University of the Basque Country
Sciences and Technology Faculty

STRUCTURAL BASIS OF GLYCOGEN BIOSYNTHESIS REGULATION IN BACTERIA

by

Natalia Comino González

Thesis Supervisor: Marcelo E. Guerin and Javier Cifuente

Structural Glycobiology Group
CIC bioGUNE
Center for Cooperative Research in Biosciences

ZTF-FCT

Universidad del País Vasco
Euskal Herriko Unibertsitatea

Natalia Comino González, Bizkaia (Spain), 2017

Distributor: Department of Biochemistry and Molecular Biology – University of the Basque Country

“We live on an island surrounded by a sea of ignorance. As our island of knowledge grows, so does the shore of our ignorance.”

John Archibald Wheeler

Agradecimientos

A mi director de Tesis Doctoral, Marcelo E. Guerin, gracias por darme la oportunidad de poder formar parte de tu laboratorio, donde a lo largo de estos años he crecido tanto a nivel profesional como personal.

A mi codirector de tesis, Javier Cifuente, del que cada día he aprendido no una, sino muchísimas cosas nuevas y que siempre ha estado dispuesto a ayudarme en todo lo que ha podido. Me ha encantado formar equipo contigo.

A la Unidad de Biofísica, el departamento de Bioquímica y Biología Molecular de la Universidad del País Vasco (UPV/EHU) y el CIC bioGUNE, instituciones donde se ha llevado a cabo este trabajo. Al Contrato de la Comisión Europea LSHP-CT-2005-018923, la Fundación Biofísica Bizkaia (FBB) y el CIC bioGUNE por su financiación.

A todos los compañeros del *Structural Glycobiology Group* (SGP): David Albesa-Jové, Beatriz Trastoy, Mª Ángela Sainz, Montse Tersa y Alberto Marina, gracias por todo vuestro apoyo, ayuda y consejos. A Saioa Urresti, de la cual tuve la oportunidad de aprender a lo largo de mi primer año de doctorado, gracias por tu paciencia y por todo lo que me enseñaste. Quisiera dar un agradecimiento muy especial a Ane Rodrigo y Rubén Sánchez, con los que he podido compartir parte de esta experiencia, época que recordaré siempre con muchísimo cariño. Por supuesto, muchísimas gracias a todas aquellas personas que en algún momento formaron parte del laboratorio o estuvieron involucradas de alguna manera en esta tesis; os deseo lo mejor en el futuro, ha sido un verdadero placer trabajar con vosotros.

Mis últimos agradecimientos son para mi familia, especialmente para mis padres. Sin su esfuerzo y apoyo no hubiera podido llegar hasta aquí, por eso y mucho más este trabajo también les pertenece; y por último a Dani, quien me ha acompañado a lo largo de esta etapa y con el que empiezo, con mucha ilusión, la próxima.

Publications

This thesis is based on the following articles, which will be referred to in the 4th and 5th chapters. A full copy of the articles has been incorporated as Annex I and II.

Cifuente JO*, **Comino N***, Madariaga-Marcos J*, López-Fernández S, García-Alija M, Agirre J, Albesa-Jové D, Guerin ME. **Structural Basis of Glycogen Biosynthesis Regulation in Bacteria**. *Structure*. 24(9):1613-22. doi: 10.1016/j.str.2016.06.023. (2016). *co-first author.

Comino N*, Cifuente JO*, Marina A, Orrantia A, Eguskiza A, Guerin ME. **Mechanistic insights into the allosteric regulation of bacterial ADP-glucose pyrophosphorylases.** *J. Biol. Chem*. pii: jbc.M116.773408. doi: 10.1074/jbc.M116.773408. (2017). *co-first author.

Other publications:

Albesa-Jové D, Romero-García J, Sancho-Vaello E, Contreras FX, Rodrigo-Unzueta A, **Comino N**, Carreras-González A, Arrasate P, Urresti S, Planas A, Guerin ME. **Structural snapshots of the mycobacterial retaining glycosyltransferase GpgS unravel the complete catalytic cycle in native state**. Submitted to *Structure.*

Rodrigo-Unzueta A, Martinez M, **Comino N**, Alzari PM, Chenal A, Guerin ME. **Molecular basis of membrane association of the phosphatidylinositol mannosyltransferase PimA from mycobacteria**. *J. Biol. Chem*. doi: 10.1074/jbc.M116.723676. (2016).

Albesa-Jové D*, **Comino N***, Tersa M*, Mohorko E, Urresti S, Dainese E, Chiarelli LR, Pasca MR, Manganelli R, Makarov V, Riccardi G, Svergun DI, Glockshuber R, Guerin ME. **The Redox State Regulates the Conformation of Rv2466c to Activate the Antitubercular Prodrug TP053**. *J Biol. Chem*. 290(52):31077-89. (2015). *co-first autor.

Albesa-Jové D, Mendoza F, Rodrigo-Unzueta A, Gomollon-Bel F, Cifuente JO, Urresti S, **Comino N**, Gomez H, Romero-Garcia J, Lluch JM, Sancho-Vaello E, Biarnés X, Planas A, Merino P, Masgrau L, and Guerin ME. **A native ternary complex trapped in a crystal reveals the catalytic mechanism of a retaining glycosyltransferase**. *Angew Chem Int Ed Engl*. 54(34):9898-9902 (2015).

Giganti D, Albesa-Jové D, Urresti S, Rodrigo-Unzueta A, Martínez MA, **Comino N**, Barilone N, Bellinzoni M, Chenal A, Guerin ME, and Alzari PM. **Secondary structure reshuffling modulates glycosyltransferase function at the membrane.** *Nat. Chem. Biol.* 11:16-8 (2015). News and Views Membrane enzymes: Transformers at the interface. Brodhun F, Tittmann K. *Nat. Chem. Biol.* 11:102-3 (2015).

Albesa-Jové D, Chiarelli LR, Makarov V, Pasca MR, Urresti S, Mori G, Salina E, Vocat A, **Comino N**, Mohorko E, Ryabova S, Pfieiffer B, Lopes Ribeiro AL, Rodrigo- Unzueta A, Tersa M, Zanoni G, Buroni S, Altmann KH, Hartkoorn RC, Glockshuber R, Cole ST, Riccardi G, Guerin ME. **Rv2466c mediates the activation of TP053 to kill replicating and non-replicating Mycobacterium tuberculosis**. *ACS Chem. Biol*. 9(7):1567-75. (2014).

Giganti D, Alegre-Cebollada J, Urresti S, Albesa-Jové D, Rodrigo-Unzueta A, **Comino N**, Kachala M, López-Fernández S, Svergun DI, Fernández JM, Guerin ME. **Conformational plasticity of the essential membrane-associated mannosyltransferase PimA from mycobacteria**. *J. Biol. Chem*. 288(41):29797- 808 (2013).

Resumen

Los polímeros de glucosa constituidos por enlaces α-1,4 y ramificaciones en α-1,6 (α-glucanos), constituyen una de las principales fuentes de reserva de carbono y energía en los organismos vivos. Los polímeros más representativos son el glucógeno, en animales, hongos, bacterias y arqueabacterias, y el almidón en organismos fotosintéticos. La ventaja de utilizar polisacáridos como compuestos de reserva es que, gracias a sus propiedades físicas y su alto peso molecular, representan una fuente osmóticamente neutra de glucosa dentro de las células. La principal diferencia entre ambos polímeros es su patrón de ramificación, el cual determina en última instancia la morfología y las propiedades físicas del gránulo resultante.

La biosíntesis de glucógeno en bacterias, al igual que en eucariotas, puede dividirse principalmente en tres fases: iniciación, elongación y ramificación. A diferencia de las células eucariotas donde la iniciación de la síntesis de glucógeno se produce mediante la autoglucosilación y la posterior acción cebadora de la enzima glucogenina, en bacterias se ha postulado que la glucogeno sintasa (GS) sería la enzima responsable de ambos procesos. Esta enzima cataliza la transferencia del residuo glucósido de un nucleótido azúcar activado al extremo no reductor de una cadena preexistente, generando un enlace α-1,4-glucosídico. Posteriormente, la enzima ramificadora (GBE) cataliza la transferencia de un segmento de cadena de α-1,4-glucano a posición α-1,6, introduciendo de este modo las ramificaciones características de este polímero. El anabolismo del glucógeno, por el contrario, se produce mediante la acción del enzima glucógeno fosforilasa (GP) que cataliza la eliminación secuencial de residuos glucosídicos desde los extremos no reductores de la molécula de glucógeno y la enzima desramificante de glucógeno (GDE) que cataliza la hidrólisis de los puntos de ramificación.

La síntesis de glucógeno requiere, por lo tanto, de un nucleótido azúcar activado que actúe como sustrato de la GS. En el caso de la síntesis de glucogeno en bacterias y almidón en plantas, evolutivamente se ha seleccionado el nucleótido azúcar ADP-glucosa (ADP-Glc) para cumplir esta función. La síntesis de ADP-Glc se lleva a cabo a través de la condensación de glucosa 1-fosfato (G1P) y ATP en presencia del catión divalente magnesio, en una reacción catalizada por la enzima ADP-glucosa pirofosforilasa (AGPase) y que resulta en la liberación de pirofosfato. La generación de

ADP-Glc en bacterias y plantas supone un paso clave y limitante en la síntesis de glucógeno y almidón, respectivamente, y conlleva un importante gasto energético. Por estos motivos, evolutivamente AGPase ha adquirido propiedades alostéricas a través de metabolitos relacionados con el estado energético celular. La naturaleza de estos metabolitos puede variar en función del organismo, pero todos comparten la característica de ser intermediarios en las principales rutas energéticas de asimilación de carbono. Por consiguiente, los reguladores alostéricos positivos representarían estados de alto contenido energético mientras que los reguladores negativos reflejarían niveles metabólicos bajos de carbón y energía.

Tanto en plantas como en bacterias AGPase es una proteína tetramérica; en el caso de plantas corresponde a un heterotetrámero compuesto por dos subunidades largas y dos cortas ($\alpha_2\beta_2$) mientras que las AGPases de origen bacteriano son homotetrámeros constituidos por protómeros de aproximadamente 50 kDa. Estos protómeros, al igual que en otros miembros de la superfamilia de nucleotidil transferasas, están constituidos por un dominio catalítico α-β-α sándwich N-terminal y un dominio regulador hélice-β C-terminal. Previamente a este trabajo, solamente dos estructuras cristalinas de AGPase habían sido resueltas; (i) la de *Agrobacterium tumefaciens* y (ii) la de *Solanum tuberosum*. En este último caso, la estructura depositada corresponde a un homotetrámero no fisiológico conformado únicamente por subunidades pequeñas (α_4). La ausencia de estructuras en complejo con reguladores alostéricos ha dificultado el estudio del mecanismo por el cual la unión de estos metabolitos regula alostericamente la actividad catalítica de la enzima. En este sentido, la identificación de los sitios alostéricos en AGPase ha sido uno de los principales desafíos en el campo a lo largo de las últimas décadas. La AGPase de *Escherichia coli* (*Ec*AGPase) ha sido el principal modelo de estudio de AGPases, no solo en relación a su actividad catalítica sino también al entendimiento de sus propiedades alostéricas. Específicamente, los reguladores alostéricos fiosiológicos de *Ec*AGPase son fructosa-1,6-bifosfato (FBP) y adenosina-difosfato (ADP), regulador alostérico positivo y negativo respectivamente. Diversos estudios bioquímicos sugieren que tanto el dominio N- como C-terminal estarían involucrados en la regulación alostérica de AGPase, sin embargo, el mecanismo a nivel molecular por el cual se produce esta regulación sigue siendo una incógnita.

Con el objetivo de profundizar en esta dirección, hemos resuelto la estructura cristalina de *Ec*AGPase en complejo con su regulador alostérico positivo (*Ec*AGPase•FBP) y negativo (*Ec*AGPase•AMP) a una resolución de 2.67 y 3.04 Å, respectivamente. La estructura cristalina de *Ec*AGPase corresponde a un homotetrámero que puede considerarse un dímero de dímeros. Los sitios de unión de ambos reguladores se encuentran parcialmente superpuestos en una profunda fisura entre los dominios N- y C- terminales de protómeros vecinos. El regulador negativo AMP se encuentra enterrado en la hendidura alostérica, con el fosfato α orientado hacia una cavidad rica en residuos cargados positivamente. De las interacciones derivadas de la unión a AMP resalta la fijación del heterociclo de adenina a través de la cadena lateral del residuo Arg130. Este importante residuo se encuentra en la hélice α7 del protómero vecino, que conforma a su vez un dímero distinto, hecho que aporta claras evidencias de interacciones inter-protoméricas derivadas de la unión a AMP. Para comprobar esta hipótesis, dado que este tipo de interacciones podrían llevar a la estabilización de la estructura cuaternaria de *Ec*AGPase en solución, decidimos estudiar su estabilidad térmica por dicroísmo circular. Gracias a estos ensayos biofísicos determinamos que *Ec*AGPase en complejo con AMP es 4.6 ºC más estable que su forma *apo,* resultado que concuerda con las observaciones estructurales. Por otro lado, FBP se une a la misma hendidura, pero en una localización más expuesta al solvente donde la Arg130 ha sufrido un importante cambio conformacional. Esta variación/modificación de las interacciones inter-protoméricas queda reflejada en los estudios de dicroísmo circular donde la unión del regulador positivo no induce la estabilización térmica de la enzima. Esta configuración estructural de *Ec*AGPase en la cual ambos reguladores se unen en localizaciones parcialmente superpuestas, explica el hecho de que la sensibilidad a la inactivación por AMP se ve afectada por la concentración de FPB. En este trabajo pudimos determinar asimismo que la adición de FBP al complejo *Ec*AGPase•AMP provoca una reversión de la estabilización térmica, lo que indica que FBP no solo es capaz de desplazar el AMP sino también de modificar el estado conformacional inducido por este regulador hacía una estructura menos estable.

El estudio detallado de las estructuras cristalinas determinó que cada hendidura alostérica comunica con el correspondiente sitio activo del mismo protómero a través de un elemento estructural que denominamos “Sensory Motif” (SM). Este motivo estructural está constituido por el loop de unión al nucleótido (NBL), el cual incluye la secuencia consenso GGxGxR involucrada en la unión del ATP, y una región compuesta por

diversos elementos de estructura secundaria relativamente cortos. La presencia de un loop flexible en el mismo protómero, denominado “Regulatory Loop 1” (RL1), que interacciona tanto con el NBL como con los elementos de estructura secundaria en el SM, indica una posible participación en la modulación de la conformación adoptada por el SM. Como se ha comentado anteriormente, la unión de AMP involucra una fuerte interacción con la cadena lateral del residuo Arg130 localizado en un protómero vecino de un dímero distinto. La hélice en la que se encuentra este residuo, α7, conecta a su vez con un loop denominado “Regulatory loop 2” (RL2) que flanquea el sitio de unión de ATP en el sitio activo. Teniendo en cuenta toda la información estructural reportada, proponemos un modelo de regulación alostérica de *Ec*AGPase en el que la unión de los reguladores alostéricos positivo y negativo modularían la actividad enzimática a través del SM y dos loop regulatorios, RL1 y RL2, mediante interacciones intra- e inter-protoméricas.

Con el objetivo de validar el sitio de unión de ambos reguladores alostéricos, diseñamos mutantes puntuales de residuos involucrados en la interacción con ambos moduladores, para posteriormente medir la estabilización térmica, actividad específica y contenido de glucógeno “*in vivo”* de cada una de las variantes. La mutación de residuos involucrados en la unión de AMP claramente impide la estabilización mediada por el inhibidor mientras que las mutantes de residuos implicados en la unión de FBP fueron incapaces de revertir el efecto estabilizador de AMP en presencia del activador. Las medidas de actividad específica mostraron que la mutante *Ec*AGPase•R130A corresponde a una variante en la que la inhibición por AMP se encuentra desregulada, hecho que concuerda con la información estructural que indica un papel fundamental de este residuo en el mecanismo inter-protomérico de transducción de la señal alostérica. Al transformar una cepa *E.coli* inactiva para el gen de *Ec*AGPase con las distintas mutantes puntuales, pudimos determinar que las actividades enzimáticas medidas para *Ec*AGPase•R130A se traducen en un incremento de la cantidad de glucógeno acumulado por las bacterias transformadas con esta variante. Posteriormente, la estructura cristalina de *Ec*AGPase•R130A se resolvió en ausencia de ligandos a una resolución de 3.09 Å. Al comparar esta nueva estructura con la correspondiente a los complejos *Ec*AGPase•AMP y *Ec*AGPase•FBP se observan claros cambios conformacionales en elementos implicados en el mecanismo alostérico propuesto. Parte del SM se encuentra en una conformación extendida invadiendo parcialmente el sitio de unión de AMP y tanto el RL1 como el RL2 adoptan conformaciones distintas a las

observadas para la enzima salvaje. Curiosamente, la estructura de *Ec*AGPase•R130A presenta una reorientación de los dímeros que conforman el homotetrámero, lo que indica una clara flexibilidad de la estructura cuaternaria de *Ec*AGPase.

Dado que diversos resultados experimentales apuntan a cambios de estructura cuaternaria asociados al mecanismo alostérico de *Ec*AGPase, decidimos focalizamos en el estudio de la estructura cuaternaria de la enzima en solución. A través de HPLC-SAXS determinamos que el complejo *Ec*AGPase•AMP es más homogéneo y compacto que el correspondiente a FBP, observándose en este último una clara pérdida de globularidad. La comparación de las estructuras cristalinas de ambos complejos con los modelos *ab initio* obtenidos por SAXS sugiere que, la adición de FBP a *Ec*AGPase en solución dispara un importante cambio conformacional a nivel de estructura cuaternaria, que no es observable por cristalografía. Con el objetivo de validar esta hipótesis resolvimos la estructura de ambos complejos por cryo-EM. Los modelos obtenidos apuntan de nuevo a cambios de estructura cuaternaria, observándose una reorientación de los dímeros en el complejo activado. Basándonos en estos resultados, la estructura cristalina del complejo *Ec*AGPase•AMP correspondería a la conformación tensa (T) de baja actividad definida por el modelo Monod-Wyman-Changeux (MWC), mientras que la estructura cristalina del complejo *Ec*AGPase•FBP representaría un primer evento de unión de FBP a este mismo estado. El estudio del estado relajado (R) de alta actividad en *Ec*AGPase requerirá en un futuro la determinación de la estructura del complejo *Ec*AGPase•FBP en solución a alta resolución.

Common Abbreviations and Symbols

Å	Ångström
ΔA	Increment of Absorbance
ΔCp	Increment of Heat capacity at constant pressure
ΔH	Increment of Enthalpy
ΔH_{VH}	Increment of Van't Hoff Enthalpy
ΔG	Increment of Gibbs free energy
ΔS	Increment of Entropy
λ	Wavelenght
θ	Ellipticity
s	second
3PGA	3-phosphoglyceric acid
ADP	Adenosine diphosphate
ADP-Glc	Adenosine diphosphate-glucose
AGPase	ADP-glucose pyrophosphorylase
AMP	Adenosine monophosphate
At	*Agrobacterium tumefaciens*
ATP	Adenosine triphosphate
BE	Branching enzyme
CA	Cycloalternan

CAFE	Cycloalternan forming enzyme
CBM	Carbohydrate-binding module
CCD	Charge-Coupled Device
CD	Circular dichroism
CGT	Cyclodextrin glucanotransferase
CIF	Crystallographic Information File
CMOS	Complementary Metal-Oxide Semiconductor
Cryo-EM	Electron cryomicroscope
CSR	Conserved sequence regions
CTF	Contrast Transfer Function
DC-SIGN	Dendritic Cell-Specific intracellular adhesion molecule-3-Grabbing Non-integrin
DED	Direct Electron Detectors
DLS	Diamond Light Source
dTTP	deoxythymidine triphosphate
Ec	*Escherichia coli*
ECA	Enterobacterial common antigen
EM	Electron Mycroscopy
EX_7E	Glu – X_7 – Glu motif
f_u	Fraction of unfolded protein
F6P	Fructose-6-phosphate

FBP ..... Fructose 1,6-biphosphate

FEG ..... Field emission gun

FWHM ..... Full Width at Half Maximum

G1P Glucose-1-phosphate

GH ..... Glycosyl hydrolase

GlcNAc ..... *N*-acetyl glucosamine

GDP ..... Guanosine diphosphate

GDP-Man ..... Guanosine diphosphate mannose

GN ..... Glycogenin

LβH ..... Left-handed parallel β-helix

G4 ..... Maltotetraose

G5 ..... Maltopentose

GDE ..... Glycogen debranching enzyme

GK ..... Glycerol kinase

GlcN ..... Glucosamine

GlgE ..... α-1,4-glucan:maltose-1-phosphate maltosyltransferase

GlmU ..... N-acetylglucosamine-1-phosphate uridyltransferase

GP ..... Glycogen phosphorylase

GS ..... Glycogen synthase

GSD ..... Glycogen storage diseases

GSG4 ..... 4-S-a-D-glucopyranosyl-4-

	thiomaltotetraose
GT	Glycosyltransferase
GT-A	Glycosyltransferase A fold
GT-B	Glycosyltransferase B fold
GT-C	Glycosyltransferase C fold
GTP	Guanosine triphosphate
HEPPSO	N-(2-hydroxyethyl)piperazine-(2-hydroxypropanesulfonic acid)
HPLC	High performance liquid chromatography
HPr	Histidine phosphocarrier protein
K	Equilibrium constant
IPTG	isopropyl β-thiogalactopyranoside
KTGGL	Lys-Tyr-Gly-Gly-Leu
LaB_6	Lanthanum hexaboride cathode
LM	Light microscopy
LCP	Left-handed polarized light
LDH	Lactate dehydrogenase
M1P	Maltose-1-phosphate
MalP	Maltodextrin phosphorylase
MGLP	Methylglucose lipopolysaccharide
MurG	D-inositol 3-phosphate glycosyltransferase
Mtb	*Mycobacterium tuberculosis*

MurG..... Pyrophosphoryl-undecaprenol N-acetylglucosamine transferase

MWC..... Monod-Wyman-Changeaux

N ..... Native state

NA..... Numerical aperture

NBL ..... nucleotide-binding loop

NDP ..... Nucleotide diphosphate

NMR ..... Nuclear magnetic resonance

N-terminal / C-terminal ..... Amino / Carboxyl terminal domain

OtsA..... Trehalose-6-phosphate synthase

Pa ..... *Pseudomonas aeruginosa*

PDB ..... Protein Data Bank

PimA.....Phosphatidyl-myo-inositol mannosyltransferase A

PK..... Pyruvate kinase

PLP..... Pyridoxal-5-phosphate

PO_4 ..... Phosphate

PPi..... Phyrophosphate

r ..... Resolution

R-state..... 'relaxed' high-activity/affinity state

RCP ..... Right-handed polarized light

REM..... Reflection electron microscopy

Rg..... Radii of gyration

RL1 ..... Regulatory loop 1

RL2 ..... Regulatory loop 2

r.m.s.d. ..... Root Mean Square Deviation

ROD ..... Rotatory Optical dispersion

SAXS ..... Small angle X-ray scattering

SBS ..... Surface/secondary binding sites

SEM ..... Scanning electron microscopy

SM ..... Sensory motif

SO_4 ..... Sulfate

SS ..... Starch synthase

S_N2 ..... Bimolecular nuclephilic substitution

S_ni ..... Intramolecular nucleophilic substitution

SNR ..... Signal to noise ratio

SNT ..... Sugar nucleotidyltransferase

St ..... *Solanum tuberosum*

T ..... Temperature

T-state ..... 'tense' low-activity/affinity state

TCA ..... Trichloroacetic acid

TDP-D-glucose ..... di-oxythymidine diphosphated-D-glucose

TG2 ..... Transglutaminase type 2

T_m ..... Temperature of melting

TEM ..... Transmission electron microscopy

T_R ..... Retention time

TreS ..... Trealose synthase

Tris ..... Tris-(hydroxymethyl)aminomethane

U ..... Unfolded state

UDP ..... Uridine diphosphate

UTP ..... Uridine triphosphate

UDP-Glc ..... Uridine diphosphate glucose

UDP-GlcNAc ..... Uridine diphosphate *N*-acetylglucosamine

UV ..... UltraViolet

WT ..... Wild Type

INDEX

INTRODUCTION

‘

1. INTRODUCCIÓN

1.1. Glycogen as a Reserve Storage Compound

The biosynthesis of α-1,4-linked and α-1,6-branched glucose polymers (α-polyglucans), defines the most widespread form of storage polysaccharides in living cells and represents a major carbon an energy reserve form in nature. The most representative α-polyglucans are glycogen and starch (Jack Preiss, 2006; Wilkinson, 1963). Glycogen is found in a majority of animal, fungal, bacterial, and archeabacterial species as discrete cytoplasmic granules, whereas starch is only found in photosynthetic eukaryotes (Figure 1.1; S. Ball et al. 1996; S. G. Ball and Morell 2003; Pfister and Zeeman 2016). The advantage of using polysaccharides as storage reserves is that they are osmotically neutral means to store glucose in cells due to their high molecular weights (Roach, Depaoli-Roach, Hurley, & Tagliabracci, 2012a).

Figure 1.1. Glycogen accumulation in ΔglgP E. coli mutant. *The arrows indicate the electron-transparent cytoplasmatic glycogen granules preferentially located in the periphery of the bacteria (Modified from Alonso-Casajús et al. 2006)*

Either the accumulation or deficiency of glycogen in humans causes glycogen storage diseases (GSDs). Commonly, GSD is produced by defective enzymes involved in glycogen metabolism and regulation, displaying distinctive phenotypes depending on the enzyme affected (Hicks, Wartchow, & Mierau, 2011). In this regard, inactivation of the muscle glycogen synthase, the enzyme involved in the α-1,4-linkages elongation, causes glycogen depletion in skeletal and heart muscle, resulting in muscle weakness and cardiomyopathy (Moslemi et al., 2010). On the other hand, deficiencies on the

enzyme responsible of the depolymerization of glycogen at the α-1,6 branching points, result in the accumulation of abnormal glycogen, impairing the function of certain organs and tissues (Cori disease; Zhai et al. 2016).

Wilkinson proposed that for a compound to be classified as having energy-storage function, it must satisfy three criteria (Wilkinson, 1959):

i. the compound should accumulate intracellularly under conditions in which the energy supply for growth of the organism is in excess,
ii. it must be utilized when supplies of energy from exogenous sources are no longer available for maintenance of growth or other processes necessary to sustain viability,
iii. the proposed storage compound must be degraded to an energy form used for some purpose which gives the cell a biological advantage to survive or attend better in the media or environment.

The occurrence of glycogen-like reserves has been reported in over 50 different bacterial species including Gram-negative, Gram-positive, archaebacteria and photosynthetic bacteria, a number of these observed as inclusion bodies of variable diameter (Iglesias & Preiss, 1992; Leadbetter & Poindexter, 1989; Jack Preiss, 2014; Shively, 1974). Based on several experimental observations bacterial glycogen fits these criteria. Its biosynthesis usually occurs in the presence of an excess carbon source and under environmental conditions of slow growth or no growth. Thus, glycogen accumulation has been shown to occur preponderantly in the stationary phase of the growth cycle due to limitations in an essential nutrient such as sulphur, nitrogen, or phosphate (M. a Ballicora, Iglesias, & Preiss, 2003; J Preiss, 1984, 2010). In this regard, *Escherichia coli* B cultures having glucose as carbon source and limited with respect to nitrogen show an inverse relationship in the rate of growth and the quantity of glycogen accumulated (Holme, 1957).

Mutants of different bacterial species, which are deficient in glycogen accumulation due to mutations in key enzymes involved on this polysaccharide synthesis, grow as well as their wild type parent strains indicating that glycogen might not be required for bacterial growth (Govons, Gentner, Greenberg, & Preiss, 1973; Steiner & Preiss, 1977). However, a more prolonged survival rate is observed in the cells

accumulating the polysaccharide compared with the glycogen-less mutants (Strange, 1968). These observations suggest that under conditions of no available carbon source, glycogen is probably utilized to preserve cell integrity, providing the energy required by the bacteria for maintenance. This 'energy of maintenance' is the energy needed for processes such as turnover of proteins and RNA, maintenance of motility and intracellular pH, chemotactic response and osmotic regulation (Wilkinson, 1963).

1.2. Bacterial glycogen structure

According to the currently accepted model, bacterial glycogen is similar to mammalian glycogen; it is a very large branched water-soluble polymer consisting of α-D-glucosyl units, these units are connected together by α-1,4-glucosidic linkages to form linear oligosaccharide chains, with α-1,6-glucosidic linkages at branching points (Figure 1.2; Manners, 1991).

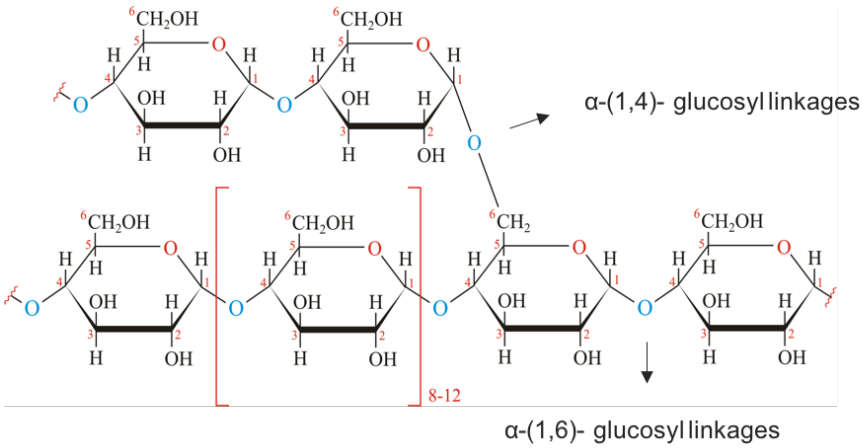


Figure 1.2. Glycogen structure. *Glycogen is a glucose homopolymer composed of approximately 90% of α-1,4-glucosidic linkages and 10% of branching α-1,6-glucosidic linkages.*

Glycogen isolated from biological sources is polydisperse, existing as a population of molecules of different sizes. In addition, α-1,6-linkages are localized in random positions and their percentage can vary depending on the bacterium and possibly on the stage of glycogen synthesis (Roach, Depaoli-Roach, Hurley, &

Tagliabracci, 2012b). It has been observed that in *Mycobacterium smegmatis* the degree of branching decreases as the glycogen granules increase in size (Manners, 1991). The molecular weights of glycogen isolated from different bacteria can also vary. However, the particular method used to isolate glycogen can significantly affect its molecular weight. The common way to analyze the chemistry of glycogen molecules is to define molecular mass distributions, average chain lengths and average branching frequencies. In bacteria, these parameters can vary depending on the source but the average length of the chains is usually of 8–12 glucose units, the α-1,6-branches accounts for 7–10% of the linkages and the molecular size of glycogen has been estimated to be about 10^7–10^8 Da (S. Ball, Colleoni, Cenci, Raj, & Tirtiaux, 2011; Leadbetter & Poindexter, 1989; Manners, 1991).

