may be one of the ways to explain how allosteric modulators communicate with the active site to affect catalytic activity, hence, contributing to the understanding of the structural basis of the allosteric regulation.

The correct orientation of the attacking nucleophile, the O1P of Glc1P is guaranteed by interaction of Lys<sup>195</sup>, which undergoes a marked conformational change observed by comparison of the apo- and ADP-Glc bound forms of the enzyme (Fig. 5.3, 5.4 and 5.5). The electrostatic repulsions between the phosphates of ATP and Glc1P are counterbalanced by a number of basic residues located in the active site, such as Arg<sup>32</sup>, Lys<sup>195</sup> and Lys<sup>42</sup> and the backbone amido groups of the Gly-rich loop. The proposed roles of Arg<sup>18</sup> and Lys<sup>25</sup> in GlnU, homologous to the *E. coli* ADP-Glc PPase Arg<sup>32</sup> and Lys<sup>42</sup>, have been reported (163), and introduction of an Arg and Lys in the homologous positions in the regulatory large subunit of the potato tuber ADP-Glc PPase converted this non-catalytic subunit into a catalytic one, demonstrating their key role in the reaction (127).

A divalent cation, preferably Mg<sup>2+</sup>, is required in this reaction, which could further neutralize the negative charges. The reported structures suggest that the metal would be coordinated by a bidentate interaction utilizing a nonbridging oxygen from the  $\alpha$ - and  $\beta$ -phosphates of ATP and the side chain oxygen atoms of two conserved residues: Asp<sup>142</sup> and Asp<sup>276</sup> in *E. coli* ADP-Glc PPase. The role of Asp<sup>142</sup> has been studied and mutations on this amino acid reduced the catalytic activity of the enzyme by four orders of magnitude (167). The important role of Asp<sup>276</sup>, suggested from observation of the reported structures, has been analyzed and presented in *Chapter 4* of this thesis. Asp<sup>276</sup> is within the set of conserved residues also interacting with the glucosyl moiety of ADP-Glc and has been thoroughly investigated here. Results indicate that not only it may be involved in binding and affinity for Glc1P but it may also play a broader role in the

active site, since mutations to other amino acids also affected the apparent affinities for ATP and  $Mg^{2+}$ , as well as reducing the  $k_{cat}$  by three orders of magnitude.

#### **5.4. Final remarks and future direction**

In summary, the work reported in this thesis provides solid evidence contributing to our better understanding of the biochemistry of the ADP-Glc PPase that is fundamental in the bacterial and plant reserve polysaccharide metabolism. A detailed analysis of the relationship between structure, function and regulation of the enzyme has been presented integrating information regarding the structural organization at the three-dimensional level and functionality of important regions in the protein, ranging from domains to individual residues.

Several questions regarding the biochemistry of ADP-Glc PPases still remain unanswered and they set the direction of future research in this field. For instance, still intriguing are the molecular determinants of ATP and  $Mg^{2+}$  binding and specificity and the conformational changes underwent upon substrate binding which are the basis of cooperative effects between substrates. Current efforts are directed to investigate them using molecular biology and protein chemistry approaches. In addition, the crystal structure of the first homotetrameric bacterial ADP-Glc PPase, the one from *Agrobacterium tumefaciens*, has been recently reported (231) however the coordinates have not yet been deposited in the Protein Data Bank. The information provided by this structure would be very valuable to understand the tetrameric organization of the homotetrameric enzyme and the structural basis of the cooperative effects between

subunits. It would also contribute greatly to the understanding of the evolution of the ADP-Glc PPases and other NDP-sugar PPases.

Another interesting goal is the application of the information obtained in the field thus far for biotechnological purposes such as the increase of starch content in crops. Manipulation of the ADP-Glc PPases in plants has proved to be an efficient way of affecting the rate of storage polysaccharide accumulation. Following this line of research, current efforts are focused on the characterization of the Glc1P site of the potato tuber enzyme, based on the results obtained with the *E. coli* enzyme, and presented here in *Chapter 4*. In addition, the study of transgenic plants bearing allosteric mutant ADP-Glc PPases, as reported before(60), would be an interesting line of research to follow. The *E. coli* N-terminal truncated allosteric mutants presented in *Chapter 3* provide an excellent alternative for this type of approach. In this case, a compromise between modified allosteric properties and heat stability would have to be achieved and interesting to be analyzed, among several other biochemical questions.

Therefore, while problems such as the mechanisms governing activity regulation and its evolution are still puzzling, the work on the *E. coli* ADP-Glc PPase presented in this thesis has paved the way for future advances in the matter and will greatly benefit not only the research on the family of the ADP-Glc PPases but of that of other NDP-sugar PPases as well.

## **BIBLIOGRAPHY**

- (1) Stryer, L. (1995) *Biochemistry*, 4th ed., Walter Freeman and Company, New York.
- (2) Voet, D., Voet, J. G., and Pratt, C. W. (2001) *Fundamentals of Biochemistry Upgrade Edition*, John Wiley & Sons, Inc., New York.
- (3) Iglesias, A. A., and Preiss, J. (1992) Bacterial glycogen and plant starch biosynthesis. *Biochemical Education* 20, 196-203.
- (4) Preiss, J. (1991) Biology and molecular biology of starch synthesis and regulation, in *Oxford Surveys of Plant Molecular and Cell Biology* (Mifflin, J., Ed.) pp 59-114, Oxford University Press, Oxford, United Kingdom.
- (5) Preiss, J. (2006) Bacterial glycogen inclusions: enzymology and regulation of synthesis, in *Microbiology Monographs* (Shively, J. M., Ed.) pp 71-108, Springer, Heidelberg, Germany.
- (6) Preiss, J., and Sivak, M. N. (1998) Starch and glycogen biosynthesis., in *Comprehensive Natural Products Chemistry* (Pinto, B. M., Ed.) pp 441-495, Pergamon Press, Oxford.
- (7) Sivak, M. N., and Preiss, J. (1998) Starch: basic science to biotechnology, in *Advances In Food and Nutrition Research* (Taylor, S. L., Ed.) pp 1-199, Academic Press, San Diego.
- (8) Ballicora, M. A., Iglesias, A. A., and Preiss, J. (2003) ADP-glucose pyrophosphorylase; a regulatory enzyme for bacterial glycogen synthesis. *Microbiol. Mol. Biol. Rev.* 67, 213-225.
- (9) Ballicora, M. A., Iglesias, A. A., and Preiss, J. (2004) ADP-glucose pyrophosphorylase, a regulatory enzyme for plant starch synthesis. *Photosynthesis Research* 79, 1-24.
- (10) Smith, A. M., Denyer, K., and Martin, C. (1997) The synthesis of the starch granule. *Annual Review of Plant Physiology and Plant Molecular Biology* 48, 449-464.
- (11) Roach, P. J. (2002) Glycogen and its metabolism. *Current Molecular Medicine* 2, 101-120.

- (12) Ball, S. G., and Morell, M. K. (2003) From bacterial glycogen to starch: Understanding the biogenesis of the plant starch granule. *Annual Review of Plant Biology* 54, 207-233.
- (13) Manners, D. J. (1991) Recent developments in our understanding of glycogen structure. *Carbohydrate Polymers* 16, 37-82.
- (14) Melendez, R., Melendez-Hevia, E., and Cascante, M. (1997) How did glycogen structure evolve to satisfy the requirement for rapid mobilization of glucose? A problem of physical constraints in structure building. *Journal of Molecular Evolution* 45, 446-455.
- (15) Strange, R. E. (1968) Bacterial "Glycogen" and survival. *Nature* 220, 606-607.
- (16) Preiss, J. (1984) Bacterial glycogen synthesis and its regulation. *Annual Review of Microbiology* 38, 419-58.
- (17) Gibbons, R. J., and Kapsimalis, B. (1963) Synthesis of intracellular iodophilic polysaccharide by *Streptococcus mitis*. *Archives of Oral Biology* 8, 319-29.
- (18) Belanger, A. E., and Hatfull, G. F. (1999) Exponential-phase glycogen recycling is essential for growth of *Mycobacterium smegmatis*. *Journal of Bacteriology* 181, 6670-6678.
- (19) Eidels, L., and Preiss, J. (1970) Carbohydrate metabolism in *Rhodopseudomonas capsulata*: enzyme titers, glucose metabolism, and polyglucose polymer synthesis. *Archives of Biochemistry and Biophysics* 140, 75-89.
- (20) Kiel, J. A., Boels, J. M., Beldman, G., and Venema, G. (1994) Glycogen in *Bacillus subtilis*: molecular characterization of an operon encoding enzymes involved in glycogen biosynthesis and degradation. *Molecular Microbiology* 11, 203-218.
- (21) Martin, M. C., Schneider, D., Bruton, C. J., Chater, K. F., and Hardisson, C. (1997) A *glgC* gene essential only for the first of two spatially distinct phases of glycogen synthesis in *Streptomyces coelicolor* A3(2). *Journal of Bacteriology* 179, 7784-7789.

- (22) Bonafonte, M. A., Solano, C., Sesma, B., Alvarez, M., Montuenga, L., Garcia-Ros, D., and Gamazo, C. (2000) The relationship between glycogen synthesis, biofilm formation and virulence in *Salmonella enteritidis*. *FEMS Microbiology Letters* 191, 31-36.
- (23) Spatafora, G., Rohrer, K., Barnard, D., and Michalek, S. (1995) A *Streptococcus mutans* mutant that synthesizes elevated levels of intracellular polysaccharide is hypercariogenic in vivo. *Infection and immunity* 63, 2556-63.
- (24) Buleon, A., Colonna, P., Planchot, V., and Ball, S. (1998) Starch granules: structure and biosynthesis. *International Journal of Biological Macromolecules* 23, 85-112.
- (25) Hizukuri, S. (1995) Starch: analytical aspects., in *Carbohydrates in food* (Eliansson, A. C., Ed.) pp 347-429, Marcel Dekker, New York.
- (26) Morrison, W. R., and Karkalas, J. (1990) Starch, in *Methods in Plant Biochemistry* (Dey, P. M., Ed.) pp 323-352, Academic Press, London.
- (27) Sachs, J. (1887) in *Lectures of the Physiology of the Plants* (Ward, H. M., Ed.) pp 304-325, Clarendon Press, Oxford.
- (28) Edwards, G. E., and Walker, D. A. (1983) C<sub>3</sub>, C<sub>4</sub>: mechanism of cellular and environmental regulation of photosynthesis, University of California Press, Berkeley.
- (29) Bandenhuizen, I. P. (1969) The Biogenesis of Starch Granules in Higher Plants, Appleton-Century Crofts, New York.
- (30) Outlaw, W. H., and Manchester, J. (1979) Guard cell starch concentration quantitatively related to stomatal aperture. *Plant Physiology* 64, 79-82.
- (31) Ritte, G., and Raschke, K. (2003) Metabolite export of isolated guard cell chloroplasts of *Vicia faba*. *New Phytologist* 159, 195-202.
- (32) Zeiger, E., Talbott, L. D., Frechilla, S., Srivastava, A., and Zhu, J. (2002) The guard cell chloroplast: a perspective for the twenty-first century. *New Phytologist* 153, 415-424.

- (33) Cori, G. T., and Cori, C. F. (1939) The activating effect of glycogen on the enzymatic synthesis of glycogen from glucose 1-phosphate. *J. Biol. Chem.* 131, 397-398.
- (34) Devlin, T. M. (1997) in *Textbook of Biochemistry with clinical correlations*, John Wiley & Son, New York.
- (35) Caputto, R., Leloir, L. F., Cardini, C. E., and Paladini, A. C. (1950) Isolation of the coenzyme of the galactose phosphate-glucose phosphate transformation. *Journal of Biological Chemistry* 184, 333-350.
- (36) Cardini, C. E., Paladini, A. C., Caputto, R., and Leloir, L. F. (1950) Uridine diphosphate glucose: the coenzyme of the galactose-glucose phosphate isomerization. *Nature* 165, 191.
- (37) Leloir, L. F. (1971) Two decades in the biosynthesis of polysaccharides. *Science* 172, 1299-303.
- (38) Leloir, L. F., Rongine de Fekete, M. A., and Cardini, C. E. (1961) Starch and oligosaccharide synthesis from uridine diphosphate glucose. *J. Biol. Chem.* 236, 636-641.
- (39) Recondo, E., and Leloir, L. F. (1961) Adenosine diphosphate glucose and starch synthesis. *Biochemical and Biophysical Research Communications* 6, 86-88.
- (40) Recondo, E., Dankert, M., and Leloir, L. F. (1963) Isolation of adenosine diphosphate D-glucose from corn grains. *Biochemical and Biophysical Research Communications* 12, 204-7.
- (41) Espada, J. (1962) Enzymic synthesis of adenosine diphosphate glucose from glucose 1-phosphate and adenosine triphosphate. *Journal of Biological Chemistry* 237, 3577-3581.
- (42) Leloir, L. F., and Goldemberg, S. H. (1960) Synthesis of glycogen from uridine diphosphate glucose in liver. *Journal of Biological Chemistry* 235, 919-923.
- (43) Sigal, N., Cattaneo, J., and Segel, I. H. (1964) Glycogen accumulation by wild-type and uridine diphosphate glucose pyrophosphorylase-negative strains of *Escherichia coli*. *Archives of Biochemistry and Biophysics* 108, 440-451.

- (44) Preiss, J., Shen, L., Greenberg, E., and Gentner, N. (1966) Biosynthesis of bacterial glycogen. IV. Activation and inhibition of the adenosine diphosphate glucose pyrophosphorylase of *Escherichia coli* B. *Biochemistry* 5, 1833-1845.
- (45) Preiss, J. (1982) Regulation of the biosynthesis and degradation of starch. *Annual Review of Plant Physiology* 54, 431-454.
- (46) Preiss, J. (1988) Biochemistry of the starch and its regulation, in *The Biochemistry of Plants* (Preiss, J., Ed.) pp 181-254, Academic Press, New York.
- (47) Preiss, J., and Romeo, T. (1994) Molecular biology and regulatory aspects of glycogen biosynthesis in bacteria, in *Progress in nucleic acid research and molecular biology* (Cohn, W. E., and Moldave, K., Eds.) pp 299-329, Academic Press, New York.
- (48) Ghosh, P., Meyer, C., Remy, E., Peterson, D., and Preiss, J. (1992) Cloning, expression, and nucleotide sequence of *glgC* gene from an allosteric mutant of *Escherichia coli* B. *Archives of Biochemistry and Biophysics* 296, 122-128.
- (49) Meyer, C. R., Bork, J. A., Nadler, S., Yirsa, J., and Preiss, J. (1998) Site-directed mutagenesis of a regulatory site of *Escherichia coli* ADP- glucose pyrophosphorylase: the role of residue 336 in allosteric behavior. *Arch. Biochem. Biophys.* 353, 152-159.
- (50) Meyer, C. R., Ghosh, P., Remy, E., and Preiss, J. (1992) Cloning, expression, and nucleotide sequence of a mutant *glgC* gene from *Escherichia coli* B. *Journal of Bacteriology* 174, 4509-4512.
- (51) Meyer, C., Ghosh, P., Nadler, S., and Preiss, J. (1993) Cloning, expression, and sequence of an allosteric mutant ADPglucose pyrophosphorylase from *Escherichia coli* B. *Arch. Biochem. Biophys.* 302, 64-71.
- (52) Preiss, J., Sabraw, A., and Greenberg, E. (1971) An ADP-glucose pyrophosphorylase with lower apparent affinities for substrate and effector molecules in an *Escherichia coli* B mutant deficient in glycogen synthesis. *Biochem. Biophys. Res. Commun.* 42, 180-186.
- (53) Ball, S., Marianne, T., Dirick, L., Fresnoy, M., Delrue, B., and Decq, A. (1991) A *Chlamydomonas reinhardtii* low-starch mutant is defective for 3-

phosphoglycerate activation and orthophosphate inhibition of ADP-glucose pyrophosphorylase. *Planta* 185, 17-26.