The storage of high amounts of free glucose within the cell would create an osmotic imbalance, for that reason, glucose units must be densely packed to increase the solubility of this polysaccharide (Figure 1.3A). Structurally, its branching pattern allows for spherical growth of the particle leading to a progressively more packed structure towards the periphery (Figure 1.3B; S. Ball et al. 2011). The size and density of the glycogen granule is self-limited to allow the interaction of the glycogen chains with the catalytic sites of the enzymes involved in its metabolism (Shearer & Graham, 2002). Mathematical modelling predicts a 42 nm maximal diameter of the glycogen granule including 55 000 glucose residues (Figure 1.3A). A higher percent of this total number of glucose residues localizes in the outer (unbranched) shell and is thus readily accessible to glycogen catabolism without debranching. *In vivo*, glycogen particles are thus present in the form of these limit size granules (macroglycogen) and also smaller granules representing intermediate states of glycogen biosynthesis and degradation (proglycogen) (Shearer & Graham, 2002). Based on its physical properties, glycogen particles are entirely hydrosoluble and, therefore, define a state where the glucose becomes less active osmotically but easily accessible for rapid mobilization through the enzymes of glycogen catabolism as if it were in the soluble phase (S. Ball et al., 2011).

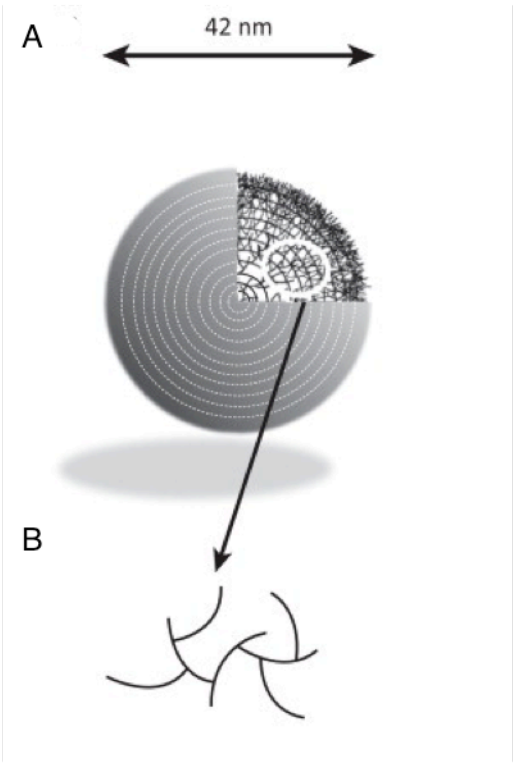


***Figure 1.3. (A, B) The schematic representation of whole glycogen granule**. (B) Enlarged views of the circled sections of the glycogen granule. The distribution of branches exemplified in (B), with two α-1,6 linkages per glucan, allows the exponential increase in the density of the particle. This leads to a predictable maximum of 42 nm for the glycogen granule displayed in (A) (Modifed from S. Ball et al. 2011).*

It is worth noting that glycogen and starch display the same chemical building unit, α-D-glucosyl units connected in both α-polyglucans by the same types of covalent bonds, displaying similar ranges of molecular mass (Manners, 1991). Starch is composed of two major polysaccharides, amylose and amylopectin. Amylose is an essentially linear α-1,4-glucan with very few α-1,6-branches (less than 1%) whereas amylopectin, the major compound, consists of α-1,4-glucan chains with approximately 5% of α-1,6-branches (Figure 1.4; Iglesias and Preiss 1992; S. Ball et al. 1996). Despite the similarities between glycogen and starch, their morphology and physical properties are quite different; due to their differences in the branch structure and molecule shape, starch defines insoluble semi-crystalline granules with sizes ranging 0.1-100 μm (S. Ball et al., 2011).

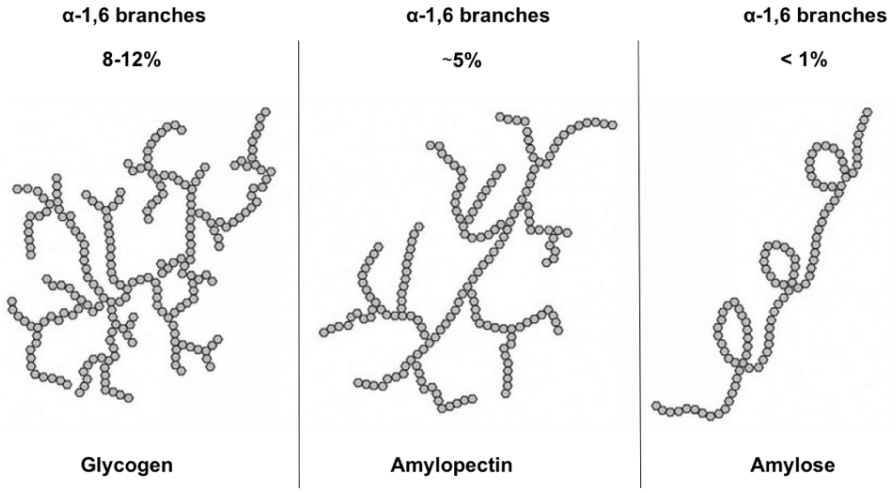


Figure 1.4. Glycogen and starch components, amylopectin and amylose, structure*. The degree of branching in glycogen is higher than in amylopectin, been practically inexistent in the case of amylose. The difference in branching pattern between glycogen and starch confers distinctive structural properties to these carbohydrate granules*

1.3. Glycogen metabolism in bacteria

In bacteria, the basic glycogen biosynthetic pathway involves the action of three enzymes: adenosine diphosphate-glucose pyrophosphorylase (AGPase; EC 2.7.7.27), glycogen synthase (GS; EC 2.4.1.21), and branching enzyme (BE; EC 2.4.1.18) (Jack Preiss, 2014). It is well established that the NDP-sugar donor for bacterial glycogen synthesis is adenosine diphosphate-glucose (ADP-Glc). ADP-Glc is synthesized from adenosine triphosphate (ATP) and α-glucose-1-phosphate (α-Glc-1-P; G1P), with the liberation of pyrophosphate (PPi), via a reaction catalyzed by the enzyme AGPase (Reaction 1) (M. a Ballicora et al., 2003).

$$ATP + G1P \leftrightarrow ADP\text{-}Glc + PPi \quad (1)$$

The next step of bacterial glycogen synthesis consists in the transfer of the activated glucose to the nonreducing end of a growing α-1,4-linked glucan (Reaction 2). A specific glycosyltransferase, GS, catalyzes this reaction (Buschiazzo et al., 2004). To date, it has been postulated that GS is able to initiate *de novo* glycogen biosynthesis

from ADP-Glc or use a pre-existent α-glucan primer as substrate for further elongation (Ugalde, Parodi, & Ugalde, 2003). However, the proposed molecular mechanism still remains a major challenge in the field.

$$\text{ADP-Glc} + (\alpha\text{-1,4-glucan})_n \rightarrow (\alpha\text{-1,4-glucan})_{n+1} + \text{ADP} \quad (2)$$

As indicated before, 10% of the total linkages found in glycogen are α-1,6-glucosyl bonds. They are formed in a reaction catalyzed by the BE (Reaction 3) which hydrolyzes an α-1,4-linkage within a pre-existing α-1,4-linked glucan and transfers a segment of chain in α-1,6 position generating a branch point (Feng et al., 2015).

$$\alpha\text{-1,4-glucan} \rightarrow \alpha\text{-1,6-branched } \alpha\text{-1,4-glucan} \quad (3)$$

In contrast, glycogen degradation is carried out by glycogen phosphorylase (GP; EC 2.4.1.1), which functions as a depolymerizing enzyme, and the debranching enzyme (DBE; EC 3.2.1.-) that catalyzes the removal of α-1,6-linked ramifications.

In bacteria, the presence of glycogen is associated with a minimum of one AGPase, one GS, one GP, one BE, and one GDE (Figure 1.5). These enzymes constitute the core pathway of glycogen metabolism and it is conserved among a broad range of bacterial species (Jack Preiss, 2014).

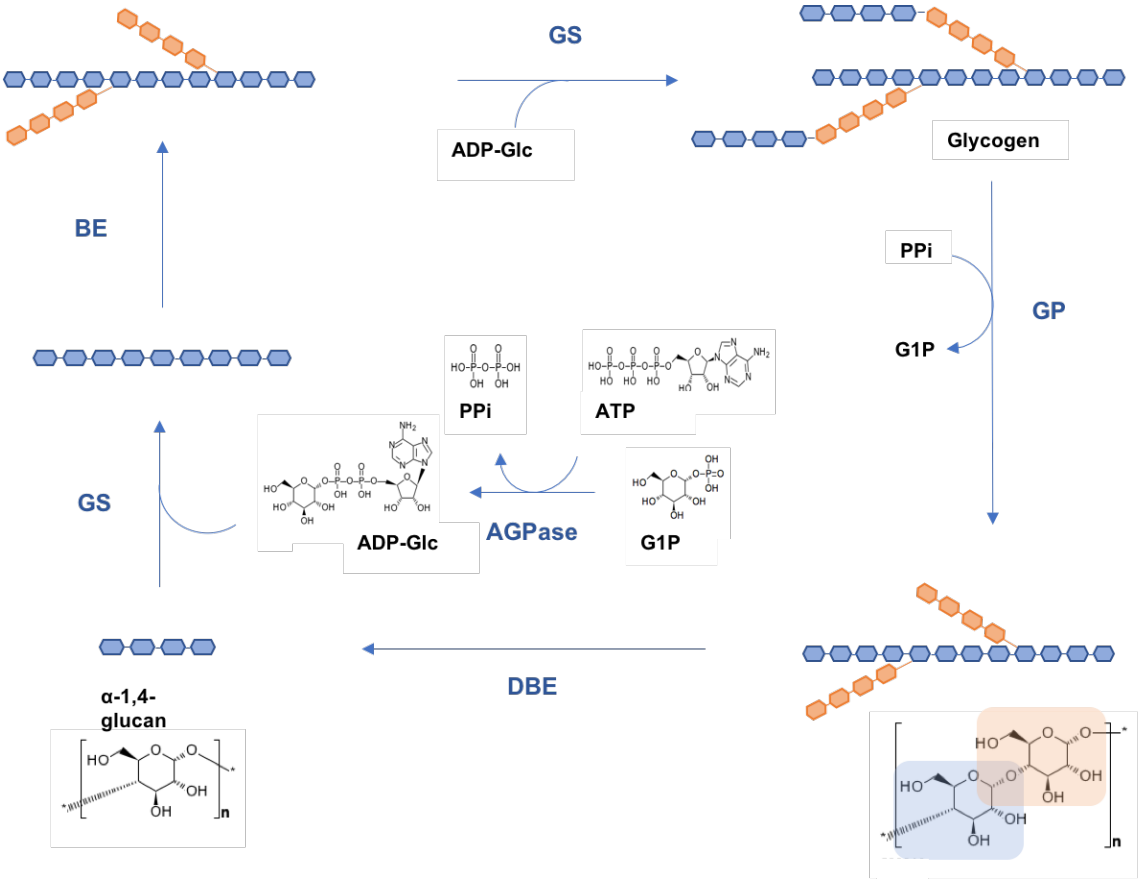


Figure 1.5. Bacterial glycogen metabolism. *Bacterial glycogen biosynthetic pathway involves the action of three enzymes: adenosine diphosphate-glucose pyrophosphorylase (AGPase; EC 2.7.7.27), glycogen synthase (GS; EC 2.4.1.21), and branching enzyme (BE; EC 2.4.1.18) whereas its degradation is carried out by glycogen phosphorylase (GP; EC 2.4.1.1) and the debranching enzyme (DBE; EC 3.2.1.-).*

In *Escherichia coli*, these enzymes are located in a cluster of 15Kb organized in two neighboring operons (Figure 1.6) (Jack Preiss, 2006; Romeo, Kumar, & Preiss, 1988). A deep study of this cluster showed that the genes coding for GS (*glgA*), AGPase (*glgC*), BE (*glgB*), isoamylase (*glgX*) and GP (*glgP*), are transcribed as two different operons, *glgBX* and *glgCAP* (Figure 1.6). Specifically, a non-coding region comprising 500pb is located between *glgX* and *glgC* genes and both operons are preceded by promotors regulated during the growth phase (Jack Preiss, 2014).

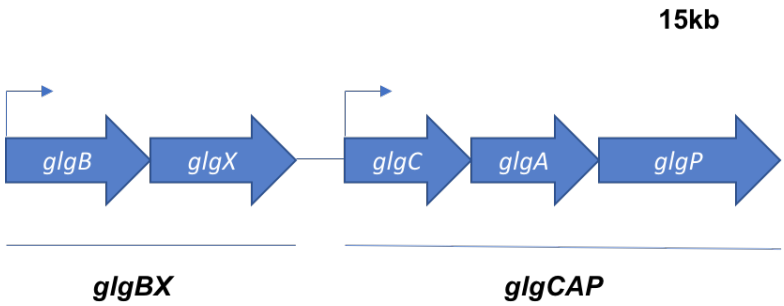


Figure 1.6. Genomic organization of E. coli glycogen metabolic enzymes in two neighboring operons; glgBX and glgCAP.

Most of the glycogen accumulated in bacteria is mediated by the GlgC-GlgA (AGPase-GS) pathway. In this regard, several studies demonstrated that glycogen-deficient or glycogen-excess mutants of several organisms are affected in AGPase, GS or BE activities (Jack Preiss, 2014).

More recently, important progress have been made in determining the physiological function of bacterial α-glucan beyond the classical point of view of glycogen as a storage reserve compound. The existence of an alternative pathway to glycogen synthesis in *Mycobacterium* species, known as the GlgE pathway, has been demonstrated. *Mycobacterium tuberculosis* synthesizes intra- and extracellular α-glucans, been the last the main constituent of the mycobacterial capsule thought to be involved in immune evasion and virulence (Mishra, Driessen, Appelmelk, & Besra, 2011). Specifically, capsular α-glucan has been identified as a novel ligand for the C-type lectin dendritic cell-specific ICAM-3-grabbing nonintegrin (DC-SIGN), involved in host immune modulation (Geurtsen et al., 2017). Originally, three separate and potentially redundant pathways were implicated in mycobacterial α-glucan biosynthesis: (i) the GlgC-GlgA pathway for classical glycogen, (ii) the glycosyltransferase Rv3032 (iii) and the TreS-Pep2-GlgE pathway for trehalose-to-glucan conversion (Figure 1.7; Kalscheuer et al., 2010). However, recent studies have shown that α-glucan in mycobacteria is exclusively assembled intracellularly utilizing the building block α-maltose-1-phosphate (M1P) as the substrate for the α-1,4-glucan:maltose-1-phosphate maltosyltransferase GlgE, with subsequent branching of the polymer by the BE GlgB. Some of this polysaccharide is then exported to form the mycobacterial α-glucan capsule. Interestingly, the mycobacterial GS GlgA is three orders of magnitude more efficient at transferring glucose from ADP-Glc to G1P than to glycogen making little or no glycogen via the

classical GlgC-GlgA biosynthetic pathway. GlgA is, therefore, a M1P-producing glucosyltransferase which generates the activated M1P substrate for the synthesis of linear α-glucan through the GlgE pathway. On the other hand, Rv3032 action appears to be restricted to the synthesis of specialized oligomeric α-glucan derivatives such as methylglucose lipopolysaccharides (MGLPs), as the deletion of Rv3032 gene has no significant impact on the levels of either intracellular or capsular α-glucans. These evidences support the fact that both cytosolic and capsular α-glucan polymers in *Mycobacterium tuberculosis*, *Mycobacterium smegmatis* and probably all other mycobacteria are predominantly, if not exclusively, synthesized by the maltosyltransferase GlgE together with the branching enzyme GlgB (Asención Diez et al., 2015; Koliwer-Brandl et al., 2016; Rashid et al., 2016). Several structures of the enzymes involve in mycobacterial α-glucan biosynthesis have been reported (Caner et al., 2013; Lindenberger, Kumar Veleti, Wilson, Sucheck, & Ronning, 2015; Pal et al., 2010; Roy et al., 2013).

Similarly, an unconventional α-glucan metabolic pathway from the food-borne pathogen *Listeria monocytogenes* has been recently described. It outcomes from the identification of cycloalternan (CA), a cyclic tetrasacharide bacterial product with alternating α-1,3- and 1,6-glucan linkages, and the related extracellular cycloalternan-forming enzyme (CAFE) in this organism. This specialized metabolic function allows the bacteria to extracellularly transform environmental α-glucans into CA, which is transport into the cell and hydrolyzed to glucose. This CA pathway is not essential for maltose/maltodextrin utilization, in fact, most bacteria directly take up prevalent α-glucan and/or secrete an amylase to digest them into mono- or disaccharides that can be imported via general transport mechanism. It is worth noting that *L. monocytogenes* must compete with the host and other members of the microbiota for limited nutrients. It has been proposed that this new pathway could be a way to convert extracellular α-glucans to a form that few competing cells can catabolically process been, therefore, involved in interspecies resource competition within the host gastrointestinal tract (Light, Cahoon, Halavaty, Freitag, & Anderson, 2016).

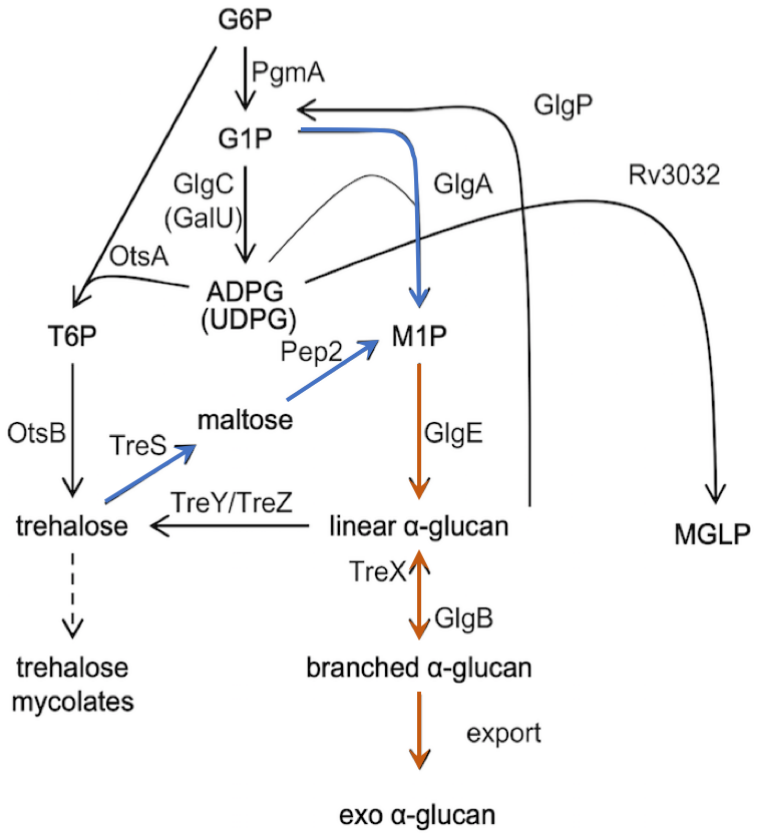


Figure 1.7. Model of GlgE-mediated intracellular and capsular α-glucan biosynthesis in mycobacteria. *The M1P building block is synthesize via two alternative routes; the TreS-Pep2 pathway which involves the isomerization of trehalose to maltose by trealose synthase (TreS, EC 5.4.99.16) and the conversion to M1P by maltokinase (Pep2, EC 2.7.1.175) and the GlgC-GlgA pathway (blue lines). GlgA and the trehalose-6-phosphate synthase (OtsA, EC 2.4.1.15) connect both routes through the shared use of ADP-Glc and the α-1,4-glucan:maltose-1-phosphate maltosyltransferase (GlgE, EC 2.4.99.16) catalyzes the α-glucan polymerization with further addition of α-1,6-branch points by GBE GlgB (orange lines).*

The characterization of novel biosynthetic pathways and functions for bacterial α-glucan which, in some cases, implies the overlap of the canonical GlgC-GlgA pathway enzymes, highlight the evolution of the field (Koliwer-Brandl et al., 2016). In the current work, we will focus our attention on the canonical GlgC-GlgA glycogen biosynthetic pathway not describing in the following sections the enzymes involved in alternative pathways.

1.4. Molecular bases of bacterial glycogen biosynthesis

The glycogen synthesis by the canonical GlgC-GlgA pathway is a three-step process including an initiation, elongation and branching step (J Preiss, 2010; Ugalde et al., 2003; Wilson et al., 2010). This section includes a detailed description of the bacterial enzymes and proposed mechanisms involved in each step of the pathway, providing only brief information regarding the analogous plant/mammal systems as reference.

1.4.1. ADP-Glc Pyrophosphorylase

ADP-Glc serves as NDP-sugar donor for bacterial and plant glycogen/starch biosynthesis and its formation is mediated by AGPase (Figure 1.8; (Recondo & Leloir, 1961; Trivelloni, Recondo, & Cardini, 1962). The activity of AGPase was firstly determined by J. Espada in 1962 (Espada, 1962) and since then, the characterization of this enzyme has been a matter of intense research in the field of glycogen/starch biosynthesis and regulation. Specifically, AGPase catalyzes the reaction between ATP and G1P to produce pyrophosphate and ADP-Glc in the presence of the divalent metal cation Mg^{2+} (Figure 1.8; Ballicora, Iglesias, and Preiss 2003). AGPase activity displays cooperative behavior and a bi-bi mechanism with sequential binding of ATP and G1P, followed by ordered release of pyrophosphate and ADP-Glc (Paule & Preiss, 1971). This reaction is reversible when isolated but in *vivo* it is shifted to ADP-Glc synthesis because of the subsequent hydrolysis of released PPi by ubiquitous pyrophosphatases in the organism (Kornberg, 1962). Evolution has led AGPase to acquire allosteric properties to control this key rate-limiting step of the glycogen biosynthetic pathway by essential metabolites in the energetic flux within the cell (J Preiss, 1978).

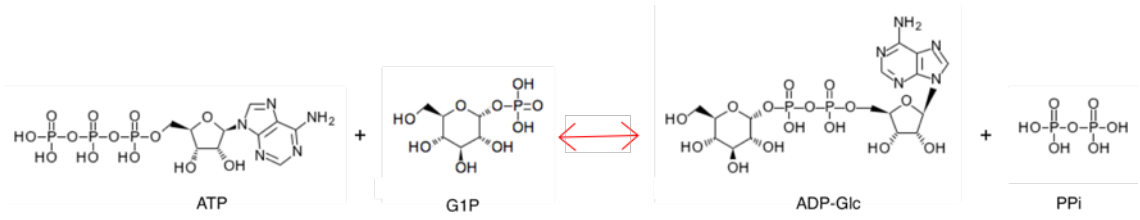


Figure 1.8. Chemical reaction catalyzed by AGPase.

AGPases from different sources have been studied in terms of their subunit composition. Plant AGPases are heterotetramers composed by small (α) and large (β) subunits ($\alpha_2\beta_2$) (Crevillén, Ballicora, Mérida, Preiss, & Romero, 2003; Giroux et al., 1996; Jin, Ballicora, Preiss, & Geiger, 2005) whereas, the native enzymes from most bacterial sources are homotetramers of ~50 kDa subunit (Cupp-Vickery, Igarashi, Perez, Poland, & Meyer, 2008; Haugen, Ishaque, Chatterjee, & Preiss, 1974; Ko, Kim, Lee, & Kim, 1996; Yung & Preiss, 1982). To date, only two crystal structures of AGPase have been reported, the *apo* state of *Agrobacterium tumefaciens* AGPase (*At*AGPase; Figure 1.9; PDB: 3BRK; Cupp-Vickery et al. 2008) and the potato tuber AGPase (*St*AGPase). In the case of *St*AGPase, the crystal structure corresponds to a non-physiological, truncated recombinant homotetrameric version of the small subunit (α4), and it was solved in the presence of (i) SO4 (PDB: 1YP2), (ii) ATP (PDB: 1YP3), and (iii) ADP-Glc (PDB: 1YP3; Jin et al. 2005).

*At*AGPase monomer comprises two different domains: a glycosyltransferase A (GTA)-like domain containing the active site and a C-terminal regulatory domain constituted by a left-handed parallel β-helix (LβH). The GTA-like catalytic domain is composed of a seven-strand central β-sheet with only one β-strand oriented antiparalell in the center of the β-sheet. This feature is surrounded on both sides by 12 amphipathic α-helices, which interact through their hydrophobic faces with the hydrophobic side chains of the central β-sheet (Figure 1.9A). The C-terminal LβH domain comprised four β-strands stacked to form a small left-handed parallel β-helix that displays a triangular coil. This structural feature shows a short β-strand on the top and two small β-strands, mixed with two short α-helices, cover the bottom of the β-helix (Figure 1.9A). The dimerization of *At*AGPase takes place by the antiparalell stacking of the LβH domains whereas the tetramerization appears from the interaction between the N-terminal domains (Figure 1.9B-C; Cupp-Vickery et al. 2008). Despite the low sequence identity shared between *At*AGPase and the *St*AGPase small subunit (31%), their overall fold is preserved and it is predicted for all bacterial and plant AGPases (Figure 1.9D; Cupp-Vickery et al. 2008; Singh, Phillips, and Thorson 2012).

AGPases are members of the superfamily of sugar nucleotidyltransferases (SNTs) and the enzymes within this family share a similar domain organization and common structural features. The GTA-like catalytic domain is composed of an α/β/α sandwich reminiscent of the dinucleotide-binding Rossmann fold (Rao & Rossmann, 1973) with a central β-sheet comprising seven β-strands flanked on both sides with

several α-helices tightly packed against the β-sheet. Often there are additional secondary structure insertions, sometimes known as subdomains, between strands β5 and β6 and strands β2 and β3, depending upon the nucleotidyltransferase family (Singh et al., 2012). On the other hand, the C-terminal auxiliary domain corresponds to a LβH feature but, based on several structures reported within this family, the orientation and length of this β-helix domain can vary along with its function (Brown et al., 1999; Kanamaru et al., 2002; Kostrewa, D'Arcy, Takacs, & Kamber, 2001; Raetz & Roderick, 1995).

Crystal structures of several other sugar nucleotide pyrophosphorylases have been reported, including; (i) the *E. coli* and *Streptococcus pneumoniae* UDP-N-acetylglucosamine pyrophosphorylase GlmU, a bifunctional enzyme that catalyzes the CoA- dependent acetylation of glucosamine-1-phosphate (GlcN-1-PO4) and the metal-dependent condensation of the resulting GlcNAc-1-PO4 with UTP to form the activated nucleotide sugar UDP-GlcNAc involved in the bacterial cell wall synthesis (PDB: 2OI5; Brown et al. 1999; Kostrewa et al. 2001; Green et al. 2012), (ii) the *Pseudomonas aeruginosa* glucose-1-phosphate thymidylyltransferases RmlA, which catalyzes the formation of di-oxythymidine diphosphated-D-glucose (TDP-D-glucose) from deoxythymidine triphosphate (dTTP) and G1P, the first reaction in the L-rhamnose biosynthetic pathway (PDB: 3ZLK; Blankenfeldt et al. 2000), and (iii) the *E. coli* glucose-1-phosphate thymidylyltransferase RffH, that catalyzes the same reaction but forms part of a different operon and biosynthetic pathway, being involved in the synthesis of enterobacterial common antigen (ECA), a cell surface glycolipid found in Gram-negative enteric bacteria (PDB: 1MC3; Sivaraman et al. 2002). The structural information derived from the GlmU, RmlA, and RffH crystal structures and the strong homology shared with the catalytic domain of AGPase, suggest a similar mechanism of action for these enzymes. The catalytic mechanism of AGPases will be further discuss in chapter 4.

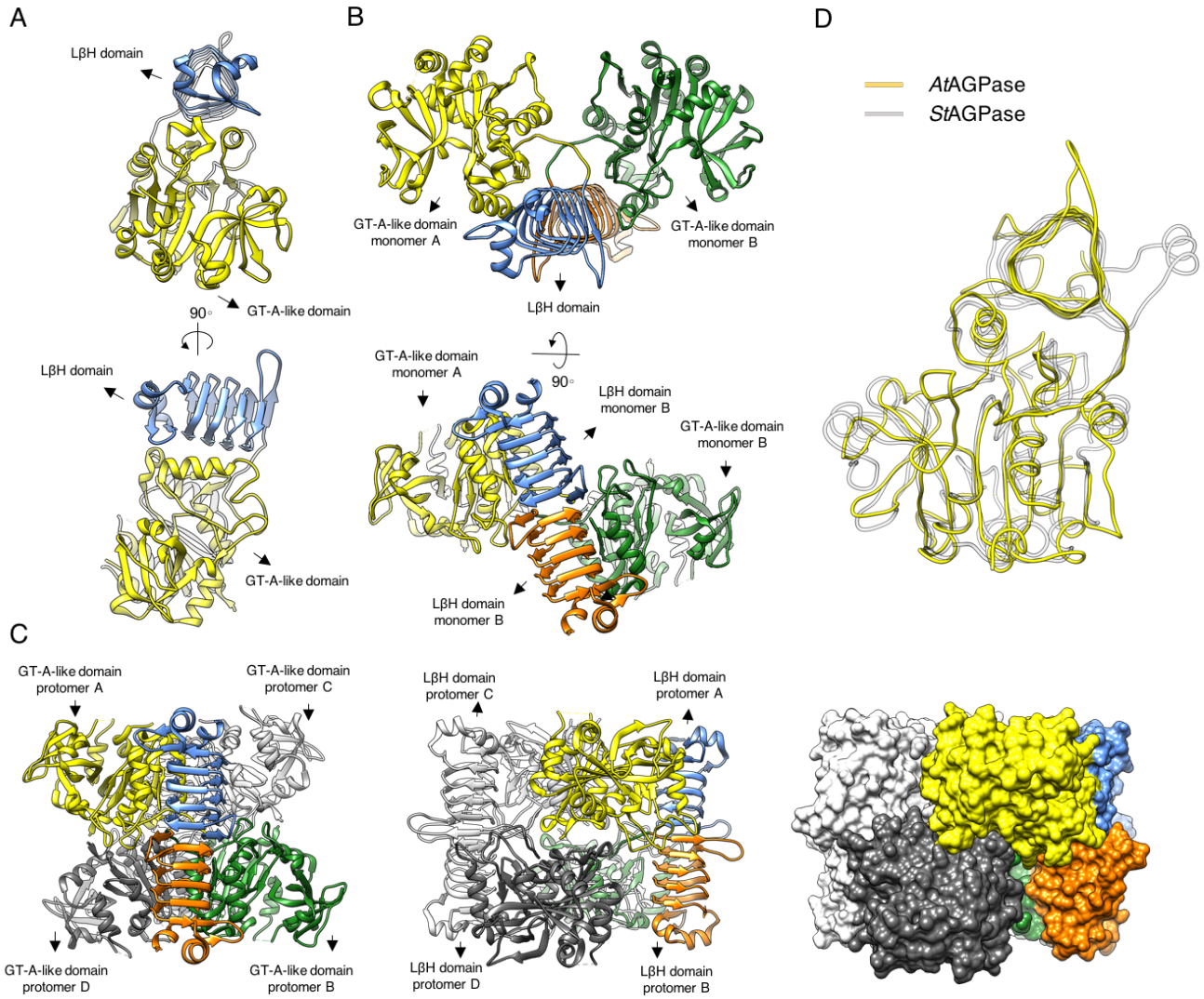


Figure 1.9. The crystal structure of AtAGPase. *(A) The overall structure of an AtAGPase protomer showing the GTA-like domain (yellow) and the LβH domain (blue). (B) Two views of the AtAGPase dimer. (C) The structure of an AtAGPase homotetramer and its surface representation. (D) Superposition of AtAGPase protomer and StAGPase small subunit.*

1.4.2. Glycogen synthase

1.4.2.1. Initiation

It is well known that in the case of yeast and mammals the initiation of glycogen biosynthesis involves the action of the enzyme glycogenin (GN), which is considered the first acceptor of glucose units from the NDP-sugar donor UDP-Glc, catalyzing an autoglycosylation reaction. Further self-glycosylation occurs up to an oligosaccharide chain containing 8–12 glucose units linked by α-1,4-glucosidic linkages. This oligosaccharide remains attached to GN and forms a primer for further elongation by GS and the action of BE. In addition, glycogen and lower molecular weight maltodextrins can also serve as primers for the GS reaction (Lomako, Lomako, & Whelan, 2004; Roach et al., 2012a).