- (54) Neuhaus, H. E., and Stitt, M. (1990) Control analysis of photosynthate partitioning. Impact of reduced activity of ADP-glucose pyrophosphorylase or plastid phosphoglucomutase on the fluxes to starch and sucrose in *Arabidopsis thaliana* (L.) Heynh. . *Planta* 182, 445-454.
- (55) Lin, T. P., Caspar, T., Somerville, C., and Preiss, J. (1988) Isolation and characterization of a starchless mutant of *Arabidopsis thaliana* L. Henyh lacking ADPglucose pyrophosphorylase activity. *Plant Physiology* 86, 1131-1135.
- (56) Lin, T. P., Caspar, T., Somerville, C., and Preiss, J. (1988) A starch deficient mutant of *Arabidopsis thaliana* with low ADPglucose pyrophosphorylase activity lacks one of the two subunits of the enzyme. *Plant Physiology* 88.
- (57) Li, L., and Preiss, J. (1992) Characterization of ADP-glucose pyrophosphorylase from a starch-deficient mutant of *Arabidopsis thaliana* (L). *Carbohydrate Research* 227, 227-239.
- (58) Giroux, M. J., Shaw, J. R., Barry, G. F., Cobb, B. G., Greene, T., Okita, T. W., and Hannah, L. C. (1996) A single mutation that increases maize seed weight. *Proceedings of the National Academy of Sciences of the United States of America* 93, 5824-5829.
- (59) Smidansky, E. D., Clancy, M., Meyer, F. D., Lanning, S. P., Blake, N. K., Talbert, L. E., and Giroux, M. J. (2002) Enhanced ADP-glucose pyrophosphorylase activity in wheat endosperm increases seed yield. *Proceedings of the National Academy of Sciences of the United States of America* 99, 1724-1729.
- (60) Stark, D. M., Timmerman, K. P., Barry, G. F., Preiss, J., and Kishore, G. M. (1992) Role of ADPglucose pyrophosphorylase in regulating starch levels in plant tissues. *Science* 258, 287-292.
- (61) Geigenberger, P., Stitt, M., and Fernie, A. R. (2004) Metabolic control analysis and regulation of the conversion of sucrose to starch in growing potato tubers *Plant Cell and Environment* 27, 655-673.
- (62) Smidansky, E. D., Meyer, F. D., Blakeslee, B., Weglarz, T. E., Greene, T. W., and Giroux, M. J. (2006) Expression of a modified ADP-glucose pyrophosphorylase

large subunit in wheat seeds stimulates photosynthesis and carbon metabolism. *Planta (in press)*.

- (63) Smidansky, E. D., Martin, J. M., Hannah, L. C., Fischer, A. M., and Giroux, M. J. (2003) Seed yield and plant biomass increases in rice are conferred by deregulation of endosperm ADP-glucose pyrophosphorylase. *Planta* 216, 656-664.
- (64) Kumar, A., C.E., L., and Preiss, J. (1986) Biosynthesis of bacterial glycogen. Primary structure of *Escherichia coli* ADPglucose: $\alpha$ -1,4-glucan, 4-glucosyltransferase deduced from the nucleotide sequence of the *glgA* gen. *Journal of Biological Chemistry* 261, 14634-14639.
- (65) Furukawa, K., Tagaya, M., Inouye, M., Preiss, J., and Fukui, T. (1990) Identification of lysine 15 at the active site in *Escherichia coli* glycogen synthase. Conservation of Lys-X-Gly-Gly sequence in the bacterial and mammalian enzymes. *J. Biol. Chem.* 265, 2086-2090.
- (66) Furukawa, K., Tagaya, M., Tanizawa, K., and Fukui, T. (1994) Identification of Lys277 at the active site of *Escherichia coli* glycogen synthase. Application of affinity labeling combined with site-directed mutagenesis. *J. Biol. Chem.* 269, 868-871.
- (67) Holmes, E., Boyer, C., and Preiss, J. (1982) Immunological characterization of *Escherichia coli* B glycogen synthase and branching enzyme and comparison with enzymes from other bacteria. *J. Bacteriol.* 151, 1444-1453.
- (68) Holmes, E., and Preiss, J. (1979) Characterization of *Escherichia coli* B glycogen synthase enzymatic reactions and products. *Archives of Biochemistry and Biophysics* 196, 436-48.
- (69) Liu, J., and Mushegian, A. (2003) Three monophyletic superfamilies account for the majority of the known glycosyltransferases. *Protein Sci* 12, 1418-1431.
- (70) MacGregor, E. A. (2002) Possible structure and active site residues of starch, glycogen, and sucrose synthases. *Journal of Protein Chemistry* 21, 297-306.
- (71) Yep, A., Ballicora, M. A., and Preiss, J. (2004) The active site of the *Escherichia coli* glycogen synthase is similar to the active site of retaining GT-B

glycosyltransferases. *Biochemical and Biophysical Research Communications* 316, 960-966.

- (72) Yep, A., Ballicora, M. A., Sivak, M. N., and Preiss, J. (2004) Identification and characterization of a critical region in the glycogen synthase from *Escherichia coli*. *J. Biol. Chem.* 279, 8359-8367.
- (73) Buschiazzo, A., Ugalde, J. E., Guerin, M. E., Shepard, W., Ugalde, R. A., and Alzari, P. M. (2004) Crystal structure of glycogen synthase: homologous enzymes catalyze glycogen synthesis and degradation. *EMBO Journal* 23, 3196-205.
- (74) Horcajada, C., Guinovart, J. J., Fita, I., and Ferrer, J. C. (2006) Crystal Structure of an Archaeal Glycogen Synthase: insights into oligomerization and substrate binding of eukaryotic glycogen synthases. *J. Biol. Chem.* 281, 2923-2931.
- (75) Baba, T., Nishihara, M., Mizuno, K., Kawasaki, T., Shimada, H., Kobayashi, E., Ohnishi, S., Tanaka, K., and Arai, Y. (1993) Identification, cDNA cloning and gene expression of soluble starch synthase in rice (*Oryza sativa* L.) immature seeds. *Plant Physiology* 103, 565-573.
- (76) Downton, W. J. S., and Hawker, J. (1973) Enzymes of starch and sucrose metabolism in *Zea Mays* leaves. *Phytochemistry* 12, 1551-6.
- (77) Macdonald, F. D., and Preiss, J. (1985) Partial purification and characterization of granule-bound starch synthases from normal and *waxy* maize. *Plant Physiology* 78, 849-852.
- (78) Ozbun, J. L., Hawker, J. S., and Preiss, J. (1971) Multiple forms of  $\alpha$ -1,4 glucan synthetase from spinach leaves. *Biochemical and Biophysical Research Communications* 43, 631-636.
- (79) Ozbun, J. L., Hawker, J. S., Greenberg, E., Lammel, C., Preiss, J., and Lee, E. Y. C. (1973) Starch synthetase, phosphorylase, ADPglucose pyrophosphorylase and UDPglucose pyrophosphorylase in developing maize kernels. *Plant Physiology* 51, 1-5.
- (80) Edwards, A., Fulton, D. C., Hylton, C. M., Jobling, S. A., Gidley, M., Rossner, U., Martin, C., and Smith, A. M. (1999) A combined reduction in activity of starch synthases II and III of potato has novel effects on the starch of tubers. *Plant Journal* 17, 251-261.

- (81) Jobling, S. A., Westcott, R. J., Tayal, A., Jeffcoat, R., and Schwall, G. P. (2002) Production of a freeze-thaw-stable potato starch by antisense inhibition of three starch synthase genes. *Nature Biotechnology* 20, 295-299.
- (82) Lloyd, J. R., Landschutze, V., and Kossmann, J. (1999) Simultaneous antisense inhibition of two starch-synthase isoforms in potato tubers leads to accumulation of grossly modified amylopectin. *Biochemical Journal* 338, 515-521.
- (83) Maddelein, M. L., Libessart, N., Bellanger, F., Delrue, B., Dhulst, C., Vandenkoornhuysse, N., Fontaine, T., Wieruszeski, J. M., Decq, A., and Ball, S. (1994) Toward an understanding of the biogenesis of the starch granule - determination of granule-bound and soluble starch synthase functions in amylopectin synthesis. *Journal of Biological Chemistry* 269, 25150-25157.
- (84) Visser, R. G., Somhorst, I., Kuipers, G. J., Ruys, N. J., Feenstra, W., and Jacobsen, E. (1991) Inhibition of the gene for granule-bound starch synthase in potato by antisense constructs. *Molecular & General Genetics* 225, 289-296.
- (85) Wattedled, F., Buleon, A., Bouchet, B., Ral, J.-P., Lienard, L., Delvalle, D., Binderup, K., Dauvillee, D., Ball, S., and D'Hulst, C. (2002) Granule-bound starch synthase: A major enzyme involved in the biogenesis of B-crystallites in starch granules. *European Journal of Biochemistry* 269, 3810-3820.
- (86) Delrue, B., Fontaine, T., Routier, F., Decq, A., Wieruszeski, J. M., Van Den Koornhuysse, N., Maddelein, M. L., Fournet, B., and Ball, S. (1992) Waxy *Chlamydomonas reinhardtii*: monocellular algal mutants defective in amylose biosynthesis and granule-bound starch synthase activity accumulate a structurally modified amylopectin. *Journal Of Bacteriology* 174, 3612-3620.
- (87) Gao, Z., Keeling, P. L., Shibles, R., and Guan, H. (2004) Involvement of lysine-193 of the conserved "K-T-G-G" motif in the catalysis of maize starch synthase IIa. *Archives of Biochemistry and Biophysics* 427, 1-7.
- (88) Imparl-Radosevich, J. M., Keeling, P. L., and Guan, H. (1999) Essential arginine residues in maize starch synthase IIa are involved in both ADP-glucose and primer binding. *FEBS Letters* 457, 357-362.
- (89) Nichols, D. J., Keeling, P. L., Spalding, M., and Guan, H. (2000) Involvement of conserved aspartate and glutamate residues in the catalysis and substrate binding of maize starch synthase. *Biochemistry* 39, 7820-7825.

- (90) Nakamura, Y., Francisco, P. B. J., Hosaka, Y., Sato, A., Sawada, T., Kubo, A., and Fujita, N. (2005) Essential amino acids of starch synthase IIa differentiate amylopectin structure and starch quality between japonica and indica rice varieties. *Plant Molecular Biology* 58, 213-27.
- (91) Horibata, T., Nakamoto, M., Fuwa, H., and Inouchi, N. (2004) Structural and physicochemical characteristics of endosperm starches of rice cultivars recently bred in Japan. *Journal of Applied Glycoscience* 51, 303-313.
- (92) Borovsky, D., Smith, E. E., and Whelan, W. J. (1976) On the mechanism of amylose branching by potato Q-enzyme. *European Journal of Biochemistry* 62, 307-312.
- (93) Baecker, P. A., Greenberg, E., and Preiss, J. (1986) Biosynthesis of bacterial glycogen: Primary structure of *Escherichia coli*  $\alpha$ -1,4-glucan:  $\alpha$ -1,4-glucan 6-glycosyltransferase as deduced from the nucleotide sequence of the *glgB* gene. *Journal of Biological Chemistry* 261.
- (94) Kiel, J. A. K. W., Boels, J. M., Beldman, G., and Venema, G. (1990) Nucleotide sequence of the *Synechococcus* sp. PCC7942 branching enzyme gene (*glg B*): Expression in *B. subtilis*. *Gene* 89, 77-84.
- (95) Kiel, J. A. K. W., Boels, J. M., Beldman, G., and Venema, G. (1991) Molecular cloning and nucleotide sequence of the branching enzyme gene (*glg B*) from *Bacillus stearothermophilus* and expression in *Escherichia coli* and *Bacillus subtilis*. *Mol. Gen. Genet.* 230, 136-144.
- (96) Kiel, J. A. K. W., Boels, J. M., Beldman, G., and Venema, G. (1992) The *glg B* gene from the thermophile *Bacillus caldolyticus* encodes a thermolabile branching enzyme. *DNA sequence : the journal of DNA sequencing and mapping* 3, 221-232.
- (97) Rumbak, E., Rawlings, D. E., Lindsay, G. G., and Woods, D. R. (1991) Characterization of the *Butyrivibrio fibrisolvens glgB* gene which encodes a glycogen-branching enzyme with starch clearing activity. *Journal of Bacteriology* 173, 6732-6741.
- (98) Takata, H., Takaha, T., Kuriki, T., Okada, S., Takagi, M., and Imanaka, T. (1994) Properties and active center of the thermostable branching enzyme from *Bacillus stearothermophilus*. *Applied and Environmental Microbiology* 60, 3096-3104.

- (99) Takata, H., Takaha, T., Okada, S., Hizukuri, S., Takagi, M., and Imanaka, T. (1996) Structure of the cyclic glucan produced from amylopectin by *Bacillus stearothermophilus* branching enzyme. *Carbohydrate Research* 295, 91-101
- (100) Schwall, G. P., Safford, R., Westcott, R. J., Jeffcoat, R., Tayal, A., Shi, Y. C., Gidley, M. J., and Jobling, S. A. (2000) Production of very-high-amylose potato starch by inhibition of SBE A and B. *Nature Biotechnology* 18, 551-554.
- (101) Garwood, D. L., Shannon, J. C., and Creech, R. G. (1976) Starches of endosperms possessing different alleles at amylose-extender locus in *Zea Mays* L. *Cereal Chemistry* 53, 355-364.
- (102) Blauth, S. L., Kim, K. N., Klucinec, J., Shannon, J. C., Thompson, D., and Gultinan, M. (2002) Identification of *Mutator* insertional mutants of starch-branching enzyme 1 (*sbe1*) in *Zea mays* L. *Plant Molecular Biology* 48, 287-297.
- (103) Blauth, S. L., Yao, Y., Klucinec, J. D., Shannon, J. C., Thompson, D. B., and Gultinan, M. J. (2001) Identification of *Mutator* insertional mutants of starch-branching enzyme 2a in corn. *Plant Physiol.* 125, 1396-1405.
- (104) Svensson, B. (1994) Protein engineering in the  $\alpha$ -amylase family: catalytic mechanism, substrate specificity, and stability. *Plant Molecular Biology* 25, 141-157.
- (105) Kuriki, T., Guan, H. P., Sivak, M. N., and Preiss, J. (1996) Analysis of the active center of branching enzyme II from maize endosperm. *Journal of Protein Chemistry* 15, 305-13.
- (106) Baba, T., Kimura, K., Mizuno, K., Etoh, H., Ishida, Y., Shida, O., and Arai, Y. (1991) Sequence conservation of the catalytic regions of amylolytic enzymes in maize branching enzyme-I. *Biochemical and Biophysical Research Communications* 181, 87-94.
- (107) Abad, M. C., Binderup, K., Rios-Steiner, J., Arni, R. K., Preiss, J., and Geiger, J. H. (2002) The X-ray crystallographic structure of *Escherichia coli* branching enzyme. *J. Biol. Chem.* 277, 42164-42170.
- (108) Kuriki, T., Stewart, D. C., and Preiss, J. (1997) Construction of chimeric enzymes out of maize endospermbranching enzymes I and II: Activity and properties. *Journal of Biological Chemistry* 272, 28999-29004.

- (109) Binderup, K., Mikkelsen, R., and Preiss, J. (2000) Limited proteolysis of branching enzyme from *Escherichia coli*. *Archives of Biochemistry and Biophysics* 377, 366-371.
- (110) Binderup, K., Mikkelsen, R., and Preiss, J. (2002) Truncation of the amino terminus of branching enzyme changes its chain transfer pattern. *Archives of Biochemistry and Biophysics* 397, 279-285.
- (111) Devillers, C. H., Piper, M. E., Ballicora, M. A., and Preiss, J. (2003) Characterization of the branching patterns of glycogen branching enzyme truncated on the N-terminus. *Archives of Biochemistry and Biophysics* 418, 34-38.
- (112) Kawabata, Y., Toeda, K., Takahashi, T., Shibamoto, N., and Kobayashi, M. (2002) Preparation of highly branched starch by glycogen branching enzyme from *Neurospora crassa* N2-44 and its characterization. *Journal of Applied Glycoscience* 49, 273-279.
- (113) Takata, H., Ohdan, K., Takaha, T., Kuriki, T., and Okada, S. (2003) Properties of branching enzyme from hyperthermophilic bacterium, *Aquifex aeolicus*, and its potential for production of highly-branched cyclic dextrin. *Journal of Applied Glycoscience* 50, 15-20.
- (114) Takata, H., Takaha, T., Nakamura, H., Fujii, K., Okada, S., Takagi, M., and Imanaka, T. (1997) Production and some properties of a dextrin with a narrow size distribution by the cyclization reaction of branching enzyme. *Journal of Fermentation and Bioengineering* 84, 119-123.
- (115) Takata, H., Takaha, T., Okada, S., Takagi, M., and Imanaka, T. (1996) Cyclization reaction catalyzed by branching enzyme. *Journal of Bacteriology* 178.
- (116) Preiss, J. (1999) Biosynthesis of bacterial and mammalian glycogen and plant starch: synthesis and regulation, in *Biorganic chemistry: Carbohydrates* (Hecht, S. M., Ed.) pp 489-554, Oxford University Press, Oxford.
- (117) Krisman, C. R., and Barengo, R. (1975) A precursor of glycogen biosynthesis:  $\alpha$ -1,4-glucan-protein. *European Journal of Biochemistry* 52, 117-123.