A major unresolved question of bacterial glycogen biosynthesis is the mechanism of initiation. Fully sequenced genomes of bacteria known to accumulate glycogen have failed to reveal the presence of glycogenin homologues (Fox et al., 1973). Initially, it was proposed the existence of an initiation reaction in bacteria based in the incorporation of glucose into a protein acceptor. These studies were based in the different solubility of protein and glycogen in trichloroacetic acid (TCA), which makes possible to differentially determine the incorporation of radio-activity from labelled NDP-sugar into protein or into glycogen (Barengo, Flawia, & Krisman, 1975). However, further experimental evidence demonstrated that this observation corresponded to the elongation of pre-existent contaminant α-1,4-glucans in the sample and not to a glycogen *de novo* synthesis phenomenon (Kawaguchi, Fox, Holmes, Boyer, & Preiss, 1978). More recently, it was observed that *A. tumefaciens* GS can not only elongate α-1,4-linked glucans, but also form the primer required for the elongation process by catalyzing its own glycosylation in a similar way as glycogenin. Interestingly, some differences between glycogenin and GS initiation mechanisms were observed. The oligosaccharides formed by GS were composed by 2-9 glucose residues and, in addition, this glycan is released from the enzyme whereas in the case of glycogenin, it remains covalently attached during the elongation and branching steps (Ugalde et al., 2003). It was proposed that, bacterial GSs use this *de novo* synthesis mechanism in the absence of available soluble α-1,4-glucans to provide itself with an initial substrate. Due to the high affinity of GS for soluble α-1,4-

glucans, when a functional concentration of this substrates is reached within the cell, GS preferentially catalyzes an elongation reaction of this oligosaccharides inducing an apparent inhibition of the initiation reaction (Ugalde et al., 2003).

1.4.2.2. Elongation

The next step in the glycogen biosynthetic pathway is carried out by GS and consists in the elongation of the α-1,4-linkages by the successive addition of α-1,4-linked glucose residues to the non-reducing end of glycogen, using NDP-glucose as the donor substrate (Figure 1.10; J Preiss 2010).

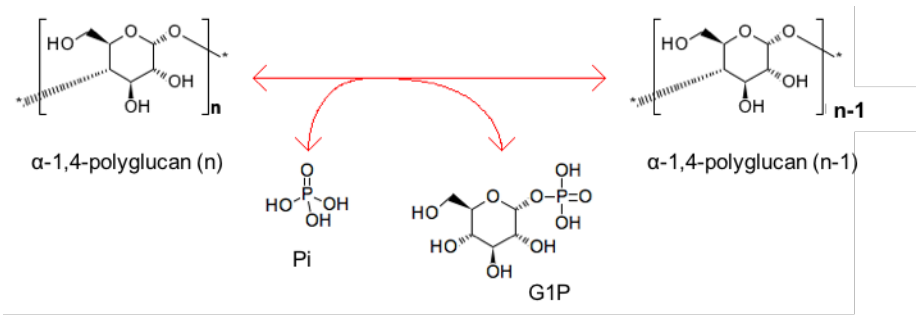


Figure 1.10. Chemical reaction catalyzed by bacterial GS.

Based on amino acid sequence similarity, glycosyltransferases (GTs) have been classified into several families (Cantarel et al., 2009). This classification system, which presently counts over 103 families of GTs is available and continuously updated at the Carbohydrate-Active enzyme database (CAZYy: www.cazy.org). A feature that makes these families particularly useful is that they correlate with several structure-based properties of the enzymes, such as an essentially conserved fold and active site geometry. Glycogen and starch synthases are classified in two large protein families: mammalian/yeast GSs included into the glycosyltransferase family 3 (GT3) and bacterial/plant GSs, which form part of the glycosyltransferase family 5 (GT5; Sheng, Jia, et al. 2009). Mammalian and yeast synthases are ≈80kDa enzymes that use UDP-glucose as sugar donor and are regulated by covalent phosphorylation and allosteric ligand binding, whereas bacterial and plant synthases have a smaller size (≈50 kDa), prefer ADP-glucose as substrate and appear to be unregulated proteins (S. G. Ball &

Morell, 2003; Gibbons, Roach, & Hurley, 2002). Sequence similarity is marginal between the ADP-glucose- and UDP-glucose-requiring enzymes, on the other hand, bacterial GSs and plant SSs share significant (30–34%) sequence homology over most of their length, consistent with a similar overall fold and enzymatic mechanism (Coutinho, Deleury, Davies, & Henrissat, 2003).

Two major structural folds have been described for the nucleotide sugar-dependent (Leloir) GTs defined as GT-A and GT-B. GSs and SSs are members of the GT-B superfamily of glycosyltransferases, which shares a common double 'Rossmann-fold' domain architecture. This 'nucleotide binding' or 'Rossmann-fold' domain contains two sets of β-α-β-α-β units (654123 topology), together forming a single parallel sheet flanked by α-helices. An extended interdomain linker peptide connects the N- and C-terminal halves describing a typical feature of GT-B enzymes (Albesa-Jové & Guerin, 2016; Urresti et al., 2012). The active site of all of these enzymes lies in the interdomain cleft and therefore, a proper relative orientation of the two domains is required for correct active site geometry and catalytic competence as demonstrated for MurG (Hu et al., 2003), PimA (Guerin et al., 2007, 2009) or MshA (Vetting, Frantom, & Blanchard, 2008). It is generally accepted that in GT-B enzymes, the nucleotide-sugar donors mainly bind to the C-terminal domain of the protein, whereas the N-terminal domain is involved in acceptor substrate recognition. Since acceptors exhibit a marked diversity of chemical structures compared with nucleotide-sugar donors, the N-terminal domains reflect this variability by showing different rearrangements of secondary structural elements (Breton, Šnajdrová, Jeanneau, Koča, & Imberty, 2006).

To date, there are two crystal structures reported of bacterial GSs, the one from *A. tumefaciens* (*At*GS; Buschiazzo et al. 2004) and from *E. coli* (*Ec*GS; Sheng, Jia, et al. 2009; Sheng, Yep, et al. 2009). Both enzymes crystalized as monomers describing the typical GT-B fold mentioned above (Figure 1.11A). Besides, comparison of *At*GS and *Ec*GS crystal structures in complex with ADP (PDB: 1RZU, 3GUH) shows that the nucleotide binds in the interdomain cleft, mostly along the C-terminal domain wall, with the adenine ring stacked against the conserved Tyr355 (equivalent to *At*GS Tyr354; Phe in animal GS and plant SS). The interaction with the N-terminal domain is restricted to Asp21, Lys15 and Gly18 all of which are conserved in both glycogen and starch synthases (Buschiazzo et al., 2004). These last residues, Lys15 and Gly18, form part of the Lys-Tyr-Gly-Gly-Leu (KTGGL) widely conserved motif loop among GSs and SSs.

This motif is located facing the nucleotide-binding pocket and is suggested to be involved in substrate binding and catalysis (Figure 1.11C; Furukawa et al. 1993; Furukawa et al. 1994). The last three residues of this motif are highly conserved in many other GT-B enzymes, including glycogen phosphorylase (GP, family GT35), maltodextrin phosphorylase (MalP, family GT35), and trehalose-6-phosphate synthase (OtsA, family GT20) (Buschiazzo et al., 2004).

Of particular interest is the *Ec*GS active closed crystal structure of the catalytic inactive mutant Glu-377-Ala containing bound ADP, glucose and the acceptor analog HEPPSO ([N-(2-hydroxyethyl) piperazine-(2-hydroxypropanesulfonic acid)]) which results in a 15.2° domain-domain closure (PDB: 3COP; Figure 1.11B). This domain motion is also observed in the crystal structure of the *wild-type* enzyme bound to ADP and glucose (PDB: 2QZS), the one in complex with ADP and HEPPSO (PDB: 3GUH) and the Glu-377-Ala mutant in complex with ADP, glucose and maltohexaose (PDB: 3CX4; Sheng, Jia, et al. 2009; Sheng, Yep, et al. 2009). Interestingly, the *At*GS crystal structure in complex with ADP (PDB: 1RZU; Buschiazzo et al. 2004) displays an open conformation, supporting the idea of the glycan acceptor-binding requirement to achieve the competent closed state. This domain closure brings the N-terminal loop containing the KTGGL motif into the vicinity of the C-terminal domain to interact with both, the substrate ADP-Glc and the acceptor molecule (Figure 1.11C).

GTs and glycoside hydrolyzes can be classified as either retaining or inverting enzymes depending on whether the anomeric configuration of the sugar donor is the same (retaining) or different (inverting) from that of the product. The reaction mechanism of inverting GTs follows a single displacement mechanism with an oxocarbenium ion-like transition state and an asynchronous S_N2 mechanism, analogous to that observed for inverting glycoside hydrolases (Lairson, Henrissat, Davies, & Withers, 2008). In contrast, the catalytic mechanism for retaining GTs is currently a matter of strong debate. It was proposed that some retaining enzymes would use a double displacement mechanism, where the attack of an enzyme nucleophile results in the formation of a covalent intermediate (Gómez, Lluch, & Masgrau, 2013; Rojas-Cervellera, Ardèvol, Boero, Planas, & Rovira, 2013). However, in the absence of a clear residue near the reaction center that could act as a nucleophile to form the glycosyl-enzyme intermediate in the majority of GTs, an alternative mechanism known as the S_Ni 'internal return', also called the S_Ni-like mechanism, has been suggested. In this mechanism, leaving group

departure and nucleophilic attack occur on the same face of the sugar (Lee et al., 2011) and involve either a short-lived oxocarbenium ion intermediate (S_Ni-like; Lairson et al. 2008; Sinnott and Jencks 1980; Persson et al. 2001; Gibson et al. 2002) or an oxocarbenium ion transition state (S_Ni; Gómez, Lluch, and Masgrau 2013; Bobovská, Tvaroška, and Kóňa 2015; Ardèvol and Rovira 2011; Lira-Navarrete et al. 2014).

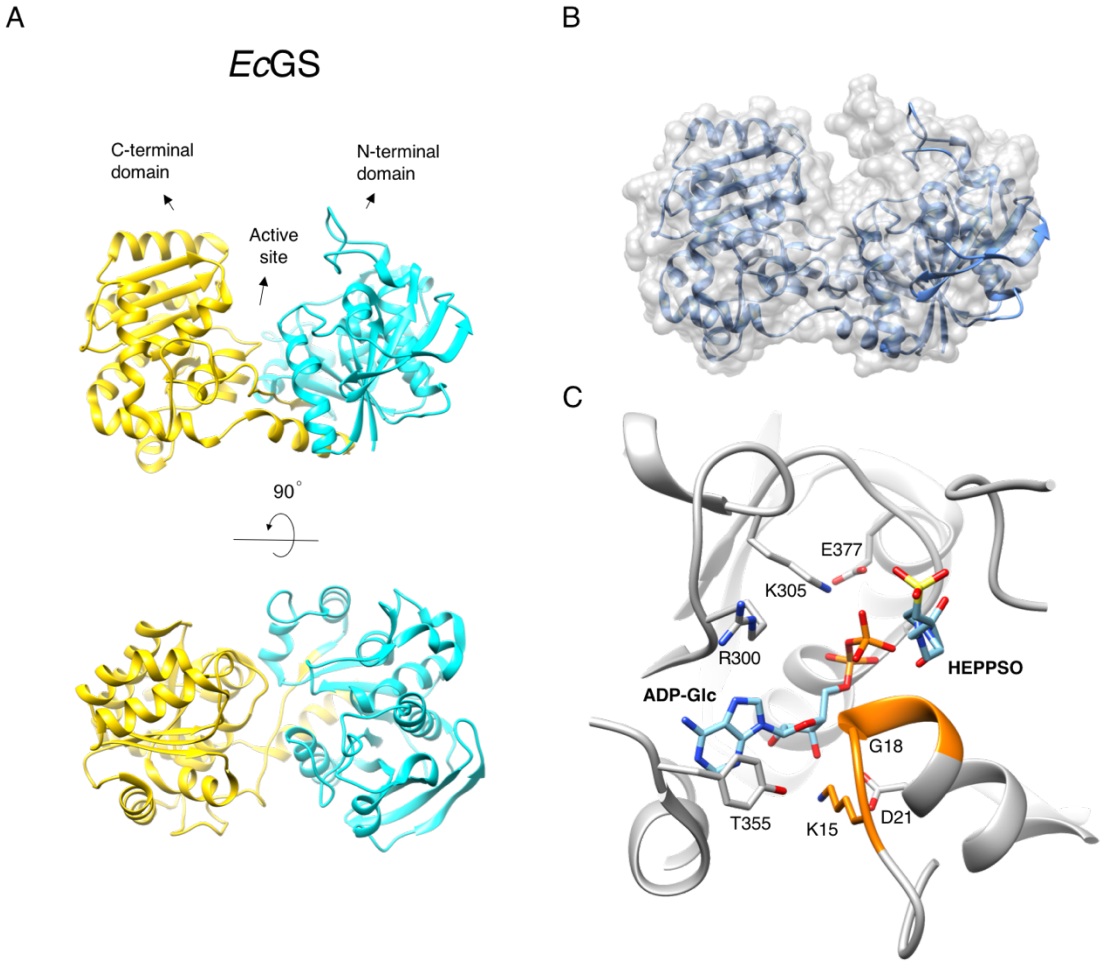


Figure 1.11. *(A) EcGS overall fold (PDB: 3D1J). (B) Structural comparison between the 'open' (schematic cartoon) and 'closed' (molecular surface representation) states of EcGS. (C) Close up view of EcGS active site in complex with ADP-Glc and HEPPSO (PDB: 3GUH) sowing only selected interactions. The conserved KTGGL motif is represented in orange.*

Based on its catalytic mechanism, GSs are retaining GTs. Glu377 (*Ec*GS numbering; equivalent to *At*GS Glu376) and its equivalents were thought/proposed to be either the catalytic nucleophile or the general acid/base catalyst for GS. It is the first residue in the E-X7-E motif of some retaining GTs, such as eukaryotic glycogen synthases (GT3), α-glucosyltransferase (GT4), starch synthases, and bacterial glycogen synthases (GT5) and its mutation per alanine result in a drastic reduction in the enzyme

activity (Yep, Ballicora, and Preiss 2004; Cid et al. 2002). However, further analysis based on the structural information available, support a different role; (i) it provides charge stabilization to Lys305, which in turn stabilizes the negative charge on the leaving group ADP among with Arg300, and (ii) it is critical in locating the acceptor glucose moiety (Sheng, Jia, et al., 2009). A similar situation can be observed for the equivalent acidic residue in MalP, OtsA (Asp), and GP where the side chain carboxyl group also makes a direct interaction with the glucose moiety and a lysine residue within the active site of these retaining GT-Bs (Gibson et al., 2002; Martinez-Fleites et al., 2006; Sprang, Goldsmith, & Fletterick, 1987; Watson et al., 1999; Watson, Schinzel, Palm, & Johnson, 1997). Although the Glu-377 side chain is not far from the β-phosphate (3.9 Å in *Ec*GS), positively charged Arg-300 and Lys-305 directly interact with the substrate phosphate group, and are therefore more likely to play the role of proton transfer agent (Figure 1.11C) (Sheng, Yep, et al., 2009).

1.4.3. Branching enzyme

The BE (EC 2.4.1.18) catalyzes the formation of α-1,6-glucosidic branch points into glycogen/amylopectin polysaccharides (Figure 1.12). This branching increases the number of non-reducing ends, making glycogen more reactive to synthesis and digestion, and is also essential for assuring its solubility in the cell. The frequency and position of the branch points are important determinants for the structure and properties of starch in plants and of glycogen in animals and bacteria (J Preiss, 2010).

The reaction catalyzed by BE proceeds in two steps: first, a preexisting α-1,4-glucan chain is cleaved and the non-reducing portion of the donor chain is covalently attached to the carboxyl group of the catalytic residue at the active site of the enzyme (Uitdehaag et al., 1999). In a second step, the glucan moiety is then transferred to the C-6 hydroxyl group of the same or another glucan chain (the acceptor chain). Through this double displacement mechanism, the α-configuration of anomeric carbon is retained, resulting in the formation of α-1,6-linkage. If a water molecule instead of the C-6 hydroxyl group of a sugar is involved in the latter step of the reaction, the enzyme serves as a hydrolase rather than a transferase. Therefore, BE is mechanistically similar to amylolytic enzymes and, in addition, they share some conserved sequence regions (CSRs), thus

being grouped into the same family (although only a limited number of BEs show hydrolytic activity; MacGregor, Janecek, and Svensson 2001).

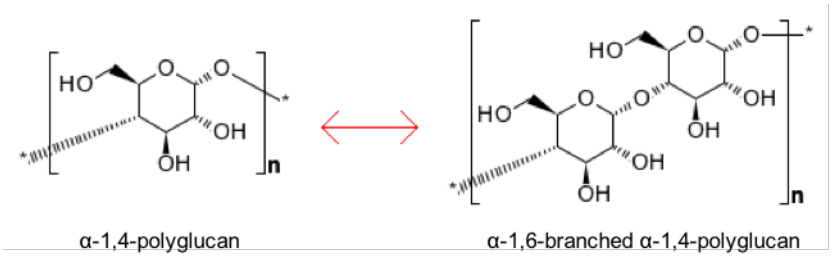


Figure 1.12. Chemical reaction catalyzed by BE.

Almost all sequence-annotated as BEs, belong to the GH13 family of glycoside hydrolases and fall either into subfamily 8 (eukaryotic BEs) or subfamily 9 (prokaryotic BEs). BEs in animals, fungi, and plants show higher sequence similarity to each other than bacterial BEs. The GH13 family is the largest glucoside hydrolase family, and comprises α-amylases, pullulanase/isoamylase, cyclodextrin glucanotransferase (CGT), and branching enzymes (Abad et al., 2002; Pal et al., 2010; Sean Froese et al., 2015; Suzuki & Suzuki, 2016). GH13 BEs are multidomain enzymes consisting on: (i) a carbohydrate-binding module 48 (CBM48), (ii) a central $(\beta/\alpha)_8$-barrel catalytic domain (also known as TIM barrel) consisting of eight parallel β-strands surrounded by eight parallel α-helices and (iii) the C-terminal domain. In some GH13_9 BEs, CBM48 is further preceded by an N domain, which is generally absent from GH13_8 BEs. Indeed, GH13_9 BEs are divided into group 1 and group 2, depending on the presence or absence, respectively, of this N domain.

The crystal structure of GH13_9 BE has been elucidated from *E. coli* (*Ec*BE; PDB: 3K1D; Abad et al. 2002) and *Mycobacterium tuberculosis* (*Mtb*BE; PDB: 1M7X; Feng et al. 2015) enzymes, both in their unliganded form (Figure 1.13). In the case of *Ec*BE, it has also been reported the crystal structure in complex with maltoheptaose (PDB: 4LPC) and maltohexaose (PDB: 4LQ1; Feng et al. 2015). Based on the reported structures, the N domain, the CBM48, and the C domain consist of approximately 100 amino acid residues each, and all adopting a β-sandwich type fold (Abad et al., 2002; Noguchi et al., 2011; Pal et al., 2010; Palomo et al., 2011; Sean Froese et al., 2015). The CBM48 domain is related to the substrate specificity, recognition, and binding whereas the N

domain seems to be involved in determining the size of the chain transferred and thus the branching pattern. On the other hand, the C domain is suggested to participate in substrate preference and catalytic capacity (Tetlow & Emes, 2014).

All members of the α-amylase family display an overall negatively charged surface which plays a crucial role in substrate binding. The crystal structures of *Ec*BE in complex with maltoheptaose and maltohexaose, respectively, show several binding sites located a the surface. Interestingly, there is no evidence for oligosaccharide binding in the active site of *Ec*BE. These binding sites can be regarded as surface/secondary binding sites (SBSs) in which oligosaccharides are bound at a certain distance of the active site. SBSs have been described for a wide range of glycoside hydrolases, and are particularly prominent in GH13 family (Cockburn et al., 2014). The feature that distinguishes BE from all other GH13 enzymes is that its reactivity is only seen with large polymers such as glycogen or amylose and not short glucan chain. The presence of multiple SBSs could be a way to anchor a complex glycogen granule and determine the chain length specificity for the branching reaction as a 'molecular ruler´. This agrees with the emerging concept of glycogen serving not only as the substrate and product of its metabolism but also as a scaffold for all acting enzymes (Feng et al., 2015).

Structural data on the GH13 family has defined amino acid residues along the central $(\beta/\alpha)_8$-barrel catalytic domain whose positions are well conserved in all α-amylases. Based on *Ec*BE numbering these residues are: Tyr300, Asp335, His340, Arg403, Asp405, Glu458, His525, and Asp526. Biochemical studies support the role of these residues in α-amylases family substrate binding and catalysis, specifically, Asp405 serves as a nucleophile in the reaction, whereas Glu458 is responsible for the protonation and deprotonation necessary on the leaving group and attacking oxygen (Tao, Reilly, & Robyt, 1989; Uitdehaag et al., 1999). In the case of *Ec*BE and *Mtb*BE the catalytic residues aspartate and glutamic acid are located in adjacent loops (Abad et al., 2002; Pal et al., 2010).

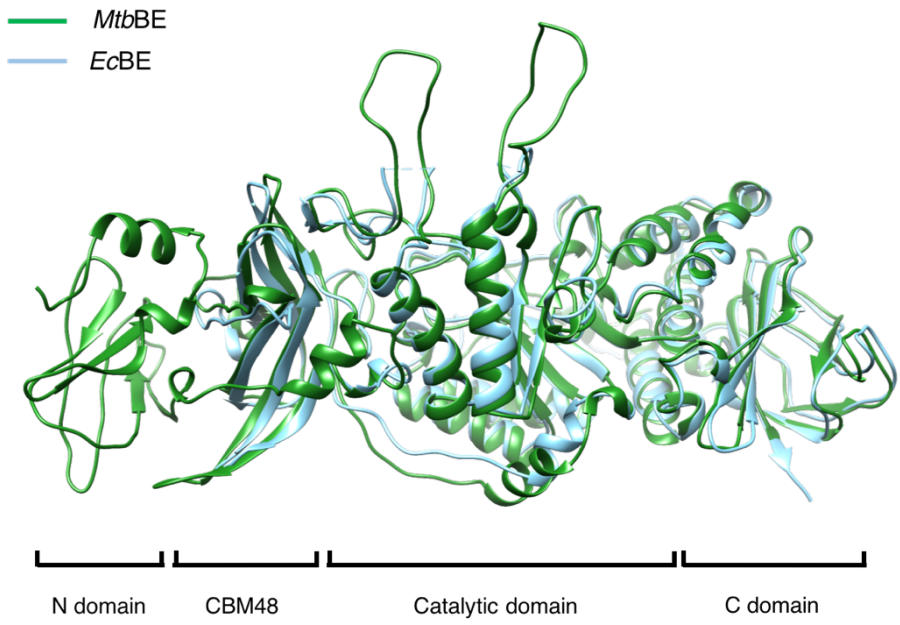


Figure 1.13. Superposition of EcBE (blue) and MtbBE (green) crystal structures highlighting the conserved domain architecture.

1.5. Molecular bases of bacterial glycogen degradation

1.5.1. Glycogen phosphorylase

Phosphorylases are key enzymes in the carbohydrate metabolism and are found in many organisms, including bacteria, yeast, plant and vertebrates. Glycogen phosphorylase (GP; EC 2.4.1.1) is encoded in the *E. coli* glycogen gene cluster *glg*BX-*glg*CAP being the expression product of the *glg*P gene and, along with the debranching enzyme (*glg*X), participates in the degradation of glycogen during extended periods of substrate deprivation. Specifically, GP catalyzes the cleavage of α-1,4-linkages at the non-reducing ends of glycogen and branched polysaccharides with retention of configuration at the anomeric carbon, releasing G1P (Figure 1.14). The cofactor pyridoxal-5-phosphate (PLP) is absolutely required for GP activity being the PLP-domain common to all α-glucan phosphorylases. This cofactor is covalently bound via a Schiff base to an active site lysine and catalysis is dependent on the ionizable phosphate residue of the enzyme-bound PLP (Kirby, 2007). This enzyme is unable to bypass or hydrolyze the α-

1,6 linkages and therefore stops two, three, or four residues from the first α-1,6 branch encountered to generate the so-called phosphorylase-limit dextrin (Alonso-Casajús et al., 2006; J. T. Park et al., 2011; Jack Preiss, 2006).

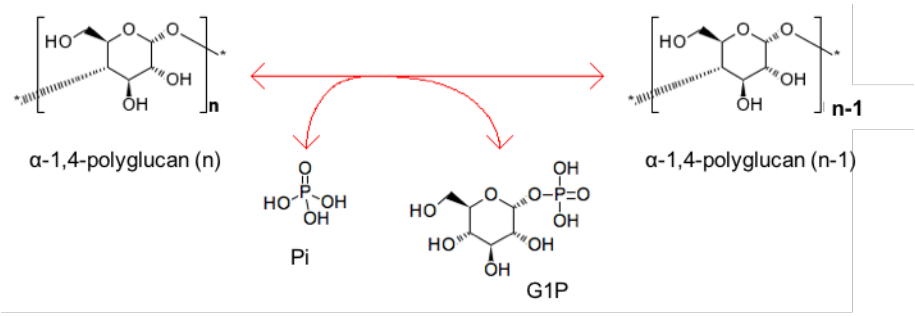


Figure 1.14. Chemical reaction catalyzed by GP.

It is well stablished that mammalian GPs are subject to regulation by two mechanisms, allosteric regulation and covalent activation through phosphorylation. In contrast, bacterial phosphorylases show no apparent regulatory properties and are thought to be exclusively controlled at gene expression level (Buchbinder, Rath, & Fletterick, 2001; Newgard & Hwang, 1989; Wilson et al., 2010). In this regard, *E. coli* GP (*Ec*GP) seems not to be allosterically regulated with respect to substrates of physiological relevance, such as AMP, ADP-Glc or glucose; however, GlgP could be allosterically modulated in *E. coli* as a result of its interaction with the histidine phosphocarrier protein (HPr). The HPr protein is an essential element in sugar transport by the bacterial phosphoenolpyruvate:sugar phosphotransferase system and it has been reported a highly specific binding with *Ec*GP. Specifically, the phosphorylated form of HPr binds with higher affinity to GlgP than the dephosphorylated form and equilibrium ultracentrifugation analysis indicated that HPr allosterically regulates the oligomeric state of glycogen phosphorylase (Alonso-Casajús et al., 2006; Seok et al., 1997; Wilson et al., 2010).

All known phosphorylases share catalytic and structural properties being included into the GT35 family. Members of this family belong to the GT-B structural superfamily of GTs (Albesa-Jové, Giganti, Jackson, Alzari, & Guerin, 2014). Only one bacterial structure has been reported within the GT35 family, the *E. coli* maltodextrin phosphorylase (MalP). MalP catalyzes the phosphorolysis of maltodextrins, but is poorly

active against the more complex, branched glycogen molecule. It is a non-allosteric, dimeric enzyme shearing a 45% sequence identity with *Ec*GP (Newgard & Hwang, 1989; Schinzel & Nidetzky, 1999). As expected from sequence comparison the overall structure of MalP shows a fold similar to that of GP from rabbit muscle, the phosphorylase study more in deep (Oikonomakos, Chrysina, Kosmopoulou, & Leonidas, 2003). The bacterial enzyme is an α/β protein that exhibits a two-domain fold, the N-terminal domain (residues 1–447) and the C- terminal domain (residues 448–797) separated by the catalytic site cleft. The dimeric structure is stabilized by inter-monomer contacts involving the central region known as the tower helices which mediates the allosteric transition between the active state (R-state) and less active state (L-state) in mammalian GP enzyme (Figure 1.15A). The active site lies between the N- and C-terminal domains, at the center of the molecule, and is accessible to the solvent via a channel ~20 Å in length that forms the substrate-oligosaccharide binding site (Figure 1.15B; O'Reilly, Watson, & Johnson, 1999; Oikonomakos et al., 2003).