- (118) Rodriguez, I. R., and Whelan, W. J. (1985) A novel glycosyl-amino acid linkage: rabbit-muscle glycogen is covalently linked to a protein *via* tyrosine. *Biochemical and Biophysical Research Communications* 132, 829-836.
- (119) Iglesias, A. A., Charng, Y. Y., Ball, S., and Preiss, J. (1994) Characterization of the kinetic, regulatory, and structural properties of ADP-glucose pyrophosphorylase from *Chlamydomonas reinhardtii*. *Plant Physiology* 104, 1287-1294.
- (120) Takata, H., Takaha, T., Okada, S., Takagi, M., and Imanaka, T. (1997) Characterization of a gene cluster for glycogen biosynthesis and a heterotetrameric ADP-glucose pyrophosphorylase from *Bacillus stearothermophilus*. *J. Bacteriol.* 179, 4689-4698.
- (121) Igarashi, R. Y., and Meyer, C. (2000) Cloning and sequencing of glycogen metabolism genes from *Rhodobacter spaeroides* 2.4.1. Expression and characterization of recombinant ADP-glucose pyrophosphorylase *Archives of Biochemistry and Biophysics* 376, 47-58.
- (122) Uttaro, A. D., Ugalde, R. A., Preiss, J., and Iglesias, A. A. (1998) Cloning and expression of the *glgC* gene from *Agrobacterium tumefaciens*: purification and characterization of the ADP-glucose synthetase. *Archives of Biochemistry and Biophysics* 357, 13-21.
- (123) Morell, M. K., Bloom, M., Knowles, V., and Preiss, J. (1987) Subunit structure of spinach leaf ADPglucose pyrophosphorylase. *Plant Physiology* 85, 182-187.
- (124) Preiss, J., and Sivak, M. N. (1998) Biochemistry, molecular biology and regulation of starch synthesis. *Genet. Eng.* 20, 177-223.
- (125) Nakata, P. A., Greene, T. W., Anderson, J. M., Smith-White, B. J., Okita, T. W., and Preiss, J. (1991) Comparison of the primary sequences of two potato tuber ADPglucose pyrophosphorylase subunits. *Plant Molecular Biology* 17, 1089-1093.
- (126) Frueauf, J. B., Ballicora, M. A., and Preiss, J. (2003) ADP-glucose pyrophosphorylase from potato tuber: site-directed mutagenesis of homologous aspartic acid residues in the small and large subunits. *Plant J.* 33, 503-511.

- (127) Ballicora, M. A., Dubay, J. R., Devillers, C. H., and Preiss, J. (2005) Resurrecting the ancestral enzymatic role of a modulatory subunit. *Journal of Biological Chemistry* 280, 10189-95.
- (128) Smith-White, B. J., and Preiss, J. (1992) Comparison of proteins of ADP-glucose pyrophosphorylase from diverse sources. *J. Mol. Evol.* 34, 449-464.
- (129) Muller-Rober, B., Nast, G., and Willmitzer, L. (1995) Isolation and expression analysis of cDNA clones encoding a small and a large subunit of ADP-glucose pyrophosphorylase from sugar beet. *Plant Molecular Biology* 27, 191-7.
- (130) Chen, B.-Y., and Janes, H. W. (1998) Multiple forms of ADP glucose pyrophosphorylase from tomato leaf. *Physiologia Plantarum* 103, 491-496.
- (131) Chen, B.-Y., Janes, H. W., and Gianfagna, T. (1998) PCR cloning and characterization of multiple ADP-glucose pyrophosphorylase cDNAs from tomato *Plant Science* 136, 59-67.
- (132) Burton, R. A., Johnson, P. E., Beckles, D. M., Fincher, G. B., Jenner, H. L., Naldrett, M. J., and Denyer, K. (2002) Characterization of the genes encoding the cytosolic and plastidial forms of ADP-glucose pyrophosphorylase in wheat endosperm. *Plant Physiol.* 130, 1464-1475.
- (133) Weber, H., Heim, U., Borisjuk, L., and Wobus, U. (1995) Cell-type specific, coordinate expression of two ADP-glucose pyrophosphorylase genes in relation to starch biosynthesis during seed development of *Vicia faba* L. *Planta* 195, 352-61.
- (134) La Cognata, U., Willmitzer, L., and Muller-Rober, B. (1995) Molecular cloning and characterization of novel isoforms of potato ADP-glucose pyrophosphorylase. *Molecular & General Genetics* 246, 538-48.
- (135) Kim, I. J., Noh, S. J., Lee, B. H., Jo, J., Kim, Y. S., and Chung, W. I. (2001) Molecular characterization of cDNA clones for ADP-glucose pyrophosphorylase from Citrus. *Biochimica et biophysica acta* 1518, 324-8.
- (136) Crevillen, P., Ventriglia, T., Pinto, F., Orea, A., Merida, A., and Romero, J. M. (2005) Differential pattern of expression and sugar regulation of *Arabidopsis thaliana* ADP-glucose pyrophosphorylase-encoding genes. *J. Biol. Chem.* 280, 8143-8149.

- (137) Crevillen, P., Ballicora, M. A., Merida, A., Preiss, J., and Romero, J. M. (2003) The different large subunit isoforms of *Arabidopsis thaliana* ADP-glucose pyrophosphorylase confer distinct kinetic and regulatory properties to the heterotetrameric enzyme. *Journal of Biological Chemistry* 278, 28508-28515.
- (138) Preiss, J. (1973) Adenosine diphosphoryl glucose pyrophosphorylase., in *The Enzymes* (Boyer, P. D., Ed.) pp 73-119, Academic Press, New York.
- (139) Preiss, J., and Romeo, T. (1989) Physiology, biochemistry and genetics of bacterial glycogen synthesis. *Adv. Microb. Physiol.* 30, 184-238.
- (140) Greenberg, E., Preiss, J. E., Van Boldrick, M., and Preiss, J. (1983) Biosynthesis of bacterial glycogen: Activator specificity of the ADPglucose pyrophosphorylase of *Rhodospseudomonas*. *Archives of Biochemistry and Biophysics* 220, 594-604.
- (141) Furlong, C. E., and Preiss, J. (1969) Biosynthesis of bacterial glycogen synthesis. VII. Purification and properties of adenosine diphosphoglucose pyrophosphorylase of *Rhodospirillum rubrum*. *Journal of Biological Chemistry* 244, 2539-2548.
- (142) Yung, S. G., and Preiss, J. (1981) Biosynthesis of bacterial glycogen: purification and structural properties of *Rhodospirillum tenue* adenosine diphosphate glucose synthetase. *Journal of Bacteriology* 147, 101-109.
- (143) Matsuno, K., Blais, T., Serio, A. W., Conway, T., Henkin, T. M., and Sonenshein, A. L. (1999) Metabolic imbalance and sporulation in an isocitrate dehydrogenase mutant of *Bacillus subtilis*. *Journal of Bacteriology* 181, 3382-3391.
- (144) Iglesias, A. A., and Podesta, F. E. (2005) Photosynthate formation and partitioning in crop plants, in *Handbook of photosynthesis, 2<sup>nd</sup> Ed.* (Pessarakli, M., Ed.) pp 525-545, CRC Press, Taylor & Francis Group, Boca Raton, FL.
- (145) Charng, Y. Y., Kakefuda, G., Iglesias, A. A., Buikema, W. J., and Preiss, J. (1992) Molecular cloning and expression of the gene encoding ADPglucose pyrophosphorylase from the cyanobacterium *Anabaena* sp. strain PCC 7120. *Plant Molecular Biology* 20, 37-47.
- (146) Iglesias, A. A., Kakefuda, G., and Preiss, J. (1991) Regulatory and structural properties of the cyanobacterial ADPglucose pyrophosphorylases. *Plant Physiol.* 97, 1187-1195.