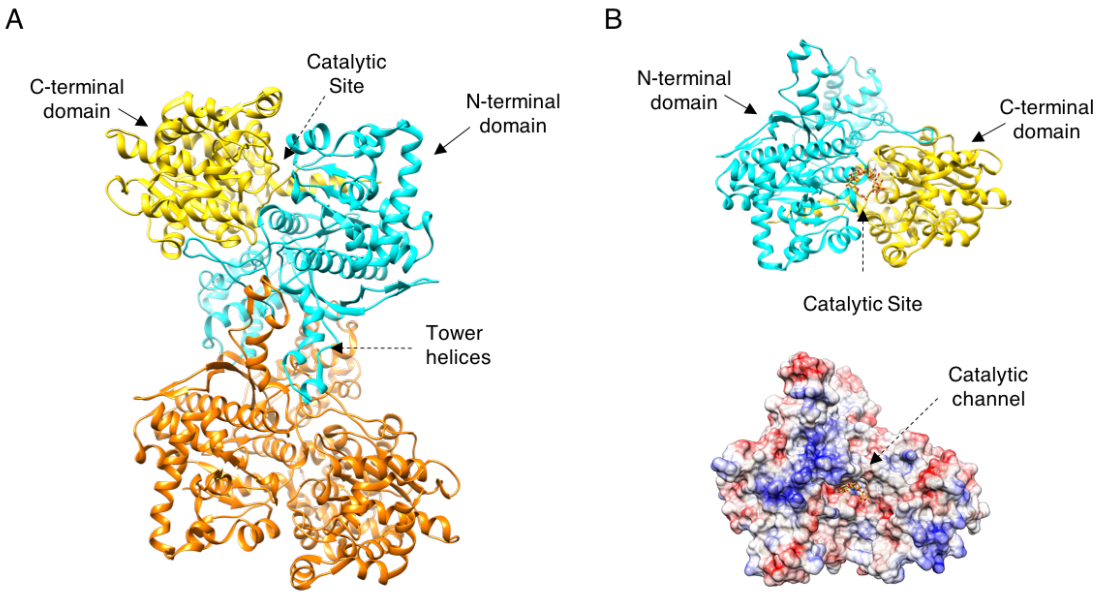


Figure 1.15. *(A) Dimeric MalP overall structure indicating in one of the monomers the N-terminal domain (cyan) and C-terminal domain (yellow) (B) MalP monomer structure with the oligosaccharide located in the active site (PDB: 1E4O). The electrostatic surface representation of the monomer shows the entrance to the catalytic cleft.*

Several structures of MalP have been reported and interestingly, the one in complex with maltose provided the first phosphorylase structure featuring an

oligosaccharide bound at the catalytic site (PDB: 1AHP; O'Reilly et al. 1997). Comparison with the structure of the enzyme alone indicates that an induced-fit mechanism closes the N- and C-terminal domains over the oligosaccharide substrate. This phenomenon is caused mostly by the movement of the 380s loop (residues His377-Trp387) which leads to the closure of the catalytic channel. (Figure 1.16A; O'Reilly et al. 1997; Watson et al. 1997; Buchbinder, Rath, and Fletterick 2001). The crystal structures in complex with maltotetraose (G4; PDB: 1L5W), maltopentose (G5; PDB: 1E4O) and a non-hydrolysable thiopentasaccharide (GSG4; PDB: 1QM5) showed similar conformations of the 380s loop to the maltose complex (Watson et al., 1999). In addition, it was reported that the binding of G1P in sub-site -1 and a tris-(hydroxymethyl)aminomethane (Tris) molecule coming from the crystallization buffer in sub-site +1, promotes the closure of the loop (Figure 1.16B; PDB: 1L5V; Geremia et al. 2002). On the other and, the structure in complex with acarbose, an unnatural cyclitol that resembles a tetrasaccharide, was close to that of the native structure indicating that the shift of the 380s loop is likely to be a feature in substrate recognition (PDB: 2ECP; O'Reilly, Watson, and Johnson 1999).

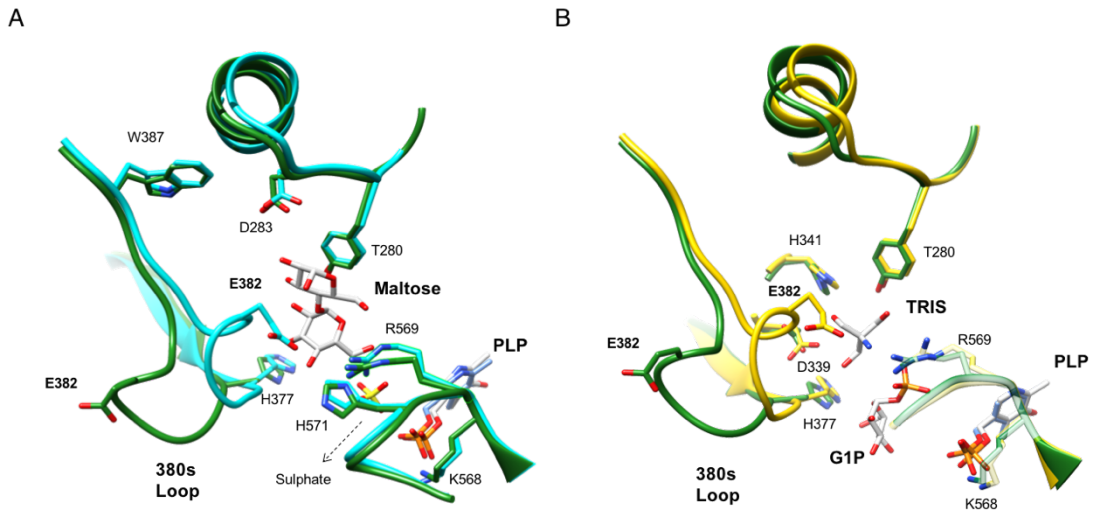


Figure 1.16. *(A) Comparison between the MalP-acarbose complex (green; PDB: 2ECP) and MalP-maltose complex (cyan; PDB: 1AHP) showing the 380s loop shift and selected residues at the binding site. In the MalP-maltose complex a sulphate anion occupies the inorganic phosphate site. (B) Comparison between the MalP-acarbose complex (green; PDB: 2ECP) and MalP-G1P-Tris complex (yellow; PDB: 1L5V) showing the 380s loop shift and selected residues at the binding site. Neither of the representations show the acarbose.*

As the catalytic domain is highly conserved among phosphorylases a common catalytic mechanism has been suggested (Schinzel & Nidetzky, 1999). Based on the MalP–GSG4-P ternary complex, it has been proposed that the α-retaining mechanism proceeds via a double displacement reaction with two steps (Figure 1.17); (i) the 5'-phosphate of the cofactor PLP promotes general acid attack by the inorganic phosphate on the glycosidic oxygen. This oxygen is protonated leading to the cleavage of the C1-O bond and formation of the oxonium-carbonium ion transition state, that is favored and stabilized by the now negatively charged phosphate; (ii) The inorganic phosphate acting as a nucleophile, attacks the carbonium oligosaccharide phosphorylase ion C1 carbon leading to formation of the product G1P with retention of configuration (Campagnolo, Campa, Zorzi, Wuerges, & Geremia, 2008; Geremia et al., 2002; Watson et al., 1999).

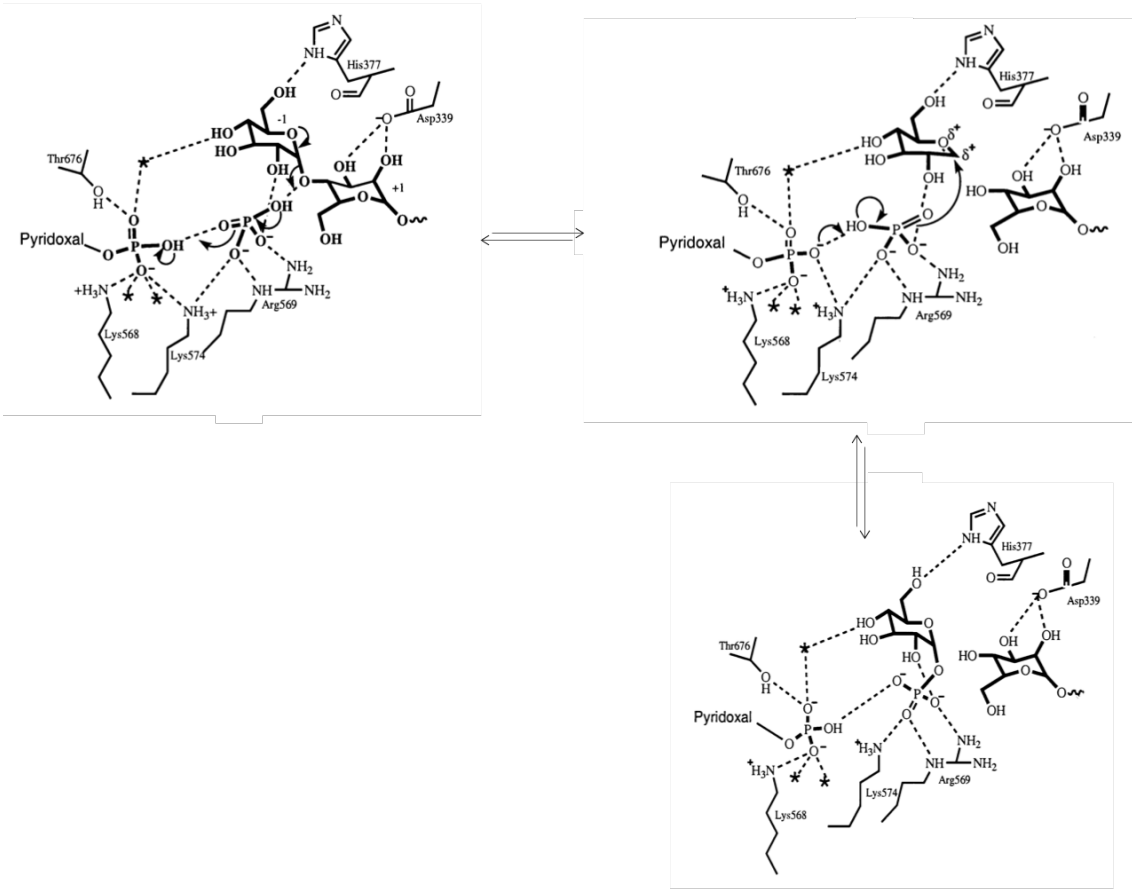


Figure 1.17. Catalytic mechanism proposed for MalP. *The oligosaccharide substrate binds in subsites -1 and +1 and subsequent sites across the catalytic cleft. His377, the first residue in the mobile region 380s loop, forms a hydrogen bound with O6 hydroxyl group of the sugar in sub-site -1 and Asp339 side chain interacts with O2/O3 hydroxyl groups of the sugar in sub-site +1. The phosphate substrate is stabilized by contacts with Arg569 and Lys574 whereas the 5´-phospate*

group of the PLP cofactor is stabilized by interactions with water molecules (show as asterisks), Lys568 and Lys574. Note that not all of the substrates contacts are shown. the 5´-phospate group of the PLP cofactor promotes a general acid attack by the inorganic phosphate on the glycosidic oxygen leading to the cleavage of the C1-O bound and formation of the oxonium-carbonium ion intermediate. (B) The sugar in sub-site -1 is polarized by the interaction with the inorganic phosphate. (C) Nucleophilic attack by the inorganic phosphate at the C1 position of the carbonium ion leads to the formation of G1P. The position of G1P was inferred from the complex of GP with heptulose-2-phosphate (PDB: 6GPB). (Modified from Watson et al. 1999).

1.5.2. Glycogen debranching enzyme

Along with GP, glycogen debranching enzyme (GDE) constitutes the core of glycogen catabolism, being both codified within *E. coli* glycogen gene cluster *glg*BX-*glg*CAP (GDE; *glg*X; EC 3.2.1.196). As mentioned above, GP catalyzes the depolymerization of the α-1,4-glucosidic linkages, however, due to the high degree of branching present in the glycogen molecule, a second type of enzyme is required to cleave the α-1,6-glucosidic bonds that remain uncleaved by GP (Figure 1.18; Jack Preiss 2006; Iglesias and Preiss 1992).

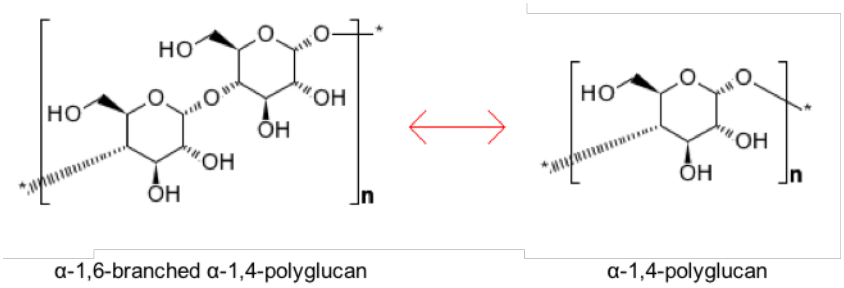


Figure 1.18. Chemical reaction catalyzed by GDE.

Sequence analyses suggest all GDEs are glucan hydrolases, probably belonging to the glycoside hydrolase GH13 family in the CAZy database; however, important functional differences arise between eukaryotic and bacterial GDEs. Eukaryotic animal and yeast GDEs are bifunctional enzymes, which have α-1,4-transferase (N-terminus)

and α-1,6-glucosidase (C-terminus) activity, in contrast, most bacteria and plant debranching enzymes contains only α-1,6-glucosidase activity (Nakayama, Yamamoto, & Tabata, 2001; Woo et al., 2008). AS GP is not able to cleave the α-1,6-linkages, it generates phosphorylase-limit dextrins that consist of maltotriose or maltotetraose chains of three to four glucose residues branching from the linear glycogen strand. GDE in *E. coli* functions by cleaving the α-1,6-glucosidic linkage between these glucose subunits and the linear glycogen chain and rely on MalP and GP to cleave the remaining α-1,4-bonds in this short-released chain. In contrast, animal and yeast GDEs transfer the maltotriose unit from the branch to the nearby linear glycogen strand using α-1,4-transferase activity and then cleave the α-1,6-glucosidic bond that connects the remaining glucose residue to the glycogen polymer (Figure 1.19; Dauvillée et al., 2005; Song et al., 2010).

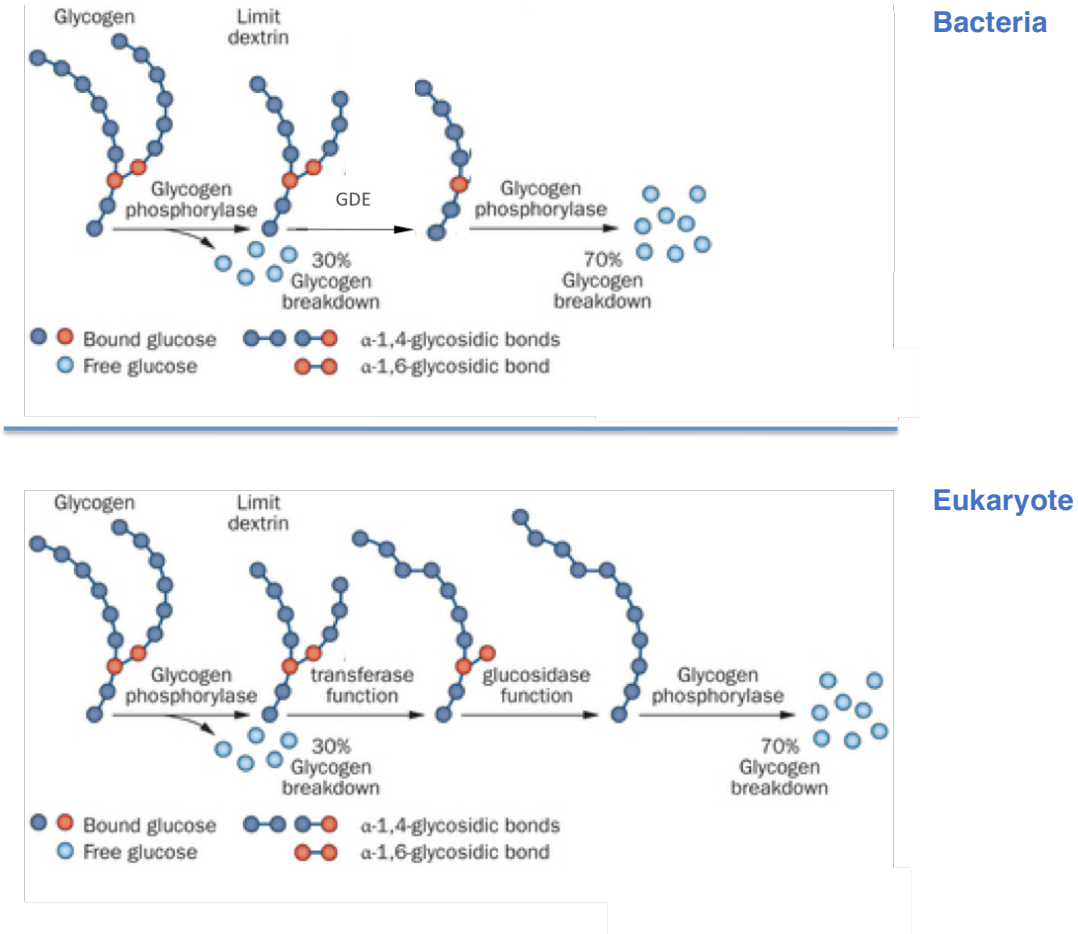


***Figure 1.19. Schematic representation of the functional differences between eukaryotic and bacterial GDE (Modified from** Lew, Guin, and Theodorescu 2015).*

Several crystal structures of GDEs have been reported, one bacterial enzyme corresponding to the *E. coli* GDE (GlgX; PDB: 2WSK; Song et al. 2010), the archeal *Sulfolobus solfataricus* tetrameric GDE (TreX; PDB: 2VNC; Woo et al. 2008), and the yeast *Candida glabrata* GDE (PDB: 5D06; (Zhai et al., 2016). The crystal structure of the monomeric GlgX shows an overall fold comprised by three major domains: (i) N-terminal CBM48 domain, (ii) a central catalytic domain and (iii) a C-terminal domain (Figure 1.20A). Both the N- and C-terminal domains are composed by several β-strands forming a β-sandwich whereas, the central domain describes the GH13 family characteristic $(\beta/\alpha)_8$-barrel catalytic motif. GlgX is a retaining enzyme whose products maintain the stereochemical orientation of their substrates. The catalytic residues Asp336, Glu371 and Asp443, within the central catalytic domain, act respectively as the nucleophile, proton donor and transition state stabilizer. (Song et al., 2010). The active site of GlgX shows an obvious cleft on top of the catalytic barrel motif with an approximate length of 26 A˚ and width of 9 A˚. The residues composing the floor and sides of the groove show good alignment with residues involved in substrate binding in other members of the GH13 family, in fact, this groove overlaps well with the one observed in TreX crystal structure, except for the much shorter cavity dimension in GlgX (Figure 1.20B; Woo et al., 2008). This structural discrepancy between TreX and GlgX correlates with the different substrate specificity among these two enzymes; GlgX does not efficiently cleave substrates longer than four glucose residues, whereas TreX exhibits higher activity for branched substrates longer than 5 glucose residues (Dauvillée et al., 2005; J.-T. Park et al., 2008). At one side of the groove, a region stands out from the bottom of the substrate channel forming helix α4 which presents residues involved in substrate binding (Figure 1.20A). This helix is only observed in GlgX and TreX crystal structures while it is missing in homologs within the GH13 family as isoamylases and pullulanases, suggesting that it could be a useful structural feature for distinguishing GDEs (Song et al., 2010).

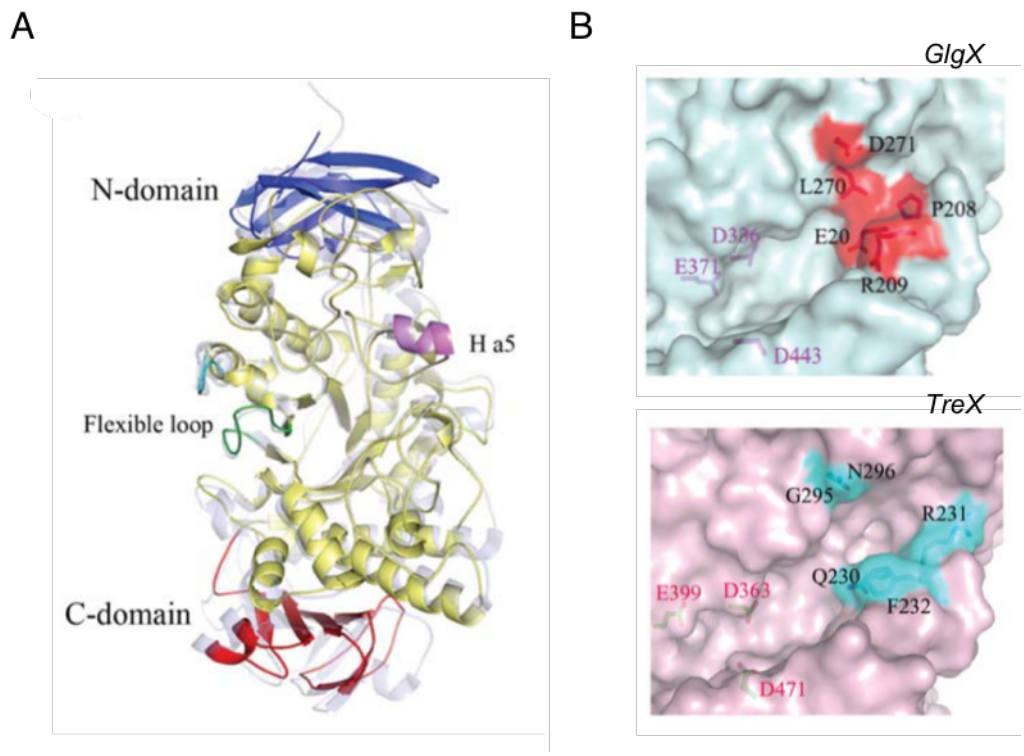


Figure 1.20. Structural comparison between GlgX and TreX. *(A) Schematic overview of GlgX superimposed on the TreX structure. N- and C- terminal domains are in blue and red, respectively. The catalytic domain of GlgX is in yellow whereas the TreX structure is in graywhite. The flexible loop above the active site containing the catalytic residue Glu371 is shown in green and the conserved helix 4 in violet. (B) GlgX and TreX substrate binding grooves showing the catalytic residues (violet). Residues that constitutes the groove barrier in GlgX are shown in red (From Song et al. 2010).*

The high structural similarity between GlgX and TreX is of interest because this archaeal GDE exhibits both the transferase and the glucosidase activity in the same catalytic cleft. TreX exists in two oligomeric states in solution, as dimer and tetramer and exhibits two different active-site configurations depending on its oligomeric state. A flexible loop above the active site containing the catalytic acid/base Glu371 in the GlgX crystal structure displays the same conformation as that observed in TreX dimer structure, however, in the tetrameric form of TreX this loop is flipped to reveal a significantly different conformation (H.-S. Park et al., 2007; J.-T. Park et al., 2008; Song et al., 2010; Woo et al., 2008). The configuration of this structural lid above the substrate binding site in the tetrameric TreX is a common motif that has been found in enzymes with α-1,4-transferase activity. It has been suggested that the inability of catalyzing transglycosylation reaction in GlgX may be related to the absence of this structural motif that could be necessary to stabilize the acceptor molecule during the transglycosylation catalysis (Song et al., 2010).

1.6. Bacterial Glycogen biosynthesis regulation

As mentioned in previous sections, plant starch and bacterial glycogen biosynthetic pathways share the initial step catalyzed by the enzyme AGPase that generates the NDP-sugar donor, ADP-Glc. The formation of ADP-Glc is consider the main regulatory step in glycogen and starch production in these organisms, and requires the investment of ATP resulting in an energetically expensive reaction. For this reason, evolution has led AGPase to acquire exquisite regulatory mechanisms based in a cooperative and allosteric behavior to control this key rate-limiting step (J Preiss, 1978). Almost all AGPases characterized so far are allosterically regulated by small effector molecules that bind to the enzyme in positions that differ from the catalytic site affecting its enzymatic activity. AGPase positive and negative allosteric regulators correspond to metabolites that can differ according to the source but they share the characteristic of being intermediates in the major carbon assimilation pathways within the organisms. AGPase activators are metabolites that represent signals of high carbon and energy content of a particular bacteria or tissue, while inhibitors of the enzyme indicate low metabolic energy levels (S. Ball et al., 2011; M. a Ballicora et al., 2003).

Kinetic studies of several AGPases of a wide range of bacterial species have been performed to determine their specific positive or negative allosteric regulators and, based on this feature, AGPases have been grouped into nine different classes (Table 1; Ballicora, Iglesias, and Preiss 2003; J Preiss 1978). In general, glycolytic intermediates such as fructose 6-phosphate (F6P), fructose 1,6-biphosphate (FBP), and/or pyruvate enhance the activity of bacterial AGPases, whereas AMP, ADP, and/or inorganic phosphate (Pi) display inhibitory properties. In contrast, AGPases from photosynthetic organisms such as plants and cyanobacteria prefer 3-phosphoglyceric acid (3PGA) as a positive signal produced by the photosynthetic activity, and Pi as the inhibitory signal. The structural overlapping observed for the activators in the different AGPase classes suggests that activator sites for distinct groups are similar or related (M. a Ballicora et al., 2003; J Preiss, 1984, 2010; Jack Preiss, 2006; Ribéreau-Gayon, Sabraw, Lammel, & Preiss, 1971).

Organism	Main carbon Utilization pathway	ADP-Glc PPase Allosteric effectors Class	Activator(s)	Inhibitor(s)
Prokaryotes				
Escherichia coli *Salmonella typhimurium* *Enterobacter aerogenes*	Glycolysis	I	Fru-1,6-bis-P	AMP
Aeromonas formicans *Micrococcus luteus* *Mycobacterium smegmatis* *Rhodopseudomonas viridis*	Glycolysis (+ Reductive carboxylic acid cycle)	II	Fru-1,6-bis-P, Fru-6-P	AMP, ADP
Serratia marcescens *Enterobacter hafniae* *Clostridium pasteurianum*	Glycolysis	III	None	AMP
Agrobacterium tumefaciens *Arthrobacter viscosus* *Chromatium vinosum* *Rhodobacter capsulata* *Rhodomicrobium vaneilli*	Entner–Doudoroff pathway	IV	Pyruvate, Fru-6-P	AMP, ADP
Rhodobacter gelatinosa *R. globiformis* *R. sphaeroides* *Rhodocyclus purpureus*	Glycolysis, Enter–Doudoroff pathway, and reductive carboxylic acid cycle	V	Pyruvate,Fru-6-P, and Fru-1,6-bis-P	AMP, P_i
Rhodospirillum rubrum *Rhodospirillum tenue*	Reductive carboxylic acid cycle	VI	Pyruvate	None
Bacillus subtilis *B. stearothermophillus*	TCA cycle during sporulation	VII	None	None
Cyanobacteria *Synechococcus* PCC 6301 *Synechocystis* Pcc 6803 *Anabaena* PCC 7120	Oxygenic Photosynthesis (Calvin cycle)	VIII	3-PGA	P_i
Eukaryotes Green algae *Chlorella fusca* *C. vulgaris* *Chlamydomonas reinhardtii*	Oxygenic Photosynthesis (Calvin cycle)	VIII	3-PGA	P_i
Higher plants Photosynthetic tissues Plant leaves; e.g.; spinach wheat, pea, *Arabidopsis*, maize, rice, etc.	Oxygenic Photosynthesis (Calvin cycle or Hatch–Slack pathway)	VIII	3-PGA	P_i
Nonphotosynthetic tissues Potato tuber, maize endosperm	Sucrose catabolism and gluconeogenesis	VIII	3-PGA	P_i
Barley and wheat endosperm	Sucrose catabolism and gluconeogenesis	IX	3-PGA and Fru-6-P reverses P_i inhibition	P_i

Table 1. Relationship between carbon metabolism and regulatory and structural properties of AGPases from different organisms (From** Ballicora, Iglesias, and Preiss 2003)**.

The identification of the allosteric binding sites in AGPase has been a matter of intense research in the field of glycogen/starch biosynthesis/regulation in the last decades. The absence of structural information reported for this enzyme in complex with allosteric effectors has led to the characterization of its regulatory mechanism through a biochemical and site-directed mutagenesis point of view. The paradigmatic *E. coli* AGPase (*Ec*AGPase) has been the subject of deep study since 60's. It is classified within class I of AGPases, which are mainly activated by FBP and inhibited by AMP (M. a Ballicora et al., 2003). Chemical modification studies of *Ec*AGPase with PLP by reduction

with $NaBH_4$ pointed to Lys39 residue, located in the N-terminal domain by homology with other AGPases, as an important amino acid implicated in the binding of the activator FBP. This observation has been further supported by direct mutagenesis, where the substitution of this residue per glutamic acid caused variations in *Ec*AGPase activation kinetics (Gardiol & Preiss, 1990; Parsons & Preiss, 1978a). The involvement of the N-terminal portion of AGPase in the allosteric regulation was also observed by the truncation of 11 amino acids, which made *Ec*AGPase insensitive to activation (C. M. Bejar, Ballicora, Iglesias, & Preiss, 2006; Wu & Preiss, 1998). On the other hand, the C-terminal domain of heterotrophic bacterial AGPases is known to play an important role for the activator specificity; this hypothesis is supported by results from experiments with the *Ec*AGPase (class I) and *At*AGPase (class IV) C-terminal chimeric enzymes in which the activator specificity was modified according to the C-terminal domain source (M. A. Ballicora, Sesma, Iglesias, & Preiss, 2002). Based on all these results, it was postulated that the allosteric regulation of AGPase might involve a rearrangement between the N- and C-terminal domains (Jack Preiss, 2014).

OBJECTIVES AND HYPOTHESIS

2. Objectives and hypothesis

Bacteria and photosynthetic eukaryotes evolutionary pathways have selected ADP-Glc as the activated glucosyl nucleotide donor for glycogen and starch biosynthesis, the most common carbon storage and energy reserve polysaccharides in nature (S. G. Ball & Morell, 2003; Recondo & Leloir, 1961). ADP-Glc biosynthesis is mediated by the enzyme AGPase, and is considered the main regulatory step in glycogen and starch production in these organisms (Espada, 1962; J Preiss, 1978). In this respect, evolution led AGPase to acquire allosteric properties to control this key rate-limiting step by essential metabolites in the energetic flux within the cell. In general, AGPase activators are metabolites that represent signals of high carbon and energy content of a particular bacteria or tissue, while inhibitors of the enzyme indicate low metabolic energy levels (S. Ball et al., 2011; Clarisa M. Bejar, Jin, Ballicora, & Preiss, 2006). The identification of AGPase allosteric binding sites has been a matter of intense research in the field of glycogen/starch biosynthesis/regulation in the last decades. To date, only two crystal structures of AGPase have been reported, that of the *Agrobacterium tumefaciens* (*At*AGPase) and the potato tuber AGPases (*St*AGPase; Cupp-Vickery et al., 2008; Jin et al., 2005). Nevertheless, no crystal structure of AGPase in complex with allosteric regulators has ever been reported.

In this work, we focused on the paradigmatic AGPase from *Escherichia coli*. Much effort has been made in the study of *Ec*AGPase allosteric regulation; experimental evidences suggest that both N- and C-terminal domains are involved in *Ec*AGPase allosteric regulation and it was postulated that this phenomenon might be determined by a combined arrangement between both domains (M. A. Ballicora et al., 2002; C. M. Bejar et al., 2006; Jack Preiss, 2014; Wu & Preiss, 1998). However, the absence of structural data has not allowed to delve into the regulatory mechanism of this enzyme at the molecular level.

Our specific aims are:

I) to solve the crystal structure of *Ec*AGPase in complex with its naturally occurring and preferred allosteric negative regulator, AMP, and positive regulator, FBP.

II) to propose a model for the allosteric regulation of *Ec*AGPase

III) to validate the structural information by direct mutagenesis and biophysical/biochemical approaches.

EXPERIMENTAL TECHNIQUES

3. EXPERIMENTAL TECHNIQUES

In this section, an overview of Circular Dichroism (CD) and Electron Microscopy (EM) is given. A careful description of the materials and other techniques used during these studies can be found in the Materials and Methods sections of each chapter.

3.1. Circular dichroism (CD)

The interaction of circulary polarized light with chromophores in optically asymmetric environments is specific and, even though the changes are subtle, they can be measured and interpreted. This fact allows the study of proteins in solution for structural characterization and structural conformational changes. In this work, we focus our CD studies on the determination of the thermal stability of *Ec*AGPase in complex with its physiologically preferred allosteric regulators.

CD, by definition, is the difference between the left circulary polarized light absorption and the right one. Plane-polarized radiation comprises two circulary vectors of equal intensity, one right-handed (RCP) and the other left-handed (LCP). A chromophore in an optically symmetrical environment absorbs the two components. When they recombine, they result in radiation oscillating again in a single plane (Eq. 3.1). However, a chromophore situated in an optically asymmetric environment will absorb unequally each of the two components to a different extent, ΔA.

$$\Delta A = A_L - A_R \qquad \text{eq.3.1}$$

As consequence, the resultant vector describes an ellipse, instead of circle. The ratio of the major and minor axes of the ellipse determines ellipticity (θ). It should be noted that $\theta = \tan^{-1}$ (OL'/OR') (Figure 3.1).

The output of CD instruments, also known as spectropolarimeters, can be either measures of the difference in absorbance between the LCP and RCP or the ellipticity (θ) in degrees.

The numerical relationship between ΔA and ellipticity (in degrees) is:

$$\theta = 32.98\ \Delta A \qquad \text{eq.3.2}$$

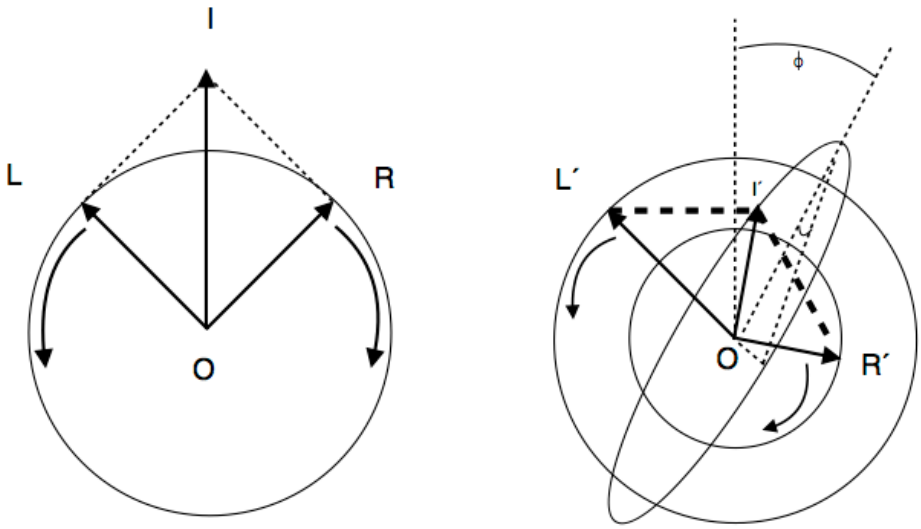


Figure 3.1 Circular dichroism phenomenon. *The relation of ellipticity to the differential absorption of circulary polarized radiation in an optically symmetrical environment (A) and in an optically asymmetrical environment (B) (From Pain 2005).*

In addition, those molecules can also present a different refractive index for the RCP and LCP components that rotate the plane of the polarized emitted light. This phenomenon is known as rotatory optic dispersion (ROD) and it is defined by the angle formed by the major axis of the ellipse and the original plane (Figure 3.1). As it was previously mentioned, this phenomenon depends on the absorption, and thus, it varies with the wavelength of the incident radiation. Generally, spectra represent CD signal (ellipticity) as a function of the wavelength. The bands of the spectra correspond to the electronic transitions happened due to either the electric field, the magnetic field or both of them above the vibrational levels of the ground state and the first excited level ones.

3.1.1 Structural protein characterization by CD

In structural studies, proteins in solution are the chromophores that absorb different energy photons. Generally, spectra are obtained from two regions of the electromagnetic radiation spectrum: Far UV (180nm - 250nm) and Near UV (250 nm – 350 nm). In the

Far UV region, the peptidic backbond is the main chromophore responsible for the signal. On the other hand, the aromatic residues and the disulfide bond are the chromophores involved in the CD signal coming from the near UV region.

1. Peptidic backbone: the amide bond is rotation limited (O=C-NH) and the partial double bond character leads to an electron pair delocalization of the nitrogen atom. The two dihedral angles, phi and psi that define the secondary structure of proteins, confer the chirality. The peptidic bond can undergo two electronic transitions:

n -> π*: Due to the generation of a magnetic dipolar moment during the transition. It is characterized by having low energy since the transition occurs from a non-bonding orbital to an antibonding one. A negative peak at 220 nm is detected.

π -> π*: Due to the electric dipolar moment transition. It involves more energy than the n -> π* transition. It is a linear transition from a bonding orbital to a antibonding π*. A broad positive signal at 190 nm is observed.

2. Aromatic residues: The asymetry of the side chain of the residues and the hydrogen bonds generated due to the solvent polarity exposure are relevant. The extinction coefficient factors for the phenilalanine and the tyrosine are small and, as a result, the transitions give as results weak bands in the spectra. Tyrosine residues show a peak between 275 and 282 nm, with a shoulder at longer wavelengths. The phenilalanine residues exhibit weaker but sharper bands at 262 and 268 nm. Interestingly, the tryptophan residue presents higher side chain asimetry and extinction coefficient factor undergoing several transitions in the 240-295 nm range. In addition, the solvent polarity has a direct effect on the energy required for the transition. Indeed, the hydrogen bonds decrease the energy required for the transitions from the ground state to the excited state. It should be noted that the fine structure in these bands arises from vibronic transitions in which different vibrational levels of the excited state are involved. (Kelly, Jess, & Price, 2005). Weaker aromatic bands are also located in the far-UV region. They can be observed in case of low secondary structure content, i.e., when the n-pi band of the backbone is weak.

3. Disulfide bonds: the absortion happens between 250-300 nm and it is low. In addition, if the protein also contains aromatic residues, the corresponding band of the disulfide bonds can be masked.

3.1.1.2. Thermal induced-unfolding process followed by Far UV CD

The thermal-induced unfolding process can be followed by CD in the far UV region providing useful information about the cooperativity during the denaturation event and stability of the folded state of the native protein that might vary upon neither denaturing agents or stabilizating substrates addition (Kelly et al., 2005). In fact, spectra can represent ellipiticity at a given wavelength (generally at 222 nm, maximun of absorption for the peptide backbond) as function of the temperature or the fraction of unfolded-folded protein as function of the temperature. The simplest model for a protein unfolding reaction is a two-state model, N < -- > U, where N is the native state defined by a 3D structure and biological function and U is the unfolded state. If N state is subjected to different conditions that might cause the unfolding process (high temperature, pH, pressure or high concentration of chemical denaturants) the organized native structure can be lost in a more or less cooperative manner to form the U state. Generally, the term "unfolded" state refers to that can refold to the native state if perturbing condition is removed, whereas "denaturation" is a process that yields an altered structure that cannot readily refold to the native state (Gore, 2000). The thermal-induced unfolding process data can be fitted in order to determine thermodynamic parameters as ΔH_{vH} or T_m (Chenal et al., 2010):

The fraction of native protein f_N that populated the native state, N, is given by:

$$f_N = \frac{N}{N+U} = \frac{1}{1+\left(\frac{U}{N}\right)} = \frac{1}{1+K_T},$$

where U is the concentration of protein in the thermal-unfolded state and K the equilibrium constant at the given temperature: $K_T = U/N$.

The free energy of unfolding ΔG at any temperature is related to the enthalpy *ΔH* and the entropy *ΔS* of the unfolding reaction

$$\Delta G = \Delta H - T\Delta S \text{ with } \Delta H = \Delta H_{vH} + \Delta Cp(T - T_m) \text{ and } \Delta S = \Delta S_m + \Delta Cp \ln(T/T_m),$$

given

$$\Delta G = \Delta H_{vH} + \Delta Cp(T - T_m) - T\left(\Delta S_m + \Delta Cp \ln(T/T_m)\right),$$

given the Gibbs-Helmholtz equation:

$$\Delta G = \Delta H_{vH} - T\Delta S_m + \Delta Cp\left(T - T_m - T\ln(T/T_m)\right)$$

where ΔCp is the variation of heat capacity at constant pressure, T_m the temperature of half denaturation (where U=N), ΔH_{vH} and ΔS_m the van't Hoff variation of enthalpy and the entropy at the melting point.

In the vicinity of the half-melting temperature ($T \approx T_m$): $(T - T_m) \approx 0$ and $\ln(T/T_m) \approx 0$. Then, the expression of the free energy of unfolding in the vicinity of the melting point ΔG_m can be approximated to:

$$\Delta G_m = \Delta H_{vH} - T\Delta S_m$$

The variation of entropy in the vicinity of the Tm can be replaced by the van't Hoff enthalpy variation. Indeed, at the melting temperature, U=N given $K_{Tm} = 1$. The equilibrium constant K being related to the free energy $\Delta G = -RT\ln K$, at the melting point: $\Delta G_m = -RT\ln K_m = 0$. Hence, $\Delta G_m = \Delta H_{vH} - T\Delta S_m = 0$ from where ΔS_m can be expressed as follows: $\Delta S_m = \frac{\Delta H_{vH}}{T}$ and used to express ΔG_m only as a function of ΔH_{vH}:

$$\Delta G_m = \Delta H_{vH} - \frac{T\Delta H_{vH}}{T_m}$$

With the ΔCp contribution, the expression is

$$\Delta G_m = \Delta H_{vH} - \frac{T\Delta H_{vH}}{T_m} + \Delta Cp\left(T - T_m - T\ln(T/T_m)\right)$$

From the free energy expression related to the equilibrium constant K,

$\Delta G = -RT\ln K$, K is given by $K = \exp^{\left(\frac{-\Delta G}{RT}\right)}$

In the vicinity of the half melting point, ΔG_m replaces ΔG, ΔCp term can be considered negligible (T-Tm -> 0 and Ln(T/Tm) -> 0) and K becomes

$$K = \exp^{\left(\frac{-1}{RT}\times\left(\Delta H_{vH} - \frac{T\Delta H_{vH}}{T_m}\right)\right)} = \exp^{\left(\frac{\Delta H_{vH}}{R}\times\left(\frac{1}{T_m}-\frac{1}{T}\right)\right)}$$

Finally, the fraction of native protein is expressed as followed, providing an estimation of ΔH_{vH} and T_m:

$$f_N = \frac{1}{1+\exp^{\left(\frac{\Delta H_{vH}}{R}\times\left(\frac{1}{T_m}-\frac{1}{T}\right)\right)}} \qquad \text{eq. 3.3}$$

The determination of the T_m values of *Ec*AGPase in complex with its allosteric regulators allow us to characterize the thermal stabilization/destabilization effect of this interaction. In combination with structural information, CD might provide insights into the allosteric regulation of *Ec*AGPase.

3.2. Electron microscopy

In this work, electron microscopy (EM) has been one of the principal techniques used for the determination of the quaternary structure of *Ec*AGPase in solution.

In the last decades, EM has become a major tool for structural biology over the molecular to cellular size range, overlapping with methods as light microscopy (LM), X-ray crystallography and nuclear magnetic resonance spectroscopy (NMR).

There are different varieties of EM, that can be classified in the way the image is generated form the specimen, as SEM (scanning electron microscopy), REM (reflection electron microscopy) and TEM (transmission electron microscopy) which is the matter of discussion in this section.

3.2.1. TEM components

TEM is analogous in some aspects to LM where the light-visible photons serve as the source of radiation and, once they pass through the specimen, they are refracted by glass optical lenses to form an image. The main difference between LM and EM is that, in EM, the radiation is a beam of electrons emitted by a source that is housed under high vacuum. Electrons are then accelerated down the microscope column and, after passing through the specimen, scattered electrons are focused by the electromagnetic lenses to form an image which can be recorded. Thus, similar to a conventional light microscope, the TEM consists of (i) an electron source, (ii) a series of lenses, and (iii) an image detecting system (Figure 3.1; Milne et al. 2013)

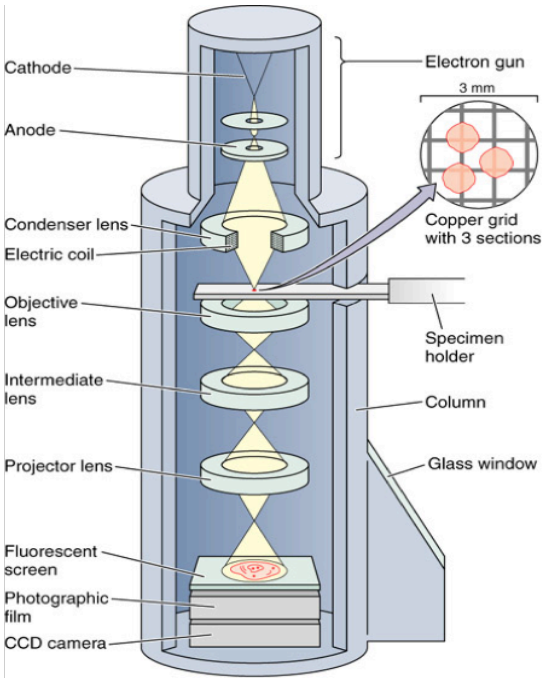


Figure 3.2 Schematic representation of a classical TEM. *(From www.ufrgs.br).*

Different types of EM electron sources are available; the usual one is a heated tungsten filament to 2000-3000 ºC, at this temperature, the thermally emitted electrons are later accelerated by an electric field between the anode and filament. Another common electron source is a lanthanum hexaboride crystal (LaB_6), which displays a low work function. This parameter describes the minimum energy needed to remove an electron from the source. For this reason, LaB_6 allows lower temperatures than tungsten, leading to greater brightness and longer life of the source. In addition, LaB_6 produces electrons from a smaller effective area of the crystal vertex, which means higher coherence of the electron beam. At present, field emission gun (FEG) is the most advanced electron source and is used in high performance microscopes. It is composed by a single crystal tungsten sharpened to give a tip radius lower than in the case of LaB6. The tip is coated with ZrO_2, which lowers the work function. These characteristics allows the FEG beam to be smaller in diameter, more coherent, and ~500X brighter, with a very small spread of energies. In all cases, thermally emitted electrons are extracted from the source by a strong potential gradient at the emitter surface (field emission), and then accelerated through voltages of 100-300 kV (Orlova & Saibil, 2011; Shimizu, Kataoka, Kawai, & Tanaka, 1975).

The basic function of electron optical lenses is to deflect electrons, as crystal lenses deflects photons in LM. Since electrons are negatively charged particles of small mass, they can be oriented by an electro-magnetic field. This way, magnetic condenser lenses convert the diverging electron beam into a parallel one illuminating the specimen. The objective lens provides the primary magnification and is the most important optical element of the electron microscope, as its aberrations play a key role in imaging. An objective aperture is placed in the back focal plane of this lens allowing the absorption of scattered electrons, thus improving the image contrast. Intermediate and projector lenses further magnify the image before the electrons arrive at the detector (Figure 3.1; (Williams & Carter, 2009).

Finally, the image can be view projecting the magnified electron image onto a fluorescent viewing screen coated with a phosphor or scintillator material, located within a chamber in the microscope (Figure 3.1). The image can be photographically recorded using a conventional photographic film composed by microscopic silver halide crystals. Nowadays, digital imaging has practically replaced recording on films and is widely used in TEM, allowing the specimen monitorization and image acquisition in real time. The most popular cameras are based on Charge-Coupled Device (CCD) or Complementary Metal Oxide Semiconductor (CMOS) sensors that convert the analogue

optical signal into digital format. However, the latest detection systems have been matter of important improvements, based on the incorporation of direct electron detectors (DEDs), which can directly count the image-forming electrons from the microscope. These latest advances have revolutionized the technique, allowing to reach the atomic resolution (Callaway, 2015).

3.2.2. Interaction of electrons with the specimen

Resolving power is one of the most significant differences between LM and EM. It is worth noting that resolution is directly influenced by the wavelength of the imaging radiation source by the next relationship,

Resolution (r) = λ/ (2NA)

where r is resolution, NA the numerical aperture and λ the imaging wavelength. Based on this equation, the shorter the wavelength, the higher the attainable resolution. In this regard, the resolution achieved with visible light (wavelengths ~ 4000–7000 Å; 1 Å = 10−10 m) is significantly less than that accomplish with electron sources in a typical TEM (wavelength of ~ 0.02 Å for operation at 300 kV) (Milne et al., 2013). In EM, the wavelength of electron radiation itself does not impose a limit on the resolution that could be obtained when imaging biological macromolecules by EM, however, the limitation in this regard is the stability of the specimen position and the radiation damage that results from the strong interaction of electrons with the organic matter. The nature of this interaction depends on the electron energy and sample composition. In this regard, electrons can either pass through the sample without any interactions or can be deflected by the specimen. These last, can interact with the electrostatic field of the nucleus, screened by the outer orbital electrons of specimen atoms, and some electrons may collide or nearly collide with the atomic nuclei, suffering high angle deflections or even backscattering (Figure 3.2A). Of the interacting electrons, some are scattered without energy loss (elastic scattering), but others transfer energy to the specimen (inelastic scattering). This energy transfer can ionize atoms in the specimen, induce X-ray emission, chemical reactions or induce secondary electron scattering, all of which can change the specimen structure. Thus, radiation damage of specimens is a significant limitation in high-resolution imaging of biological molecules (Figure 3.2B). To reduce this

phenomenon during area selection, alignment, and focusing, special “low dose” systems are used to deflect the beam until the final step of image recording (Orlova & Saibil, 2011; Williams & Carter, 2009).

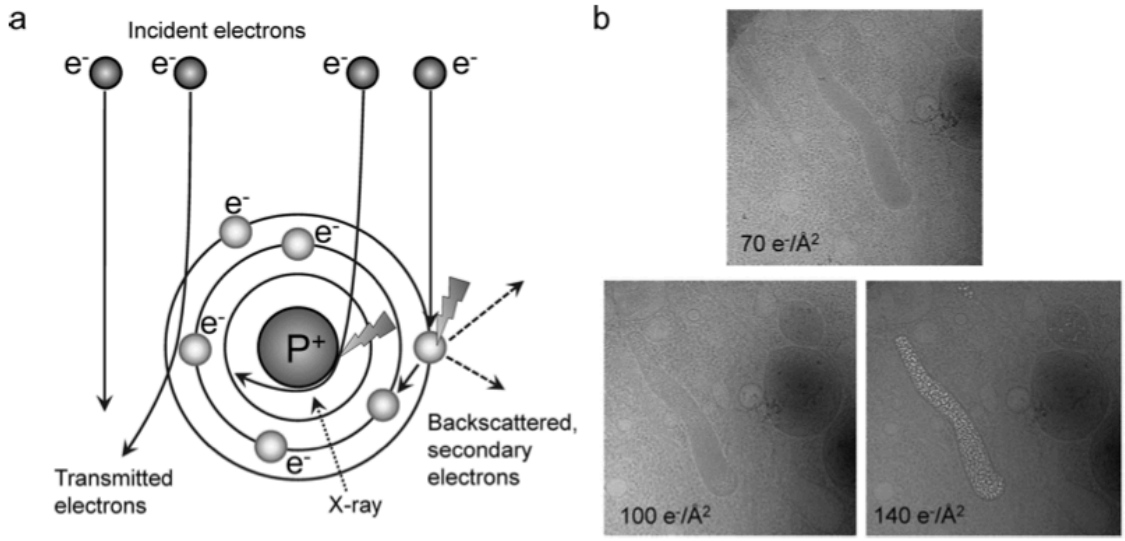


Figure 3.3 Interaction of the electron beam with the sample. *(A) Schematic representation of possible electron interactions with the specimen. (B) Effect of electron beam damage on a cryo image of a cell. The electron dose is shown on the images. Increasing dose causes damage to cellular structures (From Orlova and Saibil 2011).*

3.2.3. Specimen nature

Electron microscopy can be used to provide structural information on a range of biological specimens of different shapes, sizes, and biochemical states.

Because electrons interact strongly with matter, the column of the microscope must be kept under high vacuum to avoid unwanted scattering by gas molecules in the electron path. Therefore, the EM specimen must be in the solid state for imaging, and special preparation techniques are necessary to either dehydrate or stabilize hydrated biological samples under vacuum. It is worth noting that organic samples are predominantly made of light elements, including hydrogen, carbon, nitrogen, and oxygen, thus, images will not contain much contrast. For this reason, structural information can thus be obtained either by, (i) acquiring images from electrons scattered by stained samples or (ii) the scattering of native specimens supported in vitreous ice. In this regard, an electron cryomicroscope (cryo-TEM or cryo-EM) is a TEM with a specimen holder capable of

maintaining the sample at liquid nitrogen or liquid helium temperatures. This allows imaging specimens prepared in vitreous ice (Milne et al., 2013; Thompson, Walker, Siebert, Muench, & Ranson, 2016).

Two extreme sample groups can be performed according to its nature: biochemically purified, isolated complexes (single particles or ordered assemblies such as 2D crystals) and unique, individual objects such as tissue sections, cells, or organelles.

The sample preparation protocol, along with the data processing method, will depend on the nature of the sample: (i) single particle analysis for purified, 'homogenous' protein complexes, (ii) helical reconstruction for protein assemblies with helical symmetry, (iii) tomography for 'unique' assemblies such as organelles and cells, and (iv) 2D crystallography for proteins, significantly smaller than 150 kDa that form ordered 2D arrays (Thompson et al., 2016).

In this work, steps towards biological structure determination by EM are discussed, focusing on sample preparation and imaging of specimens for single particle analysis (single particle-EM), although many of the concepts and considerations discussed are transferrable.

3.2.4. Sample preparation

Single-particle EM depends on the computational averaging of thousands of images of identical particles, so the ideal single-particle specimen is a pure, homogenous sample. Compositional heterogeneity of the sample can be reduced biochemically using appropriate protein purification methods, however, structural arrangements and intrinsic flexibility may produce a broad range of conformations. Although novel imaging processing algorithms and detectors allows dealing with some heterogeneity in the dataset during the processing stages, the structural heterogeneity can be reduced by biochemical approaches to simplify structure determination. In this sense, solution buffer conditions can dramatically alter the appearance of a sample in the microscope being recommended to perform a throughput screening to identify optimal buffer compositions. Chemical crosslinking may also be a useful tool for reducing sample heterogeneity, however, this approach may introduce artefacts. Other way to reduce conformational flexibility include ligand-induced stabilization, where molecules such as a substrate, ligand, inhibitor or protein/nucleic acid binding partners can promote stable complex

formation, a strategy that is well established in X-ray crystallography (Cheng, Grigorieff, Penczek, & Walz, 2015; Egli, 2010).
Optimizing the sample before attempting to image it remains vital, however, the most informative way to judge the quality of a protein sample is to visualize it by negative-stain EM (see above).

3.2.5. Grid preparation

Most biological EM work is done on small support discs called grids, commonly made from a fine mesh of a metal (Cu, Ni, Ag or Au). The mesh size, or number of squares across the grid, is defined as the number of squares in one inch; 200–400 mesh grids (20-40 squares across in each direction) are the most commonly used for cryo-EM (Thompson et al., 2016).
The metal grid supports a carbon film on top, which is going to be in contact with the sample. The surface features of this film can be altered by a variety of processes, including the modification of its charge properties. This is accomplished by glow discharging; the carbon-coated grids are placed inside a partly evacuated chamber connected to a power supply. When high voltage is applied between the cathode and anode at each end of the chamber, the electron potential ionizes the gas within the chamber (Baker, Olson, & Fuller, 1999) . These negatively charged ions then deposit on the carbon, giving the carbon film an overall hydrophilic surface. This technique can change the partitioning of the sample into the grid and has to be optimized for each sample.
Once the specimen is ready, the grid must be prepared for imaging by TEM. The choice of preparation technique is ultimately determined by the nature of the sample but common methods include negative staining and vitrification, techniques that will be describe in detail through the next sections.

3.2.5.1. *Negative staining*

Negative stain is a common method for examining a specimen at room temperature. This protocol usually comprises three steps: (i) particle adsorption onto the grid, (ii) washing, and (iii) incubation with a solution of a heavy metal salt, commonly 1–2% (w/v) uranyl acetate or uranyl formate. The grid is finally blotted to ensure a thin layer of stain. The

staining process quickly dehydrates the specimen and envelops it in stain; where the sample is visualized by the absence of stain, negative staining occurs; where the sample itself becomes stained, positive staining occurs (Figure 3.3; Ohi et al. 2004).

The resulting shell of heavy metal atoms generates high contrast and signal to noise ratio (SNR), however, it only provides surface topology information. Moreover, the dehydration process can affect the sample state and staining artefacts are also probable if the metal salt solution deposition is uneven. This method is thus restricted to a modest resolution (>20 Å), nevertheless, it can be improved varying the metal grain size, as this parameter determines how well the stain envelope reflects the structure of the particle. One of the principal advantages of negative stain is the high speed of sample preparation, which makes it ideal for assessing sample purity, concentration or heterogeneity. Low-resolution 3D reconstructions from stained data can also provide starting models for higher resolution cryo-EM studies (Ohi et al., 2004; Orlova & Saibil, 2011; Thompson et al., 2016).

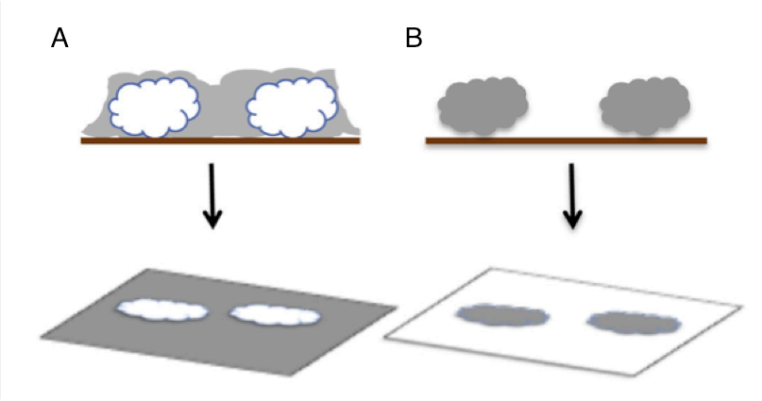


Figure 3.3. Schematic representation of (A) negative stain and (B) positive stain

3.2.5.2. *Vitrification - Plunge freezing*

Imaging of cryogenically immobilized samples by EM is known as cryo-EM. The sample must be frozen extremely rapidly, at a rate of -106 ºC/s, to allow the formation of vitreous ice. If freezing occurs too slowly, or the specimen is subsequently warmed above the temperature at which water devitrifies (-137 ºC), crystalline ice is formed which degrades image quality as this type of ice diffracts electrons. (Thompson et al., 2016). Cryo-immobilisation can be achieved following three steps: (i) particle adsorption onto the grid,

(ii) blotting to generate a thin film of liquid and (iii) rapidly freezing by plunging the grid into a cryogen (Figure 3.4). It is important to optimize the incubation time, sample distribution (blotting time) and ice thickness to achieve the best images. The most widely used cryogens are liquid ethane or propane, cooled by liquid nitrogen. The ice layer should be as thin as possible without distorting the specimen and avoiding ice contamination. Commercially available freezing apparatus offer different features that may be beneficial when working with certain samples. In this regard, the humidity control in the freezing chamber can reduce ice contamination during the process and improve the final image quality (Baker et al., 1999; Cheng et al., 2015; Milne et al., 2013; Thompson et al., 2016). One of the main advantages of this method is that cryo-immobilisation leads to the visualization of the sample in a native-like state. In this approach, low electron doses are needed to prevent radiation damage, which results in poor contrast. However, processing of data can yield high-resolution, 3D structural information.

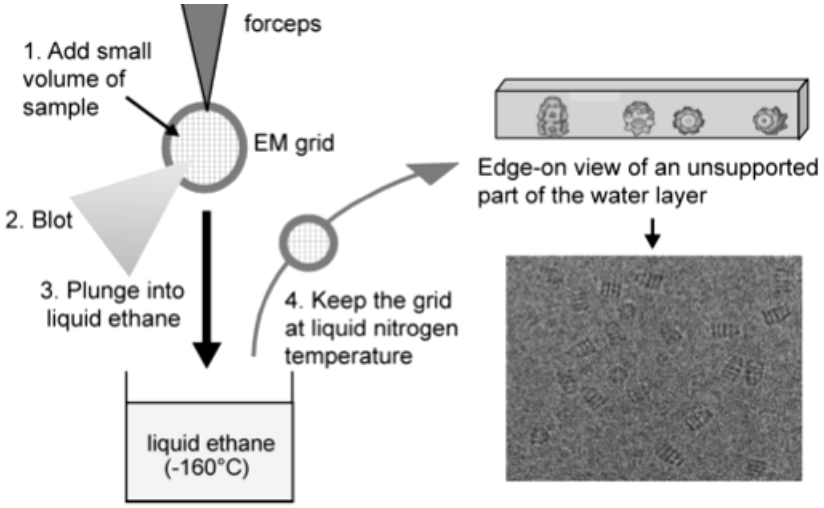


Figure 3.4. Schematic representation of plunge freezing and a vitrified layer with an example cryo-EM image (From Milne et al. 2013)

3.2.6. Image processing

In EM, the image suffers from aberrations that can be corrected by the contrast transfer function (CTF). The CTF mathematically describes how these aberrations modify the image of the sample and its accurate estimation is essential for both, the initial evaluation of micrograph quality and subsequent structure determination (Baker et al., 1999). The

parameters that have to be known to calculate this function are: acceleration voltage, spherical aberration, defocus, astigmatism, and percentage of amplitude contrast. Voltage and spherical aberration are instrument parameters that are usually used without further refinement and the contribution of the amplitude contrast is typically assumed as 5%–10% for cryo-EM images (Cheng et al., 2015). As has been mentioned before, samples are predominantly made of light elements, including hydrogen, carbon, nitrogen, and oxygen. This means that an in-focus image in cryo-EM will not contain much contrast (magnitude of variation between bright and dark features) as the electron transmission across a sample is approximately constant. The main method for generating contrast is to deliberately underfocus the objective to induce phase shifts in the scattered electrons and thus contrast in the images (Newcomb, Moyer, Lee, & Stupp, 2012). However, this approach leads to a systematic alteration of the image data and because of that, although the defocus is set during data collection, more accurate values for defocus and astigmatism have to be calculated for each micrograph.