- (147) Ballicora, M. A., Laughlin, M. J., Fu, Y., Okita, T. W., Barry, G. F., and Preiss, J. (1995) Adenosine 5'-diphosphate-glucose pyrophosphorylase from potato tuber. Significance of the N terminus of the small subunit for catalytic properties and heat stability. *Plant Physiol.* 109, 245-251.
- (148) Iglesias, A. A., Barry, G. F., Meyer, C., Bloksberg, L., Nakata, P. A., Greene, T., Laughlin, M. J., Okita, T. W., Kishore, G. M., and Preiss, J. (1993) Expression of the potato tuber ADP-glucose pyrophosphorylase in *Escherichia coli*. *J. Biol. Chem.* 268, 1081-1086.
- (149) Gomez-Casati, D. F., and Iglesias, A. A. (2002) ADP-glucose pyrophosphorylase from wheat endosperm. Purification and characterization of an enzyme with novel regulatory properties. *Planta* 214, 428-434.
- (150) Hylton, C. M., and Smith, A. M. (1992) The *rb* mutation of peas causes structural and regulatory changes in ADP-glucose pyrophosphorylase from developing embryos. *Plant Physiology* 99, 1626-1634.
- (151) Kleczkowski, L. A., Villand, P., Luthi, E., Olsen, O. A., and Preiss, J. (1993) Insensitivity of barley endosperm ADP-glucose pyrophosphorylase to 3-phosphoglycerate and orthophosphate regulation. *Plant Physiol.* 101, 179-186.
- (152) Plaxton, W. C., and Preiss, J. (1987) Purification and properties of nonproteolytic degraded ADPglucose pyrophosphorylase from maize endosperm. *Plant Physiol.* 83, 105-112.
- (153) Rudi, H., Doan, D. N. P., and Olsen, O.-A. (1997) A (His)<sub>6</sub>-tagged recombinant barley (*Hordeum vulgare* L.) endosperm ADP-glucose pyrophosphorylase expressed in the baculovirus-insect cell system is insensitive to allosteric regulation by 3-phosphoglycerate and inorganic phosphate. *FEBS Letters* 419, 124-130.
- (154) Fu, Y., Ballicora, M. A., Leykam, J. F., and Preiss, J. (1998) Mechanism of reductive activation of potato tuber ADP-glucose pyrophosphorylase. *J. Biol. Chem.* 273, 25045-25052.
- (155) Tiessen, A., Hendriks, J. H., Stitt, M., Branscheid, A., Gibon, Y., Farre, E. M., and Geigenberger, P. (2002) Starch synthesis in potato tubers is regulated by post-translational redox modification of ADP-glucose pyrophosphorylase: a novel regulatory mechanism linking starch synthesis to the sucrose supply. *Plant Cell* 14, 2191-2213.

- (156) Hendriks, J. H. M., Kolbe, A., Gibon, Y., Stitt, M., and Geigenberger, P. (2003) ADP-glucose pyrophosphorylase is activated by posttranslational redox-modification in response to light and to sugars in leaves of *Arabidopsis* and other plant species. *Plant Physiology* 133, 838-849.
- (157) Ballicora, M. A., Fu, Y., Nesbitt, N. M., and Preiss, J. (1998) ADP-Glucose pyrophosphorylase from potato tubers. Site-directed mutagenesis studies of the regulatory sites. *Plant Physiol.* 118, 265-274.
- (158) Paule, M. R., and Preiss, J. (1971) The kinetic mechanism of adenosine diphosphoglucose pyrophosphorylase from *Rhodospirillum rubrum*. *Journal of Biological Chemistry* 246, 4602-4609.
- (159) Haugen, T. H., and Preiss, J. (1979) Biosynthesis of bacterial glycogen. The nature of the binding of substrates and effectors to ADP-glucose synthase. *J. Biol. Chem.* 254, 127-136.
- (160) Kleczkowski, L. A., Villand, P., Preiss, J., and Olsen, O. A. (1993) Kinetic mechanism and regulation of ADP-glucose pyrophosphorylase from barley (*Hordeum vulgare*) leaves. *J. Biol. Chem.* 268, 6228-6233.
- (161) Lamerz, A.-C., Haselhorst, T., Bergfeld, A. K., von Itzstein, M., and Gerardy-Schahn, R. (2006) Molecular cloning of the *Leishmania major* UDP-glucose pyrophosphorylase, functional characterization, and ligand binding analyses using NMR spectroscopy. *J. Biol. Chem.* 281, 16314-16322.
- (162) Sheu, K. F., and Frey, P. A. (1978) UDP-glucose pyrophosphorylase. Stereochemical course of the reaction of glucose 1-phosphate with uridine-5'[1-thiotriphosphate]. *J. Biol. Chem.* 253, 3378-3380.
- (163) Brown, K., Pompeo, F., Dixon, S., Mengin-Lecreulx, D., Cambillau, C., and Bourne, Y. (1999) Crystal structure of the bifunctional *N*-acetylglucosamine 1-phosphate uridylyltransferase from *Escherichia coli*: a paradigm for the related pyrophosphorylase superfamily. *EMBO Journal* 18, 4096-4107.
- (164) Blankenfeldt, W., Asuncion, M., Lam, J. S., and Naismith, J. H. (2000) The structural basis of the catalytic mechanism and regulation of glucose-1-phosphate thymidylyltransferase (RmlA). *EMBO Journal* 19, 6652-6663.

- (165) Zuccotti, S., Zanardi, D., Rosano, C., Sturla, L., Tonetti, M., and Bolognesi, M. (2001) Kinetic and crystallographic analyses support a sequential-ordered Bi Bi catalytic mechanism for *Escherichia coli* glucose-1-phosphate thymidyltransferase. *Journal of Molecular Biology* 313, 831-843.
- (166) Koropatkin, N. M., Cleland, W. W., and Holden, H. M. (2005) Kinetic and structural analysis of  $\alpha$ -D-glucose-1-phosphate cytidyltransferase from *Salmonella typhi*. *J. Biol. Chem.* 280, 10774-10780.
- (167) Frueauf, J. B., Ballicora, M. A., and Preiss, J. (2001) Aspartate residue 142 is important for catalysis by ADP-glucose pyrophosphorylase from *Escherichia coli*. *J. Biol. Chem.* 276, 46319-46325.
- (168) Ballicora, M. A., Sesma, J. I., Iglesias, A. A., and Preiss, J. (2002) Characterization of chimeric ADPglucose pyrophosphorylases of *Escherichia coli* and *Agrobacterium tumefaciens*. Importance of the C- terminus on the selectivity for allosteric regulators. *Biochemistry* 41, 9431-9437.
- (169) Koropatkin, N. M., and Holden, H. M. (2004) Molecular structure of  $\alpha$ -D-Glucose-1-phosphate cytidyltransferase from *Salmonella typhi*. *J. Biol. Chem.* 279, 44023-44029.
- (170) Jin, X., Ballicora, M. A., Preiss, J., and Geiger, J. H. (2005) Crystal structure of potato tuber ADP-glucose pyrophosphorylase. *EMBO Journal*, 694-704.
- (171) Lee, Y. M., Mukherjee, S., and Preiss, J. (1986) Covalent modification of *Escherichia coli* ADPglucose synthetase with 8- azido substrate analogs. *Archives of Biochemistry and Biophysics* 244, 585-595.
- (172) Lee, Y. M., and Preiss, J. (1986) Covalent modification of substrate-binding sites of *Escherichia coli* ADP-glucose synthetase. Isolation and structural characterization of 8- azido-ADP-glucose-incorporated peptides. *Journal of Biological Chemistry* 261, 1058-1064.
- (173) Kumar, A., Tanaka, T., Lee, Y. M., and Preiss, J. (1988) Biosynthesis of bacterial glycogen. Use of site-directed mutagenesis to probe the role of tyrosine 114 in the catalytic mechanism of ADP-glucose synthetase from *Escherichia coli*. *J. Biol. Chem.* 263, 14634-14639.