3.2.7. Particle picking

Once a dataset has been collected, it is necessary to obtain individual particles from each micrograph. This process in known as particle picking and can be performed in a manual, semi-automated, and fully automated manner. The quality of the selected particles is a major factor in the subsequent analysis.

3.2.8. Alignment and 2D classification

The first step in single-particle EM structure determination is the alignment and grouping of the 2D image dataset into homogenous subsets. A micrograph often contains particles in multiple different orientations and/or conformations, and so to get more representative image averages, a method is required to group similar particle images. This is usually carried out using one of several data analysis and image classification algorithms, such as multi-variate statistical analysis and hierarchical ascendant classification, or K-means classification (Orlova & Saibil, 2011).

The reason to start with 2D analysis is that (i) 2D datasets contain image artefacts, invalid particles, or simply empty fields that should be removed; (ii) the angular distribution of the particle views is unknown and if the set is dominated by just a few views, 3D analysis is unlikely to succeed; and (iii) computational ab initio 3D structure determination requires high-SNR input data, as is present in high-quality class averages. Is essential thus to have enough images to cover all the possible orientations of the particle. The degree of relatedness between individual particle images (Figure 3.4A) is used to identify clusters of similar images within the data set. Related images are averaged to obtain characteristic projection views of the complex at much higher signal-to-noise ratios than in the original images (Figure 3.4B). Iteration of this classification process, using characteristic views of the newly generated class averages as alignment references, improves the accuracy of alignment and permits visualization of finer structural features (Figure 3.4C; Cheng et al., 2015; Orlova & Saibil, 2011; Williams & Carter, 2009)

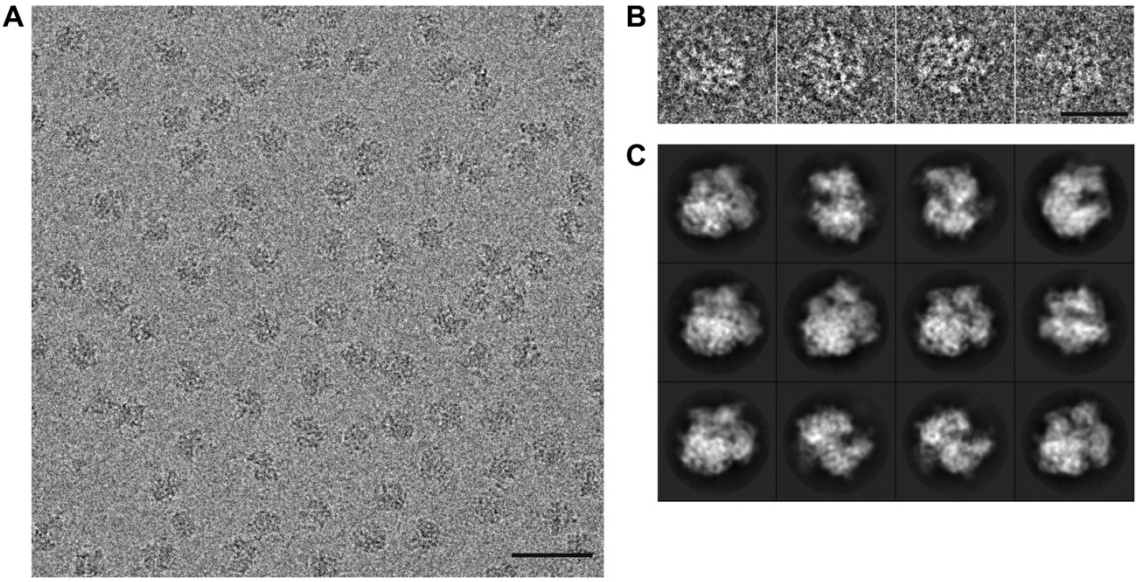


Figure 3.5 (A) Part of a recorded micrograph, (B) isolated particles and (C) class averages of the 70S T. thermophiles ribosome. *The scale bar corresponds to 20nm* (From *Bai et al. 2013).*

3.2.9. 3D reconstruction

Generation of a 3D reconstruction from the 2D electron microscopic projection views of the molecular complex is dependent upon knowing the relative orientations of all of the particles. The steps involved in computational ab initio structure determination are

mathematically complex, and take advantage of the central projection theorem, which states that, for a 3D object, the Fourier transform of each 2D projection is a central slice through the 3D Fourier transform of the object (Figure 3.6; Milne et al. 2013; Bragg 1929). Thus, once the 3D Fourier transform is built up from a collection of 2D images spanning a complete range of orientations, Fourier inversion enables recovery of the 3D structure. Initial maps can be refined to higher resolution using projection-matching refinement strategies, which modifies the orientation parameters of single-particle images (projections) to achieve a better match with reprojections computed from the current approximation of the structure. If appropriate, available atomic coordinates of corresponding protein components derived from X-ray or NMR analyses can be then fitted into the final structure. Various software packages, including EMAN (Ludtke, Baldwin, & Chiu, 1999), FREALIGN (Grigorieff, 2007), and SPIDER (Shaikh et al., 2009), were successfully used for the near-atomic resolution structure reconstructions

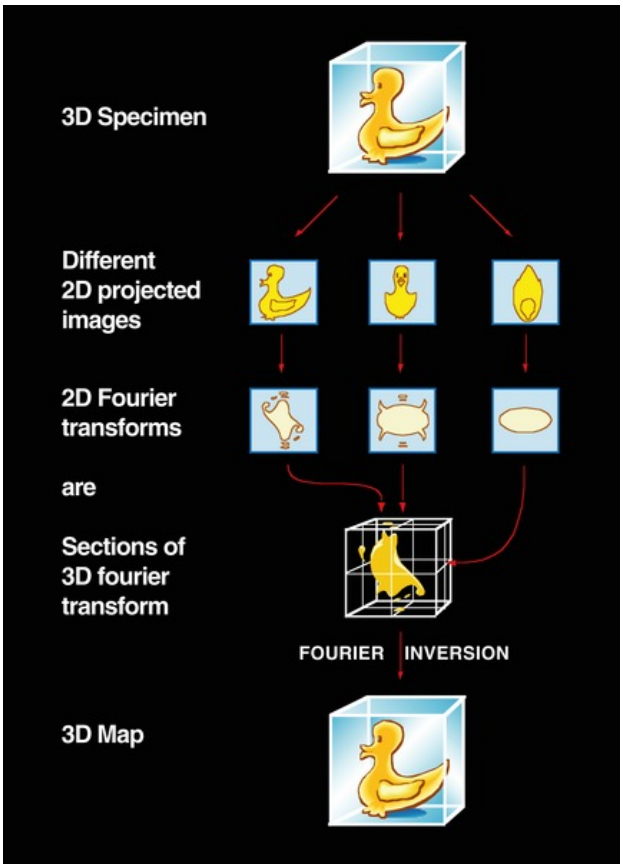
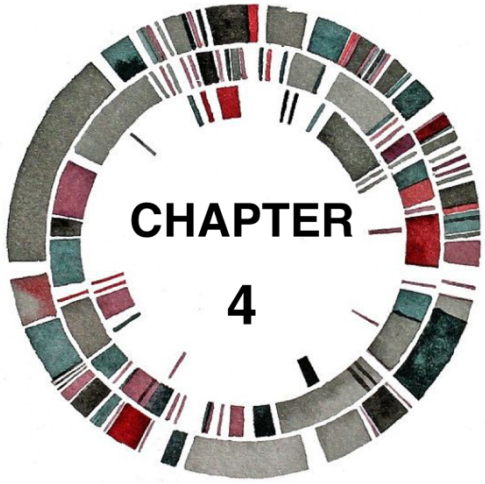


Figure 3.6 Schematic illustration of the EM 3D map reconstruction. *Projection images of the object, each with a different orientation, have 2D Fourier transforms that correspond to sections (indicated by red arrows) through the 3D Fourier transform of the original object. Thus, once the 3D Fourier transform is built up from a collection of 2D images spanning a complete range of orientations, Fourier inversion enables recovery of the 3D structure (From Milne et al. 2013).*



STRUCTURAL CHARACTERIZATION OF *Ec*AGPASE ALLOSTERIC REGULATION

4. Structural characterization of *Ec*AGPase allosteric regulation.

To advance in the comprehension of the molecular mechanism that governs AGPase allosteric regulation in bacteria, we focused in the paradigmatic AGPase from *Escherichia coli* (*Ec*AGPase) and its physiologically preferred positive and negative allosteric regulators, FBP and AMP, respectively (Ghosh & Preiss, 1966). Specifically, the aim of this study was to identify the *Ec*AGPase allosteric binding sites by X-ray crystallography and propose a model of allosteric regulation in combination with biophysical data.

The references of this chapter are based on the publication, mentioned in "Publications" section, *Structural Bases of Glycogen Biosynthesis Regulation in bacteria* (Cifuente et al., 2016).

4.1. Materials and methods

***Ec*AGPase cloning, expression and purification** — The full-length *glgC* gene from *E. coli* BL21 was amplified by standard PCR using oligonucleotide primers *glgC*_NdeI_Fwd (5´ gggaattccatatggttagtt 3´) and *glgC*_XhoI_Rev (5´ ccgctcgagtcatcgctcctg 3), Phusion DNA Polymerase (New England Biolabs) and purified genomic DNA as template. The PCR fragment was digested with NdeI and XhoI and purified by agarose gel electrophoresis. The fragment was ligated to the expression vector pET29a (Novagen) using T4 DNA ligase, generating pET29a-*Ec*AGPase. The recombinant *Ec*AGPase has no additional amino acids when compared to the native enzyme. *E. coli* BL21(DE3) cells transformed with pET29a-*Ec*AGPase were grown in 3,000 ml of LB medium supplemented with 25 μg/ml of kanamycin at 37°C. When the culture reached an OD600 of 0.8, the *Ec*AGPase expression was induced by the addition of 1 mM isopropyl β-thiogalactopyranoside (IPTG) and further incubated at 18°C overnight. All *Ec*AGPase purification procedures were carried out at 4°C. Cells were harvested by centrifugation

at 5,000 x g and resuspended in 40 ml of 50 mM Hepes pH 8.0, 5 mM MgCl2, 0.1 mM EDTA, 10% sucrose (w/v; solution A), containing protease inhibitors (Complete EDTA-free; Roche) and 10 mg/l of lysozyme (Sigma). Cells were then disrupted by sonication (five cycles of 1 min each) and centrifuged for 15 min at 20,000 x g. The supernatant was dialyzed twice against solution A by using an 100,000 Da molecular mass cutoff dialysis membrane. The solution was then applied to a DEAE FF column (5 ml; GE Healthcare) equilibrated with solution A. Elution was performed with a linear 0-0.5M NaCl gradient in 100 ml. Enzyme-positive fractions were pooled and precipitated with ammonium sulfate to 1.2 M (solution B). The resultant suspension was centrifuged for 20 min at 20,000 x g and the resulting supernatant applied into a Phenyl Shodex HIC PH-814 equilibrated in solution B. The enzyme was eluted with a linear gradient of 100% solution B to 100% solution A, in 50 ml. The most active fractions were pooled, concentrated to 10 mg/ml by an Amicon-Ultra spin concentrator (Merck Millipore) with a 100-kDa-molecular mass cutoff, and stored at -80°C.

Gel filtration — Gel filtration chromatography was performed using a Shodex KW-802.5 column equilibrated in 50 mM Tris-HCl pH 7.5, 100 mM NaCl at 0.5 ml min-1. The column was previously calibrated using gel filtration standards (Sigma) including b-amylase (200 kDa), alcohol dehydrogenase (150 kDa), bovine serum albumin (66 kDa), carbonic anhydrase (29 kDa), and cytochrome c (12.4 kDa).

***Ec*AGPase Crystallization and Data Collection** — Crystallization trials were carried out in sitting drop 96 well plates by using a mosquito crystal robot (TTP Labtech). Crystals of *Ec*AGPase•AMP•SUC were obtained by mixing 0.25 µl of *Ec*AGPase at 5 mg/ml in 5 mM ADP, 6.6 mM HEPES pH 8.0, 50 mM MgCl2, 10 mM EDTA and 12 % (w/v) sucrose with 0.25 µl of mother liquor containing 16 % (w/v) PEG 10,000 and 100 mM imidazole pH 7.8. Crystals of *Ec*AGPase•FBP were obtained by mixing 0.25 µl of *Ec*AGPase at 5 mg/ml in 3.6 mM FBP, 6.6 mM HEPES pH 8.0, 50 mM MgCl2, 10 mM EDTA and 12 % (w/v) sucrose with 0.25 µl of mother liquor containing 22 % (w/v) PEG 4,000 and 200 mM ammonium sulfate. Crystals grew in 3-15 days and were transferred to a cryo-protectant solution containing 25% ethylene glycol and frozen under liquid nitrogen. Crystals screening were performed at Diamond Light Source (DLS) (Didcot, Oxfordshire, United Kingdom). *Ec*AGPase•AMP•SUC complete dataset was collected at I04 beamline (DLS) with oscillation angle of 0.15° for a total of 1200 images using a Pilatus 6M-F pixel detector. *Ec*AGPase•FBP dataset was collected at I04-1 beamline

(DLS) with oscillation angle an 0.10° for a total of 1200 images using a Pilatus 2M pixel detector. X-ray data were processed using the program XDS23. *Ec*AGPase•AMP•SUC and *Ec*AGPase•FBP forms crystallized in space group P 21 with 16 molecules in the asymmetric unit and diffracted to a maximum resolution of 2.67 Å and 3.04 Å, respectively.

***Ec*AGPase Structure Determination and Refinement** — The crystal structure of *Ec*AGPase•AMP•SUC was solved by molecular replacement with the program Phaser24 using a tetramer from AGPase from *A. tumefaciens*, PDB atomic coordinates 3BRK (Cupp-Vickery et al., 2008), as search model and followed by automated model-building using Phenix (phenix.autobuild) (Adams et al., 2010). The *Ec*AGPase•FBP crystal structure was solved by molecular replacement using the *Ec*AGPase•AMP•SUC model. The complete models were obtained using alternate cycles of manual model-building using COOT (Emsley, Lohkamp, Scott, & Cowtan, 2010) and Phenix (phenix.refine) or Refmac5 (Murshudov et al., 2011). NCS restraints were relaxed towards the end of refinement and differences between chains subsequently modeled. The ligands topology for AMP, FBP and SUC was refined using CIFs (Crystallographic Information File) generated using PRODRG server (Schüttelkopf & Van Aalten, 2004) or Grade web server (Global Phasing Ltd.). It is worth noting that the electron density observed for the connecting region between the two phosphate groups of FBP is poor, leading to two possible conformations (i) the furanose and (ii) the open chain forms. Based on our interpretation of the maps, the furanose form is shown. During the refinement, the structure geometry was validated using Molprobity (Chen et al., 2010) and the ligands checked using pdb-care (PDB CArbohydrate REsidue check) (Lütteke & von der Lieth, 2004). Final sugar analysis and validation was performed using CCP4 package Privateer (Agirre, 2015). Atomic coordinates and structure factors have been deposited with the Protein Data Bank, accession codes 5L6V (*Ec*AGPase•AMP•SUC) and 5L6S (*Ec*AGPase•FBP). Molecular graphics and structural analyses were performed with the UCSF Chimera package (Pettersen et al., 2004).

Location of the ATP and G-1-P binding sites in EcAGPase — The location of the ATP binding site in *Ec*AGPase was determined by structural superposition with (i) the crystal structure of GlmU in complex with ATP (4K6R) (Vithani, Bais, & Prakash, 2014) and that of *St*AGPase in complex with ATP and ADPG (1YP4) (Jin et al., 2005). GlmU is an

essential bifunctional uridyltransferase that catalyzes the CoA-dependent acetylation of GlcN-1-PO4 to form GlcNAc-1-PO4 and its subsequent condensation with UTP to form pyrophosphate and UDP-GlcNAc. The G-1-P binding site in *Ec*AGPase was determined taking into account the location of the glucose moiety of sucrose in the *Ec*AGPase•AMP•SUC crystal structure.

EcAGPase Structure Analysis — Analysis of the crystal structures was carried out using Chimera (Pettersen et al., 2004). Assignment of the secondary structural elements was performed automatically using Chimera implementation of DSSP with parameters "H-bond energy cutoff = -0.5 kcal/mol", "minimum helix length = 3" and "minimum strand length = 2". Contacts were analyzed using the default settings considering a Van der Walls overlap of -0.4 Å. Structure superimposition was perform using the tool MatchMaker. Multiple sequence alignments were performed using Promals3D (Pei, Kim, & Grishin, 2008).

Circular dichroism thermal unfolding analysis — Thermal unfolding transitions were recorded on a J-810 CD spectropolarimeter (Jasco Corp., Tokio, Japan) at 222 nm by using Hellma 110-QS quartz cuvettes with a 1 mm optical path. Thermal unfolding was recorded at protein concentrations of 1 μM in 50 mM Tris-HCl pH 8,0, 150 mM NaCl between 20 °C and 90 °C with a heating rate of 1°C/min. Transitions were tentatively fitted according to equation 1 to obtain the apparent melting temperature (*Tm*) (Sotomayor-Pérez, Subrini, Hessel, Ladant, & Chenal, 2013).

$$y = \frac{\left(y_f + m_f \cdot T\right) + \left(y_u + m_u \cdot T\right) \cdot \mathrm{e}^{\left(\frac{\Delta Hm}{R}\right)\cdot\left(\frac{1}{Tm}-\frac{1}{T}\right)}}{1 + \mathrm{e}^{\left(\frac{\Delta Hm}{R}\right)\cdot\left(\frac{1}{Tm}-\frac{1}{T}\right)}}$$

(Eq. 1)

where y represents the observed CD signal at 222 nm, y_f and y_u are the y-axis intercepts and m_f and m_u the slopes of the pre- and post-transition baselines, respectively, T is the temperature in K, Tm is the melting temperature, and ΔHm is the enthalpy change of unfolding at Tm. Curve fitting was performed with Prism GraphPad.

4.2. Results and discussion

4.2.1. Overall structure of *Ec*AGPase

The paradigmatic *Ec*AGPase was purified to homogeneity following the protocol described in Materials and methods section. It is important to emphasize that the *Ec*AGPase construct has no additional amino acids when compared with the native full-length enzyme (Figure 4.1A). SDS-PAGE and analytical gel filtration confirmed that *Ec*AGPase displays the homotetrameric structure expected for bacterial AGPases in solution (Figure 4.2B; Jack Preiss 2014).

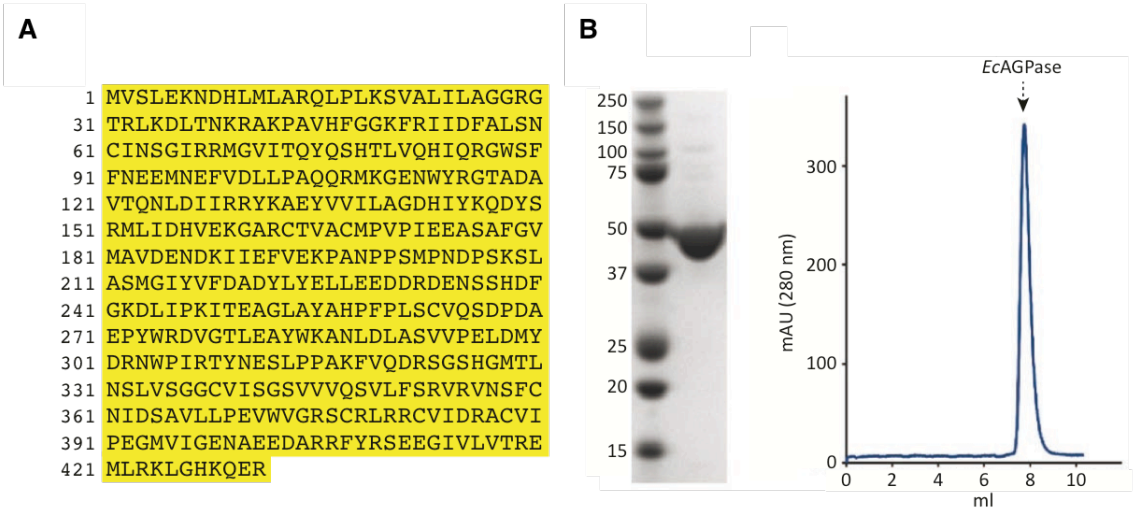


Figure 4.1. Recombinant production of EAGPase. *(A) Amino acid sequence of EcAGPase construct used for heterologous production in E. coli. EcAGPase was expressed in E. coli as the native full-length protein without additional amino acids. (B) SDS-PAGE showing purified EcAGPase with an apparent monomeric molecular mass of ~50 kDa. The analytical gel filtration confirms that the enzyme is a homotetramer in solution. The retention time of 7.75 min observed for EcAGPase corresponds to an apparent molecular mass of 198.3 kDa, which agrees within the experimental error of gel filtration experiments with the calculated mass of the homotetramer of 194.8 kDa.*

The crystal structures of *Ec*AGPase were solved by molecular replacement using a tetramer of *At*AGPase (PDB: 3BRK) in two different states, including the complexes with its naturally occurring and preferred allosteric negative regulator AMP

(*Ec*AGPase•AMP•SUC) and positive regulator FBP (*Ec*AGPase•FBP; Preiss et al., 1966). In addition, the *Ec*AGPase•AMP•SUC structure displayed sucrose (SUC) located in the active site of the enzyme. *Ec*AGPase•AMP•SUC and *Ec*AGPase•FBP forms crystallized in space group P 2_1 with 16 molecules (431 residues each) in the asymmetric unit and diffracted to a maximum resolution of 2.67 and 3.04 Å, respectively. *Ec*AGPase crystallized as a homotetramer with each protomer (48.7 kDa) composed of two domains, the N-terminal glycosyltransferase A (GT-A)-like domain (residues 1–315) containing the active site, and the C-terminal regulatory domain (residues 316–431) containing a left-handed parallel β helix (LβH; residues 316–396; Figure 1A). The GT-A-like domain consists of one Rossmann fold domain (residues 1– 315; Pelissier et al., 2010). The core is composed of a central β sheet comprising seven β strands (β5, β4, β1, β8, β14, β10, and β15, of which β14 is antiparallel) flanked on both sides with several α helices (Figure 4.2). In contrast, the LβH domain (residues 316–396; Raetz and Roderick, 1995) is built of short β strands (β17-β30) oriented parallel to each other and describing a triangular prism (Figure 4.2).

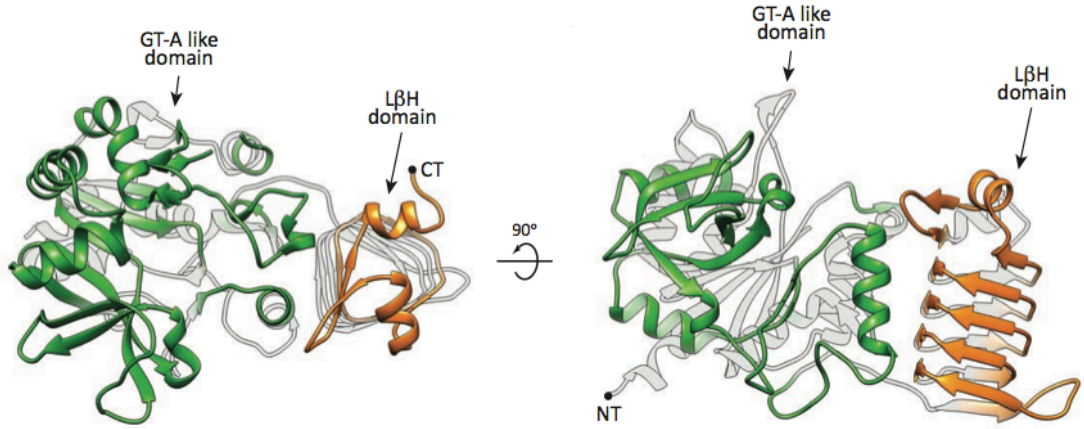


Figure 4.2. The structure of a EcAGPase monomer. *Cartoon representation of EcAGPase showing the secondary structural elements, as defined by CHIMERA. β-strands and α-helices are shown as arrows and cylinders, respectively. The N-terminal GT-A like domain consists of one Rossmann fold domain (green; residues 1-315). The core is composed of a central β-sheet comprising seven β-strands (β5, β4, β1, β8, β14, β10, β15 of which β14 is antiparallel) alternating with α-helices. The N-terminus region (residues 1 to 10) has no electron density indicating conformational flexibility. The C-terminal domain comprises the LβH (orange; residues 316 to 396), built of short β-strands (β17-30) oriented parallel to each other, and describing a triangular prism.*

The *Ec*AGPase protomers build into a physiological and functional homotetrameric structure (194.8 kDa) that can be viewed as a dimer of dimers (Figure 4.3). The most important contribution to the dimer inter-face is the triangular base of the LβH prism of each protomer (strands β17, β18, and β19) resulting in two antiparallel β sheets.

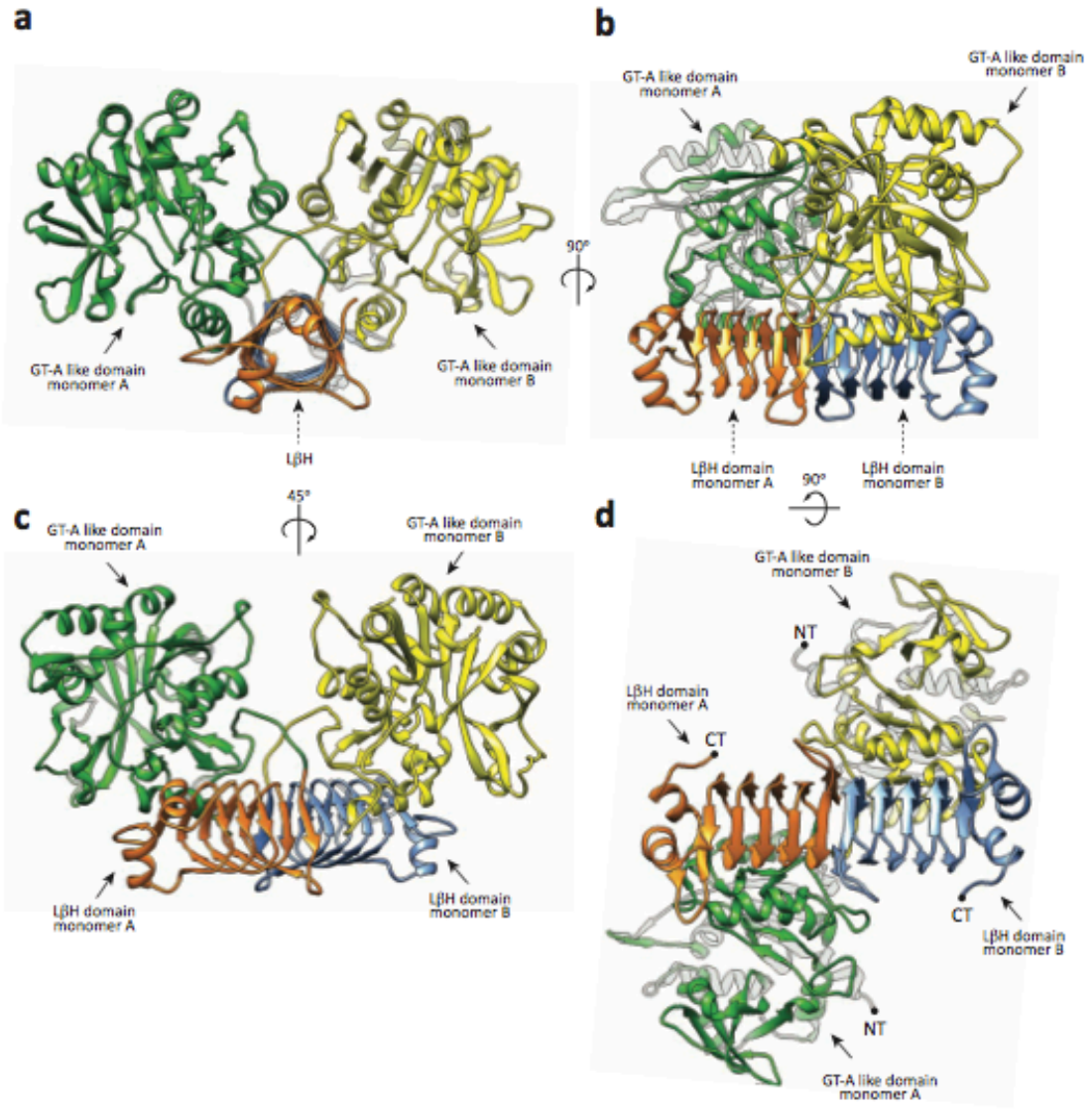


***Figure 4.3*. Structure of *EcAGPase* dimer.** *A-D. Cartoon representing the three-dimensional configuration of an EcAGPase dimer. The most important contribution to the dimer interface is the triangular base of the LβH prism of each monomer (strands β17, β18 and β19), resulting in two anti-parallel β-sheets.*

In contrast, the tetramer assembles mainly by interactions between the N-terminal GT-A-like domains from different dimers. Specifically, as depicted in Figure 4.4, three adjacent α helices, α1 (residues 10–19), α6 (residues 93–95), and α8 (residues 149– 159) from the GT-A-like domain of protomer A interact with the equivalent structural elements of the GT-A-like domain of protomer D. Moreover, α5 (residues 78–87), α7 (residues 117–131), and β5 (residues 98–103) interact with the equivalent structural elements of the GT-A-like domain of protomer C, strongly contributing to anchoring both dimers in a competent tetramer configuration. The resulting architecture allows the protomers to communicate with each other, from which cooperativity emerges.