- (174) Hill, M. A., Kaufmann, K., Otero, J., and Preiss, J. (1991) Biosynthesis of bacterial glycogen. Mutagenesis of a catalytic site residue of ADP-glucose pyrophosphorylase from *Escherichia coli*. *J. Biol. Chem.* 266, 12455-12460.
- (175) Parsons, T. F., and Preiss, J. (1978) Biosynthesis of bacterial glycogen. Incorporation of pyridoxal phosphate into the allosteric activator site and an ADP-glucose-protected pyridoxal phosphate binding site of *Escherichia coli* B ADP-glucose synthase. *J. Biol. Chem.* 253, 6197-6202.
- (176) Parsons, T. F., and Preiss, J. (1978) Biosynthesis of bacterial glycogen. Isolation and characterization of the pyridoxal-P allosteric activator site and the ADP-glucose-protected pyridoxal-P binding site of *Escherichia coli* B ADP-glucose synthase. *J. Biol. Chem.* 253, 7638-7645.
- (177) Fu, Y., Ballicora, M. A., and Preiss, J. (1998) Mutagenesis of the glucose-1-phosphate-binding site of potato tuber ADP-glucose pyrophosphorylase. *Plant Physiol.* 117, 989-996.
- (178) Ball, K., and Preiss, J. (1994) Allosteric sites of the large subunit of the spinach leaf ADPglucose pyrophosphorylase. *Journal of Biological Chemistry* 269, 24706-24711.
- (179) Morell, M., Bloom, M., and Preiss, J. (1988) Affinity labeling of the allosteric activator site(s) of spinach leaf ADP-glucose pyrophosphorylase. *J. Biol. Chem.* 263, 633-637.
- (180) Sheng, J., and Preiss, J. (1997) Arginine(294) is essential for the inhibition of *Anabaena* PCC 7120 ADP-glucose pyrophosphorylase by phosphate. *Biochemistry* 36, 13077-13084.
- (181) Charng, Y. Y., Iglesias, A. A., and Preiss, J. (1994) Structure-function relationships of cyanobacterial ADP-glucose pyrophosphorylase. Site-directed mutagenesis and chemical modification of the activator-binding sites of ADP-glucose pyrophosphorylase from *Anabaena* PCC 7120. *J. Biol. Chem.* 269, 24107-24113.
- (182) Sheng, J., Charng, Y. Y., and Preiss, J. (1996) Site-directed mutagenesis of lysine382, the activator-binding site, of ADPglucose pyrophosphorylase from *Anabaena* PCC 7120. *Biochemistry* 35, 3115-3121.

- (183) Frueauf, J. B., Ballicora, M. A., and Preiss, J. (2002) Alteration of inhibitor selectivity by site-directed mutagenesis of Arg(294) in the ADP-glucose pyrophosphorylase from *Anabaena* PCC 7120. *Arch. Biochem. Biophys.* 400, 208-214.
- (184) Gardiol, A., and Preiss, J. (1990) *Escherichia coli* E-39 ADPglucose synthetase has different activation kinetics from the wild-type allosteric enzyme. *Archives of Biochemistry and Biophysics* 280, 175-180.
- (185) Gomez-Casati, D. F., Igarashi, R. Y., Berger, C. N., Brandt, M. E., Iglesias, A. A., and Meyer, C. R. (2001) Identification of functionally important amino-terminal arginines of *Agrobacterium tumefaciens* ADP-glucose pyrophosphorylase by alanine scanning mutagenesis. *Biochemistry* 40, 10169-10178.
- (186) Laughlin, M. J., Chantler, S. E., and Okita, T. W. (1998) N- and C-terminal peptide sequences are essential for enzyme assembly, allosteric, and/or catalytic properties of ADP-glucose pyrophosphorylase. *Plant J.* 14, 159-168.
- (187) Wu, M. X., and Preiss, J. (1998) The N-terminal region is important for the allosteric activation and inhibition of the *Escherichia coli* ADP-glucose pyrophosphorylase. *Arch. Biochem. Biophys.* 358, 182-188.
- (188) Wu, M. X., and Preiss, J. (2001) Truncated forms of the recombinant *Escherichia coli* ADP-glucose pyrophosphorylase: The importance of the N-terminal region for allosteric activation and inhibition. *Arch. Biochem. Biophys.* 389, 159-165.
- (189) Ballicora, M. A., Frueauf, J. B., Fu, Y., Schurmann, P., and Preiss, J. (2000) Activation of the potato tuber ADP-glucose pyrophosphorylase by thioredoxin. *J. Biol. Chem.* 275, 1315-1320.
- (190) Guerin, M., and Parodi, A. J. (2003) The UDP-glucose:glycoprotein glucosyltransferase is organized in at least two tightly bound domains from yeast to mammals. *J. Biol. Chem.* 278, 20540-20546.
- (191) Hammet, A., Pike, B. L., Mitchelhill, K. I., Teh, T., Kobe, B., House, C. M., Kemp, B. E., and Heierhorst, J. (2000) FHA domain boundaries of the Dun1p and Rad53p cell cycle checkpoint kinases. *FEBS Letters* 471, 141-146.

- (192) Hallet, B., Sherratt, D. J., and Hayes, F. (1997) Pentapeptide scanning mutagenesis: random insertion of a variable five amino acid cassette in a target protein. *Nucleic Acids Research (Online)* 25, 1866-7.
- (193) Yep, A., Bejar, C. M., Ballicora, M. A., Dubay, J. R., Iglesias, A. A., and Preiss, J. (2004) An assay for adenosine 5'-diphosphate (ADP)-glucose pyrophosphorylase that measures the synthesis of radioactive ADP-glucose with glycogen synthase. *Analytical Biochemistry* 324, 52-59.
- (194) Govons, S., Vinopal, R., Ingraham, J., and Preiss, J. (1969) Isolation of mutants of *Escherichia coli* B altered in their ability to synthesize glycogen. *Journal of Bacteriology* 97, 970-972.
- (195) Lindqvist, L., Kaiser, R., Reeves, P. R., and Lindberg, A. A. (1994) Purification, characterization, and high performance liquid chromatography assay of *Salmonella* glucose-1-phosphate cytidyltransferase from the cloned *rfbF* gene. *J. Biol. Chem.* 269, 122-126.
- (196) Ning, B., and Elbein, A. D. (2000) Cloning, expression and characterization of the pig liver GDP-mannose pyrophosphorylase. Evidence that GDP-mannose and GDP-Glc pyrophosphorylases are different proteins. *European Journal of Biochemistry* 267, 6866-6874.
- (197) Peneff, C., Ferrari, P., Charrier, V., Taburet, Y., Monnier, C., Zamboni, V., Winter, J., Harnois, M., Fassy, F., and Bourne, Y. (2001) Crystal structures of two human pyrophosphorylase isoforms in complexes with UDPGlc(Gal)NAc: role of the alternatively spliced insert in the enzyme oligomeric assembly and active site architecture. *EMBO J.* 20, 6191-6202.
- (198) Weissborn, A. C., Liu, Q., Rumley, M. K., and Kennedy, E. P. (1994) UTP: alpha-D-glucose-1-phosphate uridylyltransferase of *Escherichia coli*: isolation and DNA sequence of the *galU* gene and purification of the enzyme. *Journal of Bacteriology* 176, 2611-2618.
- (199) Westerholm-Parvinena, b. A., Vernosa, I., and Serranob, L. (2000) Kinesin subfamily UNC104 contains a FHA domain: boundaries and physicochemical characterization. *FEBS Letters* 486, 285-290.
- (200) Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goetze, N. M., Olson, B. J., and Klenk, D. C. (1985) Measurement of protein using bicinchoninic acid [published erratum

appears in *Anal Biochem* 1987 May 15;163(1):279]. *Analytical Biochemistry* 150, 76-85.