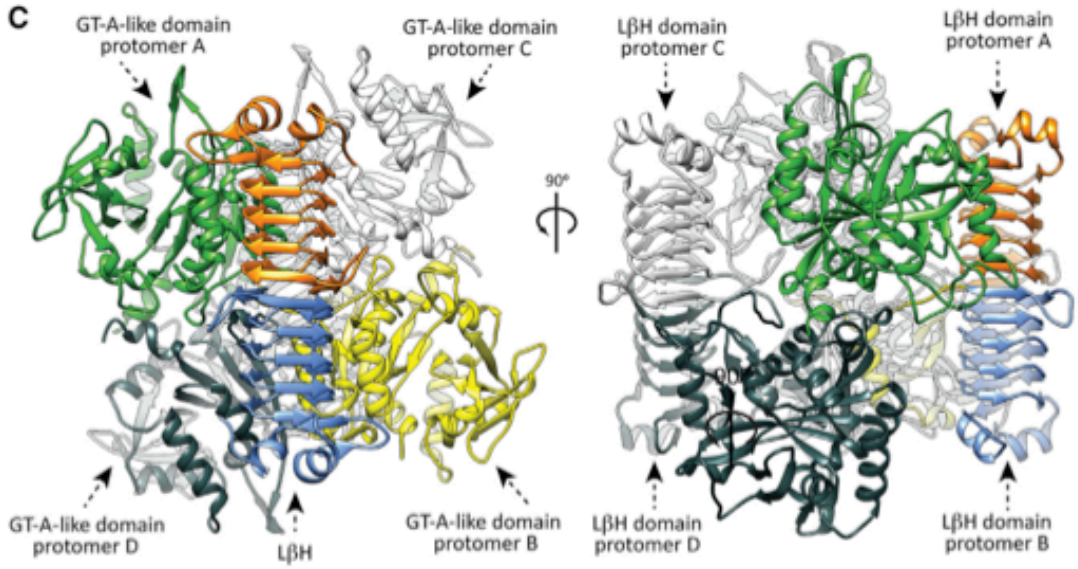


Figure 4.4. Structure of EcAGPase tetramer. *The EcAGPase monomers (48,7 kDa) build into a physiological and functional homotetrameric structure (194.8 kDa). EcAGPase is a homotetramer that can be viewed as a dimer of dimers[54], with the dimer mainly built by interactions between the LβH C-terminal domains, and the tetramer assembling mainly by interactions between the N-terminal GT-A like domains from different dimers. Specifically, three adjacent α-helices, α1 (residues 10 to 19), α6 (residues 93 to 95) and α8 (residues 149 to 159) from the GT-A like domain of monomer A (green) interact with the equivalent structural elements of the GT-A like domain of monomer D (black). Moreover α5 (residues 78 to 87), α7 (117 to 131) and β5 (98 to 103) interact with the equivalent structural elements of the GT-A like domain of monomer C (grey), strongly contributing to anchoring both dimers in a competent tetramer configuration.*

4.2.2. The active site of *Ec*AGPase

The active site is located in a deep cleft of the GT-A-like domain, as observed in other nucleotide sugar pyrophosphorylases (Cupp-Vickery et al., 2008; Jin et al., 2005). The crystal structure of *Ec*AGPase•AMP•SUC revealed the presence of SUC in the C-terminal region of the active site (residues 163– 315). SUC is clearly visible in the electron density maps and is present in all four active sites of the *Ec*AGPase homotetramer. The glucose moiety binds to a deep pocket (the ‘‘sugar binding pocket’’ according to Brito et al., 2011, defined by three β strands, β11 (residues 179–183), β12 (residues 189–194), and β13 (residues 208–212), and four loops: β12-β13 (residues 194–208), α9-β11 (residues 177–179), α9-α12 (residues 261– 280), and β8-α8 (residues 140–149). The O2 and O4 atoms of the glucose ring make hydrogen bonds with the side chain of Glu194 and the main chain carbonyl atom of Ser212, respectively. The glucose O6 is hydrogen bonded with the lateral chain of His143. Importantly, the replacement of His143, Glu194, and Ser212 per alanine displayed lower apparent affinity for G1P compared with the wild-type *Ec*AGPase (Clarisa M. Bejar et al., 2006). Several aromatic residues, including Phe178, Phe192, Tyr216, and Trp274, are important for building the walls of the cavity. Interestingly, the structural comparison of the *Ec*AGPase•AMP•SUC complex with that of the glucose-1-phosphate thymidylyltransferase from *Pseudomonas aeruginosa* (*Pa*RmlA) in complex with G1P (PDB: 1G0R; (Blankenfeldt et al., 2000) revealed that the glucose moieties superimpose very well (Figure 4.5A). According to this configuration, the phosphate group of G1P is bound to Lys195, a conserved and essential residue for the enzymatic activity (Hill, 1991). Recent studies in the human UDP-glucose pyrophosphorylase from *Leishmania major* showed that the equivalent residue Lys255 is affected during the catalytic cycle, being relevant for G1P stabilization (Führing et al., 2013). Taken into account all these experimental data, we clearly defined the location of G1P in the active site of *Ec*AGPase. The ATP binding site is located in the N-terminal region of the GT-A-like domain (residues 20–162; Figure 4.5B). The crystal structures of *Ec*AGPase•AMP•SUC and *Ec*AGPase•FBP revealed the presence of PO_4 and SO_4 ions, respectively, in the ATP binding site. Both anions are bound in equivalent positions, making strong interactions with the lateral and main chains of Arg32 and the main chain of Thr31. The structural comparison of both *Ec*AGPase complexes with that of the N-acetylglucosamine-1-phosphate uridyltransferase from Mycobacterium tuberculosis (*Mt*GlmU) in complex with ATP (PDB: 4K6R; Vithani, Bais, and Prakash 2014) and the GDP-Man

pyrophosphorylase from *Thermotoga maritima* in complex with GTP (PDB: 2X60; Pelissier et al. 2010) revealed that the anions superimpose with the ATP γ-PO_4. According to this configuration, ATP accommodates into the active site of *Ec*AGPase, in close contact to the essential catalytic Lys42 and favorably positioned to receive the G1P (Figure 4.5B; Ballicora et al. 2005; Ballicora et al. 2007).

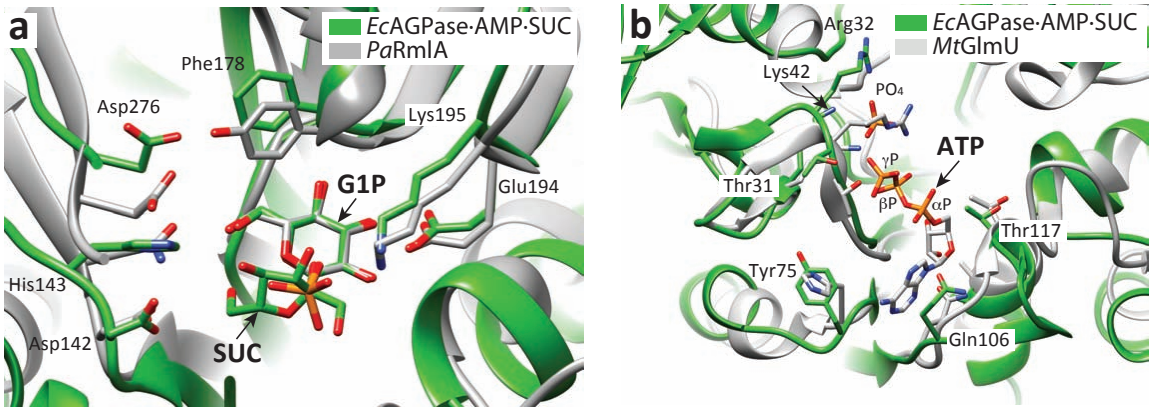


Figure 4.5. The ATP and G1P binding sites in EcAGPase. *(A) Structural superposition of EcAGPase•AMP•SUC complex with that of RmlA in complex with G1P (PDB: 1G0R) defining the position of G1P, and showing key interactions of SUC with selected residues of EcAGPase.(B) Structural superposition of EcAGPase•AMP•SUC complex with that of the GlmU in complex with ATP (PDB: 4K6R) defining the entry position of ATP.*

Based on the structural information reported here and the biochemical data regarding the involvement of these residues in *Ec*AGPase catalytic mechanism, we proposed a model in which Lys42 and Lys195 polarize G1P and ATP anionic phosphorous groups. This polarization increases the nucleophilic nature of the sugar phosphate oxygen attacking the ATP α-phosphate group, leading to the liberation of PPi (Figure 4.6A). The condetation reaction is held in the presence of cation Mg^{2+} (magnesium-dependent activity), which minimizes the charge repulsion between anionic phosphorous groups and has been proved to induce nucleophile activation in other nucleotidiltransferases (Swift, Ong, & Amaro, 2012). In nature, the reaction is pull forward by the hydrolysis of PPi by inorganic pyrophosphatases, since the reaction have a ΔG~0, otherwise the reverse reaction could also occur (Figure 4.6B).

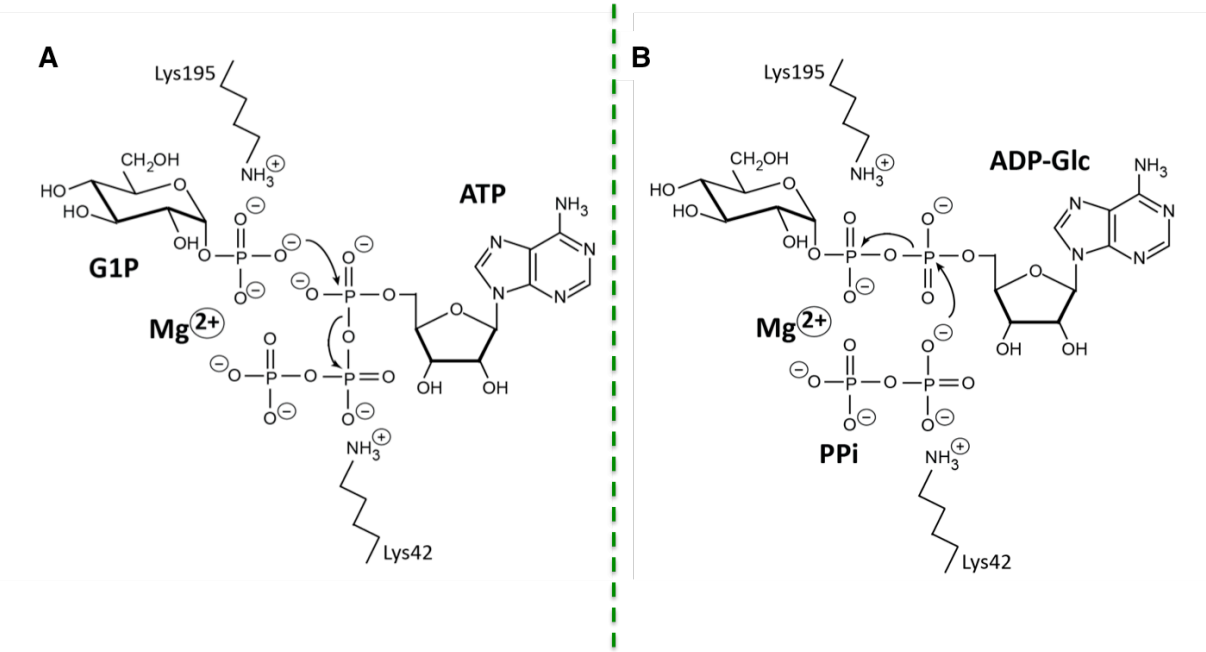


Figure 4.6. Proposed catalytic mechanism for EcAGPase. *(A) The formation of the nucleotide activated donor ADP-Glc occurs as a condensation reaction between ATP and a G1P. (B) Reverse reaction leading to the formation of G1P.*

4.2.3. The AMP allosteric binding site

The identification of the physiological positive and negative regulatory sites at the molecular level in AGPases has been a long-standing question and a matter of intense research in the field of glycogen/starch biosynthesis/regulation. In the *Ec*AGPase•AMP•SUC crystal structure, AMP is clearly visible in the electron density maps and is present in all four allosteric sites located in the corresponding clefts between the N-terminal GT-A-like and C-terminal LβH domains of neighboring protomers from different dimers (Figures 4.7). Specifically, AMP is deeply buried into a cleft mainly defined by (i) the N-terminal β2-β3 hairpin (residues 46–52), α5, and the connecting loop α2-α3 (residues 37–42), (ii) the C-terminal α15 (residues 419–425) and the connecting loops β28-β29 (residues 384–388) and β25-β26 (residues 367–371), and (iii) the N-terminal α7 from a neighboring protomer. The α-PO4 group occupies a cavity rich in positively charged residues including Arg40 (α3), His46 and Arg52 (β2-β3 hairpin),

Thr79 (α5), and Arg386 (LβH; Figure 4.7A). The adenine heterocycle is stabilized by a strong stacking interaction with Arg130 (α7) from the GT-A-like domain of the neighboring protomer and van der Waals interactions with Arg419 (α15) and Arg386 from the LβH domain. The side chain of Glu420 (α15) forms an important salt bridge with Lys39, communicating the α2-α3 loop with α15. In addition, a strong hydrogen-bonding interaction of the adenine N6 nitrogen with the side-chain carboxylate group of Glu270 (LβH domain) may account for the nucleotide specificity. Finally, the D-ribose O2 atom makes a hydrogen bond with the side chain of Arg130 being also at a van der Waals distance from the Lys39 side chain (Figure 4.7B)

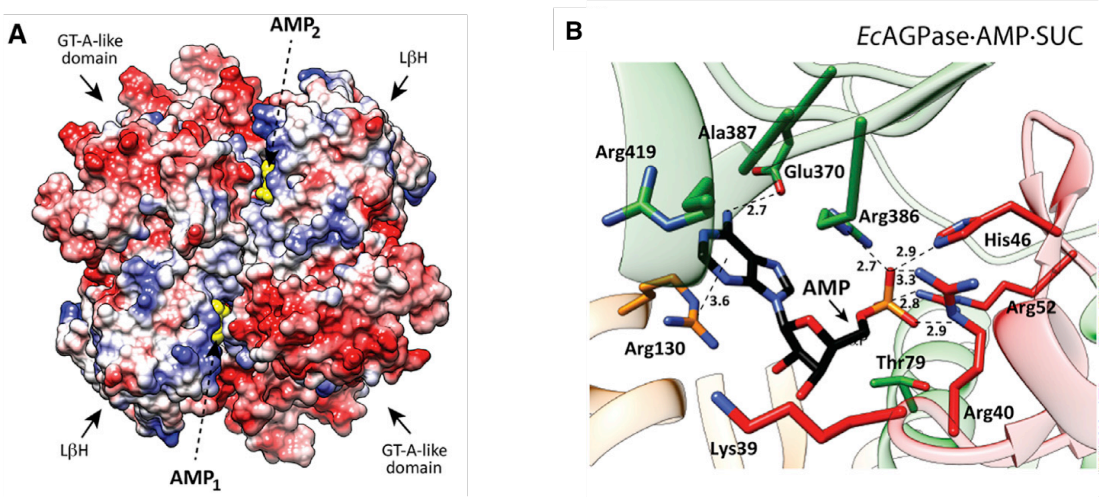


Figure 4.7. Location of the AMP binding site. *(A) Surface representation of the EcAGPase tetramer in complex with AMP (yellow spheres) in which each individual monomer displays it surface colored by coulombic potential, revealing the positively charged nature of the allosteric site. (B) Close-up view of the AMP binding site, showing key interactions with selected residues.*

Interestingly, a crosstalk event between protomers from the same and different dimers suggests that AMP interactions might lead to the stabilization of the quaternary structure of *Ec*AGPase in solution. Supporting this notion, thermal unfolding followed by the far-UV circular dichroism (CD) signal at 222 nm indicated important differences in protein stability between the *apo Ec*AGPase and the *Ec*AGPase•AMP complex. The apparent melting temperatures (T_m) of *Ec*AGPase and *Ec*AGPase•AMP were 71.2 ºC and 75.8 ºC, respectively, indicating that the AMP-complexed form is approximately 4.6 ºC more stable than the unliganded form (Figure 4.8).

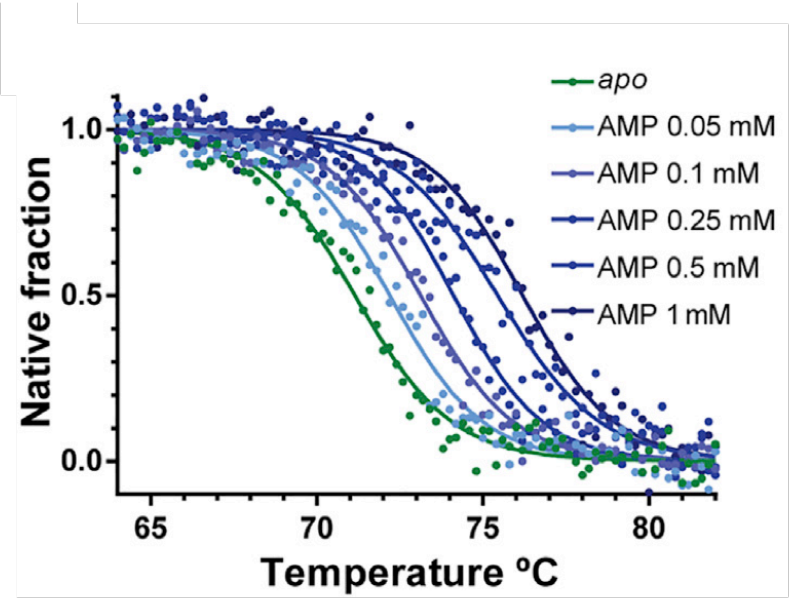


Figure 4.8. EcAGPase thermal unfolding transitions recorded at 222 between 20 °C and 90 °C. *Native fraction of EcAGPase plotted versus temperature for the apo state (green), and EcAGPase at different concentrations of AMP (in a blue scale), The corresponding fitted two-state sigmoidal curves of the unfolding events are also shown.*

4.2.4. The FBP binding site: Partial overlapping with the AMP binding site

The EcAGPase•FBP crystal structure reveals that FBP binds into the same cleft as AMP, but is located in a more solvent-exposed environment, with no evident interactions with neighboring protomers (Figures 4.9A). The FBP binding site comprises the last C-terminal residues of the enzyme (residues 420–431), with the FBP making important interactions with positively charged residues located in one side of α15 (Figure 4.9B). Supporting the relevance of the C terminus in the recognition of FBP, a protein chimera containing the N terminus (271 residues) of *At*AGPase, activated by fructose 6-phosphate and pyruvate, and the C terminus (153 residues) of *Ec*AGPase, retained the selectivity for FBP (M. A. Ballicora et al., 2002). Moreover, a variant of *Ec*AGPase in which the two last C-terminal residues, Glu430 and Arg431, were removed, became less sensitive to FBP activation (Wu & Preiss, 2001).

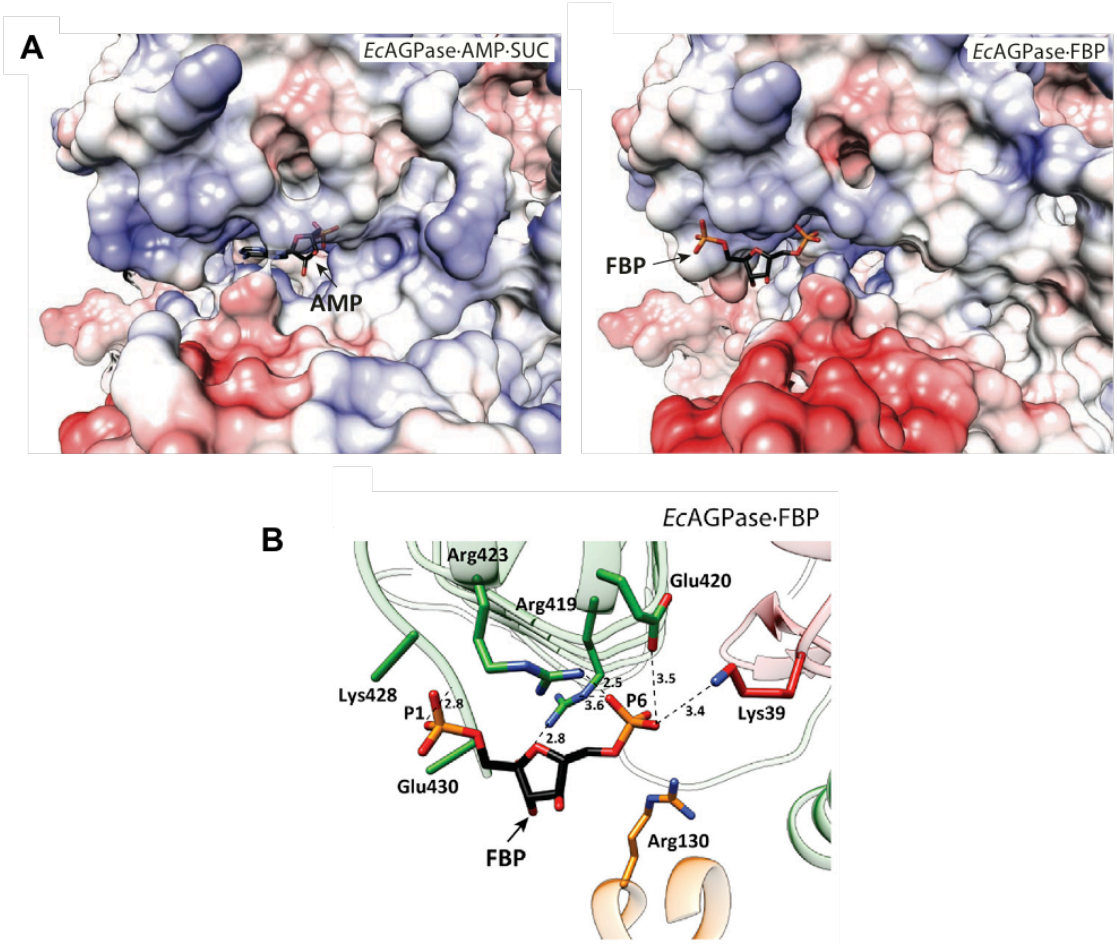


Figure 4.9. Location of the FBP binding site. *(A) FBP binds into the same cleft as AMP, but is located in a more solvent-exposed environment. (B) Close-up view of the FBP binding site, showing key interactions with selected residues.*

The structural comparison of the EcAGPase•AMP•SUC and EcAGPase•FBP crystal structures revealed that the AMP and FBP binding sites partially overlap. FBP binding promotes important local conformational changes in the allosteric site when compared with the AMP complex (Figure 4.10). Specifically, the Lys39 side-chain coordinates the O1 atom of the FBP PO4 group at position 6, whereas the side chain of Glu420 makes a hydrogen bond with the O3 of the fructose ring. Biochemical studies demonstrated the important role of Lys39 in the binding and the mechanism of activation of *Ec*AGPase by FBP (Gardiol & Preiss, 1990). Interestingly, Lys39 showed protection to the covalent modification with pyridoxal-PO_4 (PLP), by reduction with $NaBH_4$, in the presence of FBP (Parsons & Preiss, 1978b). The modification of Lys39 with PLP resulted in an enzyme with a permanently enhanced activity, even in the absence of FBP. This result suggests that the Schiff base formed between Lys39 and PLP might result in binding of the PLP phosphate group to the allosteric site, mimicking FBP binding and

contributing to the permanent locking of the enzyme in the activated state. In addition, the side chain of Arg423 positions its guanidinium group in close contact with the O5 atom of the FBP phosphate group at position 6, the O6 atom of which makes a strong hydrogen bond with the main chain of Gln429. This structural configuration allows the fructose to be positioned in close proximity to α15, allowing the side chain of Arg419 to make an important hydrogen bond with the sugar ring O2 atom. Moreover, the side chain of Arg130 from the neighboring protomer completely changes its conformation and the last two residues, Glu430 and Arg431, become structured in the *Ec*AGPase•FBP complex.

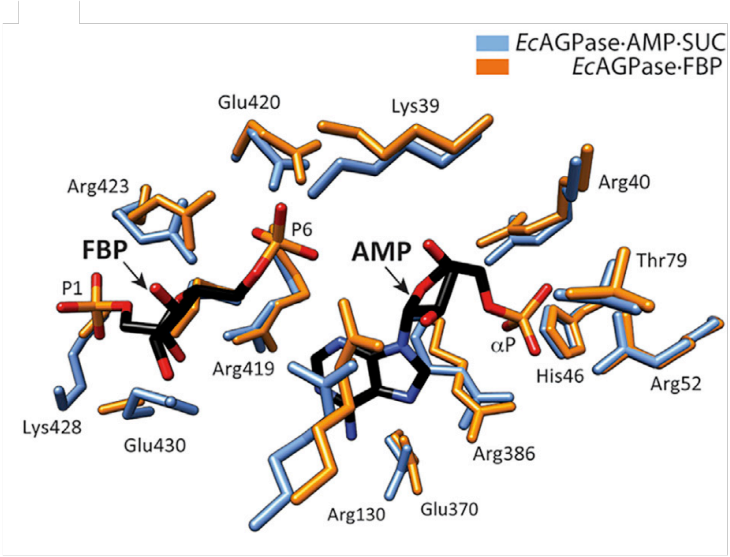


Figure 4.10. Structural superposition of the EcAGPase•AMP•SUC and EcAGPase•FBP, showing the partial overlapping of AMP and FBP binding sites.

The prominent conformational change of Arg130 side chain suggests that FBP interactions might not lead to the stabilization of the quaternary structure of *Ec*AGPase in solution. The T_m value of *Ec*AGPase•FBP was 72.0 ºC, indicating the formation of a less-stable complex than that observed for *Ec*AGPase•AMP (Figue 4.11A). Moreover, the addition of FBP to the *Ec*AGPase•AMP complex triggered a clear reduction in the *Tm* values as revealed by the CD experimental data, indicating that FBP not only is able to compete with AMP but also to modify the structural arrangement of the *Ec*AGPase•AMP complex, leading to the occurrence of a less-stable structure (Figure 4.11B). The structural configuration of the *Ec*AGPase regulatory site, in which the AMP and FBP binding sites partially overlap, accounts for the fact that sensitivity to inhibition by AMP is modulated by the concentration of the activator FBP (J Preiss, 1978). In

addition, the experimental data indicate that the *Ec*AGPase•FBP complex is markedly less stable and more flexible/dynamic than the *Ec*AGPase•AMP complex (Figueroa et al., 2011).

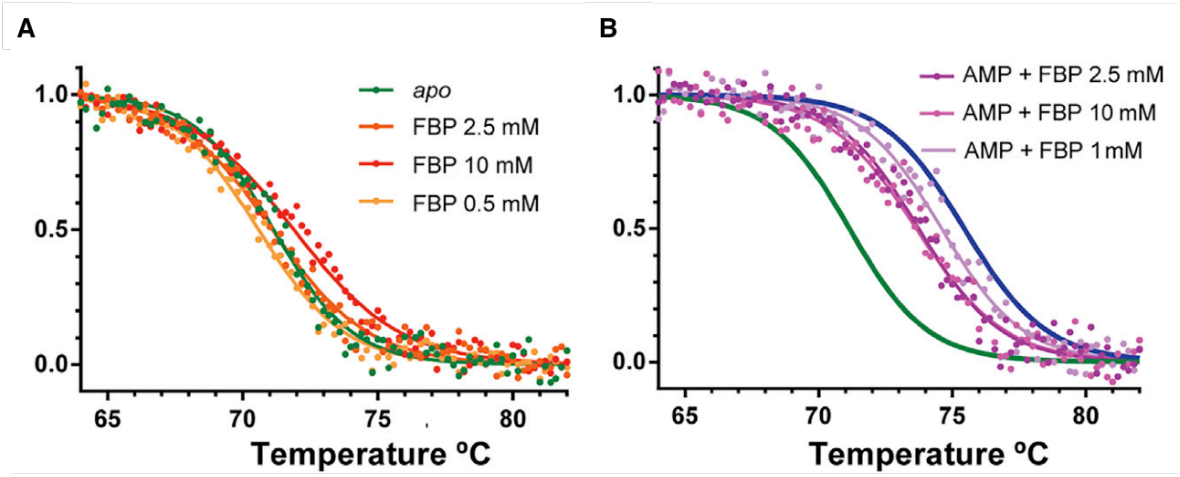


Figure 4.11. EcAGPase thermal unfolding transitions recorded at 222 between 20 °C and 90 °C. *Native fraction of EcAGPase plotted versus temperature for the apo state (green), FBP (A, in orange and red scale), or AMP and FBP (B, in a purple scale). In (B), the curve for EcAGPase•AMP (where concentration of AMP = 0.5 mM) is indicated for reference in blue. The corresponding fitted two-state sigmoidal curves of the unfolding events are also shown.*

4.2.5. A model for *Ec*AGPase allosteric regulation

Close inspection of the *Ec*AGPase•AMP•SUC and *Ec*AGPase•FBP crystal structures revealed how the allosteric and active sites are connected each other. The AMP allosteric site communicates with the active site (Figure 4.12A) of the same protomer through a region comprised of 27 residues, which we have defined as the sensory motif (SM), located between β1 and α4 (residues 26–52; Figure 4.12B). The SM is constituted by (i) the nucleotide-binding loop (NBL) (residues 26–33) including the GGxGxR consensus sequence involved in ATP binding; followed by (ii) a segment rich in short secondary structure elements (residues 34–52) including α2 (residues 34–37), α3 (residues 42–44), β2 (residues 46–47), and β3 (residues 51–52), the last arranged in the form of a β hairpin (residues 46–52). In addition to the NBL, two side-chain residues of the SM face the active site playing a prominent role in ATP recognition and catalysis:

Arg32 (α2) interacts with the γ-PO4, whereas Lys42 (α3) is proposed to participate in the electrostatic stabilization of the transition state (Figure 4.5; Führing et al. 2013; Ballicora et al. 2005). In contrast, Arg40 (α3), His46, and Arg52 (β2-β3 hairpin) face the AMP allosteric site, strongly interacting with α-PO4. The flexible β4-α5 loop (residues 73–77; regulatory loop 1; RL1 thereafter) interacts with both the NBL and the segment rich in secondary structure elements located on the same protomer, likely modulating their conformations. This loop also connects with α helix 5 (residues 78–87), of which Thr79 interacts with the α-PO4 of AMP. Strikingly, the AMP binding site not only connects with the active site by intra-protomer interactions, but also through inter-protomer crosstalk. The adenine heterocycle of AMP makes an important stacking interaction with the side chain of Arg130 from a neighboring protomer of a different dimer. This important residue is located inside α7 (residues 117–131) and further communicates with a long loop (residues 104–116; regulatory loop 2; RL2 thereafter) that flanks the ATP binding pocket in the active site (Figure 4.5). In addition, the β2-β3 hairpin directly interacts with the loop connecting the N- and C-terminal domains (residues 292–315) of the neighboring protomer of the same dimer. In contrast, the FBP allosteric site communicates with the active site mainly through the SM of the same protomer, involving a key interaction of Lys39 with the O1 atom of the FBP phosphate group at position 6. This interaction directly modulates the conformation of the catalytic Lys42 located in the same loop, which is essential for the reaction to take place (Figure 4.6; Ballicora et al. 2007). We propose a model in which the positive and negative energy reporters regulate AGPase catalytic activity via intra- and interprotomer crosstalk, with the SM and two critical regulatory loops RL1 and RL2 flanking the ATP binding site playing a prominent role.