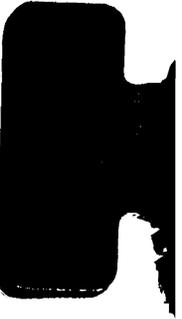
- (201) Haugen, T. H., Ishaque, A., and Preiss, J. (1976) Biosynthesis of bacterial glycogen. Characterization of the subunit structure of *Escherichia coli* B glucose-1-phosphate adenylyltransferase (EC 2.7.7.27). *Journal of Biological Chemistry* 251, 7880-7885.
- (202) Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-685.
- (203) Fiser, A., Do, R. K., and Sali, A. (2000) Modeling of loops in protein structures. *Protein Science*, 1753-1773.
- (204) Marti-Renom, M. A., Stuart, A., Fiser, A., Sánchez, R., Melo, F., and Sali, A. (2000) Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.*, 291-325.
- (205) Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.*, 779-815.
- (206) Luthy, R., Bowie, J., and Eisenberg, D. (1992) Assessment of protein models with three-dimensional profiles. *Nature* 356, 83-5.
- (207) Bowie, J., Luthy, R., and Eisenberg, D. (1991) A method to identify protein sequences that fold into a known three-dimensional structure. *Science* 253, 164-70.
- (208) Baecker, P. A., Furlong, C. E., and Preiss, J. (1983) Biosynthesis of bacterial glycogen. Primary structure of *Escherichia coli* ADP-glucose synthetase as deduced from the nucleotide sequence of the *glgC* gene. *J. Biol. Chem.* 258, 5084-5088.
- (209) Uttaro, A. D., and Ugalde, R. A. (1994) A chromosomal cluster of genes encoding ADP-glucose synthetase, glycogen synthase and phosphoglucomutase in *Agrobacterium tumefaciens*. *Gene* 150, 117-122.
- (210) Palenik, B., Brahamsha, B., Larimer, F. W., Land, M., Hauser, L., Chain, P., Lamerdin, J., Regala, W., Allen, E. E., McCarren, J., Paulsen, I., Dufresne, A.,

Partensky, F., Webb, E. A., and Waterbury, J. (2003) The genome of a motile marine *Synechococcus*. *Nature* 424 1037-1042.

- (211) Nelson, K. E., Clayton, R. A., Gill, S. R., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J. D., Nelson, W. C., Ketchum, K. A., McDonald, L., Utterback, T. R., Malek, J. A., Linher, K. D., Garrett, M. M., Stewart, A. M., Cotton, M. D., Pratt, M. S., Phillips, C. A., Richardson, D., Heidelberg, J., Sutton, G. G., Fleischmann, R. D., Eisen, J. A., White, O., Salzberg, S. L., Smith, H. O., Venter, J. C., and Fraser, C. M. (1999) Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. *Nature* 399, 323-329.
- (212) Tettelin, H., Nelson, K. E., Paulsen, I. T., Eisen, J. A., Read, T. D., Peterson, S., Heidelberg, J., DeBoy, R. T., Haft, D. H., Dodson, R. J., Durkin, A. S., Gwinn, M., Kolonay, J. F., Nelson, W. C., Peterson, J. D., Umayam, L. A., White, O., Salzberg, S. L., Lewis, M. R., Radune, D., Holtzapple, E., Khouri, H., Wolf, A. M., Utterback, T. R., Hansen, C. L., McDonald, L. A., Feldblyum, T. V., Angiuoli, S., Dickinson, T., Hickey, E. K., Holt, I. E., Loftus, B. J., Yang, F., Smith, H. O., Venter, J. C., Dougherty, B. A., Morrison, D. A., Hollingshead, S. K., and Fraser, C. M. (2001) Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science* 293, 498-506.
- (213) Heidelberg, J. F., Eisen, J. A., Nelson, W. C., Clayton, R. A., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J. D., Umayam, L., Gill, S. R., Nelson, K. E., Read, T. D., Tettelin, H., Richardson, D., Ermolaeva, M. D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I., Sellers, P., McDonald, L., Utterback, T., Fleischmann, R. D., Nierman, W. C., White, O., Salzberg, S. L., Smith, H. O., Colwell, R. R., Mekalanos, J. J., Venter, J. C., and Fraser, C. M. (2000) DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* 406 477-483.
- (214) Fleischmann, R. D., Alland, D., Eisen, J. A., Carpenter, L., White, O., Peterson, J. D., DeBoy, R. T., Dodson, R. J., Gwinn, M. L., Haft, D. H., Hickey, E. K., Kolonay, J. F., Nelson, W. C., Umayam, L. A., Ermolaeva, M. D., Salzberg, S. L., Delcher, A., Utterback, T. R., Weidman, J. F., Khouri, H. M., Gill, J., Mikula, A., Bishai, W., Jacobs, W. R., Jr., Venter, J. C., and Fraser, C. M. (2002) Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. *J. Bacteriol.* 184, 5479-5490.
- (215) White, O., Eisen, J. A., Heidelberg, J., Hickey, E. K., Peterson, D., Dodson, R. J., Haft, D. H., Gwinn, M., Nelson, W. C., Richardson, D., Moffat, K. S., Qin, H., Jiang, L., Pamphile, W., Crosby, M., Sheng, M., Vamathevan, J., Lam, P., McDonald, L., Utterback, T., Zalewski, C., Makarova, K. S., Aravind, L., Daly,

- M. J., Minton, K. W., Fleischmann, R. D., Ketchum, K. A., Nelson, K. E., Salzberg, S. L., Smith, H. O., Venter, J. C., and Fraser, C. M. (1999) Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. *Science* 286, 1571-1577.
- (216) Wang, S. M., Lue, W. L., Yu, T. S., Long, J. H., Wang, C. N., Eimert, K., and Chen, J. (1998) Characterization of ADG1, an *Arabidopsis* locus encoding for ADPG pyrophosphorylase small subunit, demonstrates that the presence of the small subunit is required for large subunit stability. *Plant Journal* 13, 63-70.
- (217) Hannah, L. C., Shaw, J. R., Giroux, M. J., Reyss, A., Prioul, J. L., Bae, J. M., and Lee, J. Y. (2001) Maize genes encoding the small subunit of ADP-glucose pyrophosphorylase. *Plant Physiology* 127, 173-183.
- (218) Zabawinski, C., Van Den Koornhuysse, N., D'Hulst, C., Schlichting, R., Giersch, C., Delrue, B., Lacroix, J.-M., Preiss, J., and Ball, S. (2001) Starchless mutants of *Chlamydomonas reinhardtii* lack the small subunit of a heterotetrameric ADP-Glucose pyrophosphorylase. *J. Bacteriol.* 183, 1069-1077.
- (219) Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* 15, 51-59.
- (220) Sanchez, R., and Sali, A. (1997) Advances in comparative protein-structure modelling. *Current Opinion in Structural Biology* 7, 206-214.
- (221) Sánchez, R., and Sali, A. (2000) Comparative protein structure modeling, in *Protein Structure Prediction: Methods and Protocols* (Webster, D. M., Ed.) pp 97-129, Humana Press Inc., Totowa, NJ.
- (222) Tramontano, A. (1998) Homology modeling with low sequence identity. *Methods* 14, 293-300.
- (223) Rossmann, M. G., Liljas, A., Branden, C. I., and Bansazak, L. J. (1975) Evolutionary and structural relationships among dehydrogenases, in *The Enzymes* (Boyer, P. D., Ed.) pp 61-102, Academic Press, New York.
- (224) Taroni, C., Jones, S., and Thornton, J. M. (2000) Analysis and prediction of carbohydrate binding sites. *Protein Engineering, Design and Selection* 13, 89-98.

- (225) Vyas, N. K. (1991) Atomic features of protein-carbohydrate interactions. *Current Opinion in Structural Biology* 1, 732-740.
- (226) Sujatha, M. S., Sasidhar, Y. U., and Balaji, P. V. (2004) Energetics of galactose- and glucose-aromatic amino acid interactions: Implications for binding in galactose-specific proteins. *Protein Science* 13, 2502-2514.
- (227) Vogt, G., Woell, S., and Argos, P. (1997) Protein thermal stability, hydrogen bonds, and ion pairs. *Journal of Molecular Biology* 269, 631-643.
- (228) Greenberg, E., and Preiss. (1965) Biosynthesis of bacterial glycogen. II. Purification and properties of the adenosine diphosphoglucose:glycogen transglucosylase of *Arthrobacter* species NRRL B1973. *Journal of Biological Chemistry* 240, 2341-2348.
- (229) Shen, L., and Preiss, J. (1964) The Activation and Inhibition of Bacterial Adenosine-Diphosphoglucose Pyrophosphorylase. *Biochemical and Biophysical Research Communications* 17, 424-429.
- (230) Morrison, R. T., and Boyd, R. N. (1990) *Organic Chemistry, Fifth Edition*, Addison-Wesley Iberoamericana, Wilmington, Delaware.
- (231) Cupp-Vickery, J. R., Igarashi, R. Y., and Meyer, C. R. (2005) Preliminary crystallographic analysis of ADP-glucose pyrophosphorylase from *Agrobacterium tumefaciens*. *Acta crystallographica. Section F, Structural biology and crystallization communications* 61, 266-8.



MICHIGAN STATE UNIVERSITY



3 1293 02845