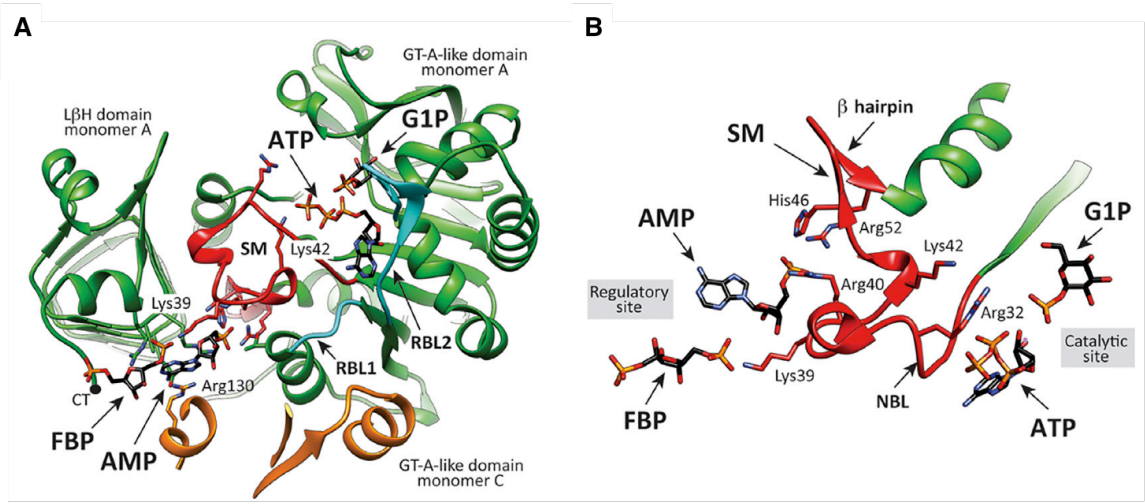


Figure 4.12. Cartoon representing the key structural elements involved in EcAGPase allosteric regulation*. Protomers A and C of the complex EcAGPase•AMP•SUC are shown in green and orange, respectively. AMP and the superimposed FBP molecules are shown in the allosteric site. The SM is shown in red, whereas the RBL1 and RBL2 loops are shown in cyan. G1P and ATP are shown in the active site. (B) Close view of the SM.*

4.2.6. The allosteric sites are essentially preserved in AGPases

The crystal structure of *At*AGPase has been solved in the presence of SO_4 (PDB: 3BRK; Cupp-Vickery et al., 2008). The *Ec*AGPase primary sequence shares a 55% identity with *At*AGPase. The overall fold between *Ec*AGPase and *At*AGPase is essentially preserved with (i) a root-mean-square deviation (RMSD) of 2.77 A˚ for the monomer and (ii) an RMSD of 2.92 A˚ for the tetramer. Importantly, multiple amino acid sequence alignments among the bacterial AGPase family, weighted by structural alignment of *Ec*AGPase and *At*AGPase, strongly support a common mechanism for the regulation of the enzymatic activity (Figure 4.13). The positively charged residues Arg40, His46, Arg52, and Arg386, involved in the binding of the α-PO4 moiety of the negative regulator AMP, are highly conserved. Interestingly, close inspection of the *At*AGPase crystal structure shows that SO_4 superimposes well with the α-PO4 of AMP in *Ec*AGPase (Cupp-Vickery et al., 2008). In addition, Thr79 and Arg130, involved in the nucleoside ring binding of AMP, are mostly conserved within the enterobacteria family. The C-terminal region 419RxMLRKLxxKQER431, involved in FBP binding, and the key residue Lys39 are also

conserved among enterobacteria AGPases that use FBP as a positive regulator. Importantly, critical residues that participate in the SM and RL1 and RL2 loops are also preserved.

*St*AGPase is composed of two α and two β subunits, also referred to as small and large subunits, respectively, to form an α2β2 heterotetramer. The α subunit of AGPases is highly conserved in higher plants (85%–95% identity), whereas the β subunit is less conserved (50%–60% identity). In the *St*AGPase, the α and β subunits share 53% identity (Jin et al., 2005). Importantly, the two subunits have different functions: α is the catalytic subunit, whereas β is the regulatory subunit. The crystal structure of a non-physiological, truncated recombinant homotetrameric version of the small subunit (α4) of *St*AGPase was solved in the presence of (i) SO_4 (PDB: 1YP2), (ii) ATP (PDB: 1YP3), and (iii) ADP-Glc (PDB: 1YP3; Jin et al. 2005). The *Ec*AGPase primary sequence shares a 31% identity with *St*AGPase. Although the overall fold between *Ec*AGPase and *St*AGPase is preserved, with (i) an RMSD of 4.35 Â for the monomer and (ii) an RMSD of 5.20 Â for the tetramer, clear differences can be found both in the GT-A-like and the LβH domains, as revealed by the structurally weighted alignment (Figure 4.13). The structural comparison of *Ec*AGPase with *St*AGPase, an enzyme that is negatively regulated by Pi, shows that SO_4 binds to equivalent residues Arg40 and Arg52 and to a lysine occupying an equivalent position to Arg486 (Jin et al., 2005). Thus, the positively charged pocket responsible for the binding of (i) the AMP α-PO_4 in bacterial AGPases and (ii) the Pi in plant AGPases seems to be conserved in both families, being essential for the negative regulation of most AGPases. Interestingly, Lys39 was also observed in several plant AGPases that use 3-phosphoglyceric acid (3PGA) as a positive regulator, suggesting that the PO_4 groups might be coordinated in a similar manner to FBP in *Ec*AGPase.

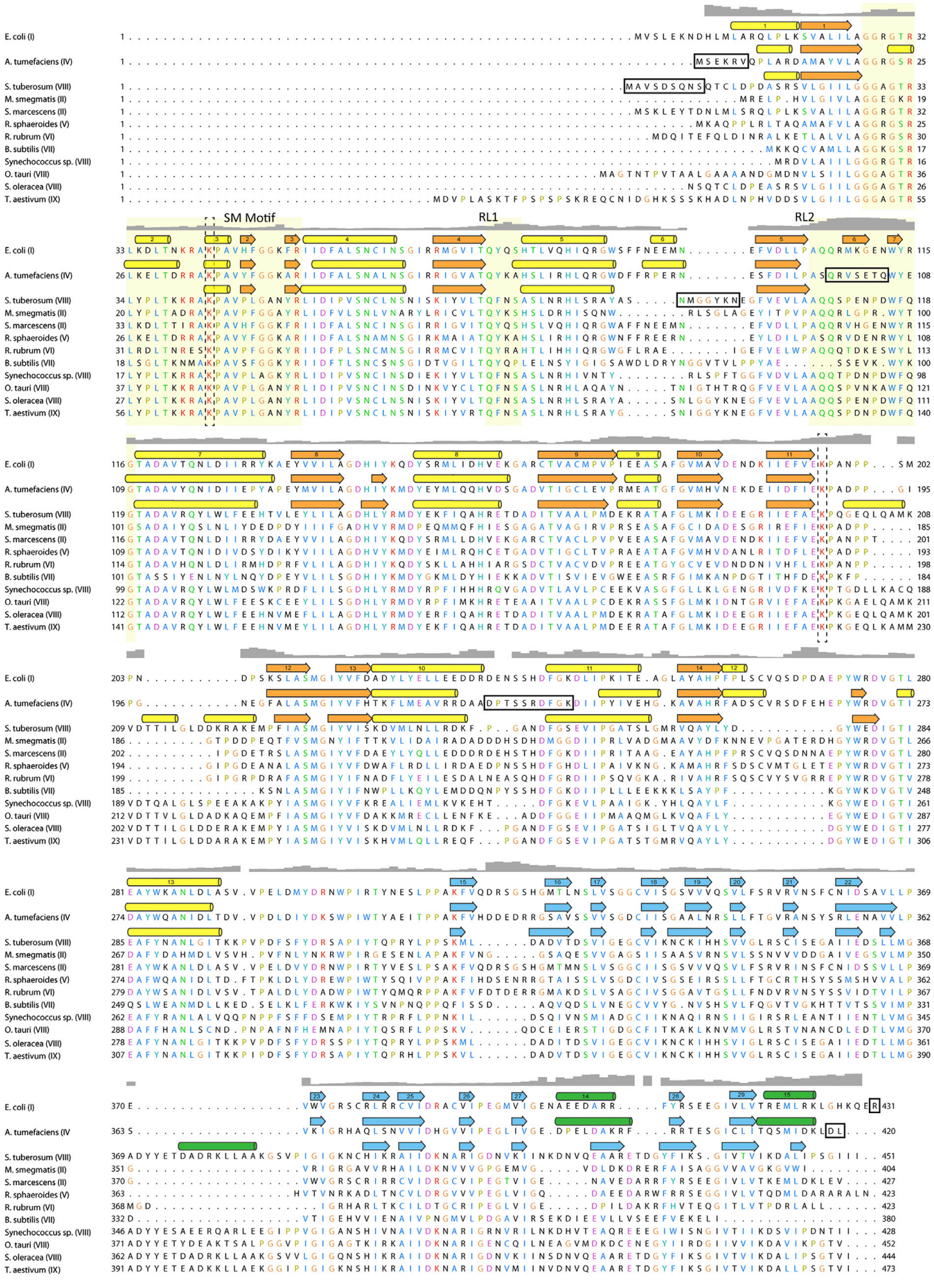

SM Motif
RL1
RL2
E. coli (I)
A. tumefaciens (IV)
S. tuberosum (VIII)
M. smegmatis (II)
S. marcescens (II)
R. sphaeroides (V)
R. rubrum (VI)
B. subtilis (VII)
Synechococcus sp. (VIII)
O. tauri (VIII)
S. oleracea (VIII)
T. aestivum (IX)

Figure 4.13. Structure-Weighted Sequence Alignment of EcAGPase with other AGPases. *Structural alignment between the crystal structures of EcAGPase (PDB: 5L6V; UniProt: P0A6V1), AtAGPase (PDB: 3BRK; UniProt: P39669), and StAGPase (PDB: 1YP3; UniProt: P23509). The secondary structure elements corresponding to the GT-A-like domain are shown in yellow (α-helices) and orange (b helices); and to the LbH domain in green (a helices) and blue (b helices). Residues with poor electron density are highlighted as full boxes. The SM and the RL1 and RL2 loops are highlighted in yellow. Catalytic residues are highlighted as dotted boxes. The RMSD value is shown for each residue. Amino acid sequences of selected AGPases were aligned to the structure alignment: Mycobacterium smegmatis (class II; UniProt: A0R2E1), Serratia marcescens (class II; UniProt: A0A0U6P844), Rhodo- bacter sphaeroides (class V; UniProt: Q9RNH7), Rhodospirillium rubrum (class VI; UniProt: Q9ZFN4), Bacillus subtilis (class VII; UniProt: P39122), Synechococcus sp. (class VIII; UniProt: Q2JU94), Ostreococcus tauri (class VIII; UniProt: Q6PYZ7), Spinacia oleracea (class VIII; UniProt: Q43152), and Triticum aestivum (class IX; UniProt: P30523)*

MECHANISTIC INSIGHTS INTO *Ec*AGPase ALLOSTERIC REGULATION

5. Mechanistic insights into *Ec*AGPase allosteric regulation

As it has been shown in the previous chapter, the report of *Ec*AGPase crystal structures in complex with its physiologically preferred allosteric regulators, AMP and FBP, allowed us to identify the allosteric binding sites and propose a model of regulation for this paradigmatic enzyme. In this regard, we define four common allosteric clefts between the GT-A-like and LβH domains of neighboring protomers, in which both allosteric modulators bind to partially overlapping sites. This structural configuration of the *Ec*AGPase regulatory site accounts for the fact that sensitivity to inhibition by AMP is modulated by the concentration of the activator FBP (Gardiol & Preiss, 1990; Gentner & Preiss, 1967). Specifically, each allosteric cleft is communicated with the corresponding active site of the same protomer through a region defined as the 'Sensory Motif' (SM), a complex structural element constituted by the nucleotide-binding loop NBL, including a G-rich motif involved in ATP binding, and a segment rich in short secondary structure elements. Based on this structural information, we proposed a model, in which the binding of the positive and negative energy reporters regulates *Ec*AGPase catalytic activity through the SM and two critical regulatory loops RL1 and RL2 flanking the active binding site, via intra-protomer interactions and inter- protomer crosstalk.

Considering the important structural information already available, the next step was to attempt to validate the allosteric binding sites and regulatory model proposed for *Ec*AGPase. For this aim, we carefully explore the consequences of single point mutations in the allosteric cleft of *Ec*AGPase, in key residues involved in FBP and AMP binding, looking for changes in the allosteric properties, stabilization and their impact in enzymatic activity. Finally, we explore how those mutations impact in the yield of glycogen in the physiological environment of the host.

The references of this chapter are based on the publication, mentioned in "Publications" section, *Mechanistic insights into the allosteric regulation of bacterial ADP-glucose pyrophosphorylases* (Comino et al., 2017).

5.1. Materials and methods

Materials — The pET22b:*Ec*AGPase•K39A, pET22b:*Ec*AGPase•R40A, pET22b:*Ec*AGPase•R40E, pET22b:*Ec*AGPase•K42A, pET22b:*Ec*AGPase•H46A, pET22b:*Ec*AGPase•R130A, pET22b:*Ec*AGPase•K195A, pET22b:*Ec*AGPase•R386A, pET22b:*Ec*AGPase•R419A and pET22b:*Ec*AGPase•R423A plasmids carrying out the *Ec*AGPase mutants were synthesized/sequenced by ATG:biosynthetics (Merzhausen, Germany) using the pET22b:*Ec*AGPase clone as template (please see *Ec*AGPase cloning section), and further expressed and purified to apparent homogeneity as described for the recombinant *Ec*AGPase enzyme.

EcAGPase cloning, expression and purification — The full-length *glgC* gene from *E. coli* BL21 was amplified by standard PCR using oligonucleotide primers *glgC*_NdeI_Fwd (5´GGGAATTCCATATGGTTAGTT 3´) and *glgC*_XhoI_Rev (5´ CCGCTCGAGTCA TCGCTCCTG 3´), Phusion DNA Polymerase (New England Biolabs) and purified genomic DNA as template. The PCR fragment was digested with NdeI and XhoI and purified by agarose gel electrophoresis. The fragment was ligated to the expression vector pET22b (Novagen) using T4 DNA ligase, generating pET22b:*Ec*AGPase. The recombinant EcAGPase has no additional amino acids when compared to the native enzyme. *E. coli* BL21(DE3) cells transformed with pET22b:*Ec*AGPase or the corresponding mutants were grown in 2,000 ml of LB medium (5 g yeast extract, 10 g peptone tryptone, 10 g NaCl) supplemented with 100 µg/ml carbenicillin at 37°C. When the culture reached an OD^{600} of 0.8, *Ec*AGPase expression was induced by the addition of 1 mM isopropyl β-thiogalactopyranoside (IPTG) and further incubated at 18°C for 20 h. Cells were harvested by centrifugation at 5,000 x g and resuspended in 40 ml of 50 mM Hepes pH 8.0, 5 mM $MgCl_2$, 0.1 mM EDTA, 10% sucrose (w/v; solution A), containing protease inhibitors (Complete EDTA-free; Roche) and 10 mg/l of lysozyme (Sigma). Cells were then disrupted by sonication (24 cycles of 10 s each) and centrifuged for 20 min at 20,000 x g. The supernatant was dialyzed twice against solution A by using a 100,000 Da molecular mass cutoff dialysis membrane. The solution was then applied to a Q Sepharose packed column (30 ml; GE Healthcare) equilibrated with solution A. Elution was performed with a linear 0-0.5M NaCl gradient in 100 ml. Enzymatically active fractions were pooled and incubated with 1.2 M ammonium sulphate final concentration (solution B). The resultant suspension was centrifuged for 20 min at 32,000 x g and the supernatant applied into a Phenyl Shodex HIC PH- 814 equilibrated in solution B. The

enzyme was eluted with a linear gradient of 100% solution B to 100% solution A in 50 ml. The protein was dialyzed against solution A with a 100-kDa-molecular mass cutoff overnight, and stored at -80°C. All the mutants were purified following the same protocol.

Thermal unfolding — Thermal unfolding transitions were recorded on a J-815 CD spectropolarimeter (Jasco Corp., Tokyo, Japan) equipped with a water-cooled Peltier unit. Measurements were carried out at 222 nm by using Hellma 110-QS quartz cuvettes with a 1-mm optical path length. Samples were 2 µM EcAGPase or EcAGPase mutants in 50 mM Tris-HCl pH 8.0, 100 mM NaCl. Thermal dependences of the ellipticity were monitored in a range from 30 to 90°C at 222 nm. Temperature was increased stepwise by 1°C/min. Ligands effects were assessed in the same conditions for the following concentrations: AMP 0.25 mM, 0.5 mM and 1 mM; FBP 0.5 mM and 2.5 mM. Transitions were normalized and tentatively fitted according to equation 1 to obtain the apparent melting temperature (T_m), (Sotomayor-Pérez et al., 2013).

$$y = \frac{\left(y_f + m_f \cdot T\right) + \left(y_u + m_u \cdot T\right) \cdot \mathrm{e}^{\left(\frac{\Delta Hm}{R}\right)\cdot\left(\frac{1}{Tm}-\frac{1}{T}\right)}}{1 + \mathrm{e}^{\left(\frac{\Delta Hm}{R}\right)\cdot\left(\frac{1}{Tm}-\frac{1}{T}\right)}}$$

(Eq. 1)

where y represents the observed CD signal at 222 nm, y_f and y_u are the y-axis intercepts and mf and m_u the slopes of the pre- and post-transition baselines, respectively, T is the temperature in K, T_m is the melting temperature, and ΔHm is the enthalpy change of unfolding at Tm. Curve fitting was performed with KaleidaGraph 4.5.2 (Synergy Software).

EcAGPase enzymatic assay — The enzymatic activity of *Ec*AGPase and *Ec*AGPase mutants was monitored using a micro plate colorimetric end-point malachite green phosphate assay (Fusari, Demonte, Figueroa, Aleanzi, & Iglesias, 2006). *Ec*AGPase catalyzes the reaction of ATP and G1P to produce ADP-Glc and PPi. This reaction is coupled with an inorganic pyrophosphatase (PPase), which catalyzes the hydrolysis of PPi to orthophosphate (Pi). Pi dosage is determined by the formation of a phosphomolybdate- malachite green complex:

1) Enzymatic coupled reaction

α-Glucose-1P + ATP → ADP-glucose + $P_2O_7^{4-}$ (Eq. 2)

$P_2O_7^{4-}$ + H2O → 2 PO_4^{3-} + $2H^+$ (Eq. 3)

2) Stop reaction by complexation of Mg^{2+} with EDTA

3) Molybdate/Malachite Green (MG) reaction

$H_3PO_4 + 12\ H_2MoO_4 \rightarrow H_3PO_4(MoO_3)_{12} + 12\ H_2O$ (Eq. 4)

$H_3PMo_{12}O_{40} + HMG^{2+} \rightarrow [MG^+](H_3PO_4(MoO_3)_{12}) + 2H^+$ (Eq. 5)

4) Addition of sodium citrate for reaction stabilization

Specifically, samples contained 8 μg/ml of *Ec*AGPase and 1 μg/ml of PPase (Sigma), 0.5 mM ATP, 0.5 mM G1P, 50 mM Tris-HCl pH 7.5, 2 mM $MgCl_2$ and 100 mM NaCl, with or without the addition of the allosteric regulators AMP, FBP or AMP/FBP at 0.5 mM, in a final volume of 60 μl. Reactions were incubated at 20°C and stopped at given time by the addition of 10 μl of 50 mM EDTA pH 8.0. 20 μl of molybdate solution (42 g/L ammonium molybdate tetrahydrate in 4N HCl; Fluka) were added to the reaction mixture, and incubated at 20°C for 3 min. Then, 60 μl of Malachite Green solution (0.52 g/L malachite green oxalate salt in water; Sigma) were added to the mixture and incubated at 25°C for 5 min. Finally, the solution was stabilized by the addition of 60 μl of sodium citrate (50 g/L in water; Sigma), and measured at 620 nm in a Spectra Max M2 plate reader. Measurements were performed in quintuplicates. *Ec*AGPase and *Ec*AGPase mutants were stored in 1 mg/ml aliquots in 50 mM Tris-Cl pH 7.5 and 100 mM NaCl at -80 °C for single use. Specific activities were stable during a period of 6 months.

Overexpression of EcAGPase mutants in a glgC knock-out E. coli strain — *E. coli* strain K-12 carrying out a *glgC* gene deletion (*E. coli* K-12 Δ*glgC*; kanamycin resistant strain; Keio collection; Dharmacon GE) was transformed with pTARA plasmid (chloramphenicol resistant plasmid; Wycuff and Matthews 2000; Addgene), to provide a controlled expression of T7 RNA polymerase. The resulting strain was subsequently transformed with pET22b:*Ec*AGPase, pET22b:*Ec*AGPase•K39A, pET22b:*Ec*AGPase•R40A, pET22b:*Ec*AGPase•R40E, pET22b:*Ec*AGPase•K42A, pET22b:*Ec*AGPase•H46A, pET22b:*Ec*AGPase•R130A, pET22b:*Ec*AGPase•K195A,

pET22b:*Ec*AGPase•R386A, pET22b:*Ec*AGPase•R419A or pET22b:*Ec*AGPase•R423A mutants. *E. coli* K-12 Δ*glgC* cells transformed with pTARA and pET22b:*Ec*AGPase or the corresponding mutants were further synchronized in 10 ml of LB medium supplemented with 34 µg/ml chloramphenicol, 25 µg/ml kanamycin and 100 µg/ml carbenicillin at 30°C (Richard H Baltz, Julian Davies, 2010). After two passages, bacterial cultures were diluted 1:50. 900 µl aliquots were removed every hour and kept on ice. After 4hs, protein expression was induced by the addition of 1 mM isopropyl β-thiogalactopyranoside (IPTG) and 100 µg/ml arabinose. 900 µl aliquots were removed every hour and kept on ice for subsequent steps.

Glycogen extraction — Glycogen was extracted as previously described with minor modifications (J Preiss, Greenberg, & Sabraw, 1975). 400 µl culture aliquots of the *E. coli* K-12 Δ*glgC* strain transformed with pTARA and pET22b:*Ec*AGPase or the corresponding mutants, were centrifuged at 1200 x g for 10 min, and the supernatant discarded. Pellets were resuspended in 100 µl of 50% KOH, and further incubated at 95°C during 30 min. 300 µl of ice cold 95% ethanol was added to the the sample and centrifuged 1200 x g for 30 min. The pellet was air-dried stored at -20°C.

Glycogen measurement — The production of glycogen in the *E. coli* K-12 Δ*glgC* strain transformed with pTARA and pET22b:*Ec*AGPase or the corresponding mutants was determined by two colorimetric methods (Archibald et al., 1961; Krisman, 1962). Glycogen pellets were resuspended in 100 µl of water by vigorous shaking during 5 min followed by the addition of 200 µl of 0.2 % anthrone (Sigma) in 95% H_2SO_4. Samples were then incubated at 95°C during 20 min. 200 µl aliquots were transferred to a 96 well-plate and measured at 650 nm using a Spectra Max M2 plate reader. Glucose standards were treated in a similar manner in parallel. Since anthrona is a general reagent for the detection of sugars, a second less sensitive but specific assay was used. Glycogen pellets were washed with 20 µl of saturated NH_4Cl and dry heated at 95°C to remove the excess of ammonia. 200 µl of freshly prepared iodine reagent, a modified Lugol's iodine staining, obtained by mixing 20 µl of 2.6% I_2 and 26% KI in water, with 5 ml of saturated $CaCl_2$ were added. Samples were transferred to a 96 well-plate and measured at 450 nm using a Spectra Max M2 plate reader. Measurements were performed in 4 and 5 replicates for the iodimetric and anthrone methods, respectively.

EcAGPase Crystallization and Data Collection — Crystallization trials were carried out in sitting drop 96 well plates by using a mosquito crystal robot (TTP Labtech). Crystals of *Ec*AGPase•R130A were obtained by mixing 0.25 μl of EcAGPase•R130A at 6.3 mg/ml in 50 mM Tris-HCl pH 7.5, 100 mM NaCl with 0.25 μl of mother liquor containing 14% polyethylene glycol 3.350, 140 mM magnesium formate, 30% ethylene glycol. Crystals grew in 13 days and were frozen under liquid nitrogen. Crystals screening were performed at Diamond Light Source (DLS) (Didcot, Oxfordshire, United Kingdom). *Ec*AGPase•R130A complete dataset was collected at I04 beamline (DLS) with oscillation angle of 0.15° for a total of 1200 images using a Pilatus 6M-F detector. *Ec*AGPase•R130 form crystallized in space group P21 with 8 molecules in the asymmetric unit and diffracted to a maximum resolution of 3.09 Å.

EcAGPase Structure Determination and Refinement — The crystal structure of *Ec*AGPase•R130A was solved by molecular replacement with the program Phaser (McCoy et al., 2007) using a tetramer from the crystal structure of *Ec*AGPase•AMP•SUC, PDB atomic coordinates 5L6V (Cifuente et al., 2016), as search model. The final structure was obtained using alternate cycles of manual model-building using COOT (Emsley et al., 2010) and Phenix (phenix.refine) or Refmac5 (Murshudov et al., 2011). NCS restraints were calculated automatically during refinement and differences between chains subsequently modelled. During the refinement the structure geometry was validated using Molprobity (Chen et al., 2010). Atomic coordinates and structure factors have been deposited with the Protein Data Bank, accession code 5MNI (*Ec*AGPase•R130A). Molecular graphics and structural analyses were performed with the UCSF Chimera package (Pettersen et al., 2004).

Enterobacterial AGPases alignment — A representative group of enterobacterial AGPase protein sequences from different species was obtained from UNIPROT database and sequence aligned using Clustal-Omega. Middle distance BLOSUM62 tree was performed using Jalview.

5.2. Results and discussion

5.2.1. Dissecting the structural determinants of EcAGPase allosteric regulation

It was shown in the previous chapter that *Ec*AGPase•AMP complex was *ca.* 4.6 °C more stable than the unliganded form of the enzyme. Moreover, the addition of FBP to the *Ec*AGPase•AMP complex triggered a clear reduction in the T_m values, indicating that FBP is able to compete with AMP and to modify the structural arrangement of the *Ec*AGPase•AMP complex, leading to a less stable structure. To further advance on the understanding of the molecular mechanism of *Ec*AGPase allosteric regulation, we studied the contribution of a series of chemical derivatives of AMP and FBP to the stabilization of the enzyme. To this end, the ellipticity of *Ec*AGPase in the absence and presence of AMP, orthophosphate, D-ribose 5-phosphate, adenine, FBP and fructose 6-phosphate (F6P), was monitored at 222 nm as a function of temperature. The T_m of *Ec*AGPase was 65.6°C (Figure 5.1A). Upon the addition of AMP, the T_m value increased to 73.6°C. In contrast, the presence of orthophosphate, D-ribose 5-phosphate, adenine did not significantly affect the T_m value of *Ec*AGPase, with 69.0°C, 69.7°C and 69.5°C values, respectively (Figure 5.1A). These experimental observations correlate well with the structural configuration of the AMP allosteric site as visualized in the *Ec*AGPase•AMP•SUC crystal structure (Figure 5.2). Specifically, the α-PO4 of AMP localizes in a pocket comprised by Arg40, His46, Arg52 and Arg386 of the same protomer, with no evident interactions with adjacent protomers of the homotetramer (Figure 5.2B). Similarly, the ribose moiety of D-ribose 5-phosphate provides two additional interactions with Lys39 and Thr79, also located in the same protomer. Interestingly, the *Ec*AGPase•AMP•SUC crystal structure revealed that the adenine heterocycle is stabilized by a strong stacking interaction with Arg130 (α7) from the GT-A-like domain of a neighbouring protomer (Figure 5.2A and 5.2B). However, the nucleobase alone did not provide any measurable stabilization effect.

Different moieties of the AMP chemical scaffold have been studied regarding their modulatory effects on *Ec*AGPase activity. Orthophosphate was characterized as a weak inhibitor of *Ec*AGPase, requiring a much higher concentration to achieve similar inhibitory levels than AMP (Gentner & Preiss, 1967). Most of the AGPase crystal

structures reported to date revealed the presence of either orthophosphate or sulphate anions located in the corresponding binding pocket of the α-PO4 moiety of AMP in the regulatory site of *Ec*AGPase (Cupp-Vickery et al., 2008). Interestingly, structural evidence also revealed (i) the presence of orthophosphate in the active site of *Ec*AGPase, (ii) and sulphate, a phosphate mimic, in the active site of *Solanum tuberosum* AGPase (*St*AGPase), with the ATP substrate in a non-catalytically active conformation (PDB: 1YP3; Jin et al. 2005). This structural information might suggest a competitive inhibitory effect of these anions with the *Ec*AGPase substrates at high concentrations, however, an intra-protomeric inhibitory effect cannot be discarded. Finally, the presence of D-ribose 5-phosphate or adenine did not alter the activity of *Ec*AGPase (Preiss, Shen, and Partridger 1965; Ribéreau-Gayon et al. 1971).

Altogether, these experimental observations indicate that the complete AMP scaffold is required for the stabilization of the enzyme, strongly supporting the notion that AMP inhibitory properties are inherently linked to the inter-protomer crosstalk in the negative allosteric mechanism of *Ec*AGPase. Finally, the addition of FBP to the *Ec*AGPase•AMP complex produced a clear reduction in the T_m value, whereas the presence of F6P did not significantly affect the T_m value of the *Ec*AGPase•AMP complex. Taking into consideration that F6P does not modulate *Ec*AGPase activity (Preiss 1978) together with its inability to alter *Ec*AGPase•AMP stability, these observations points towards the PO4 group at position 1 of fructose as a key player to enhance the enzymatic activity, to modulate AMP inhibition and to reverse AMP stabilization effect (Figure 5.1B and 5.2C).

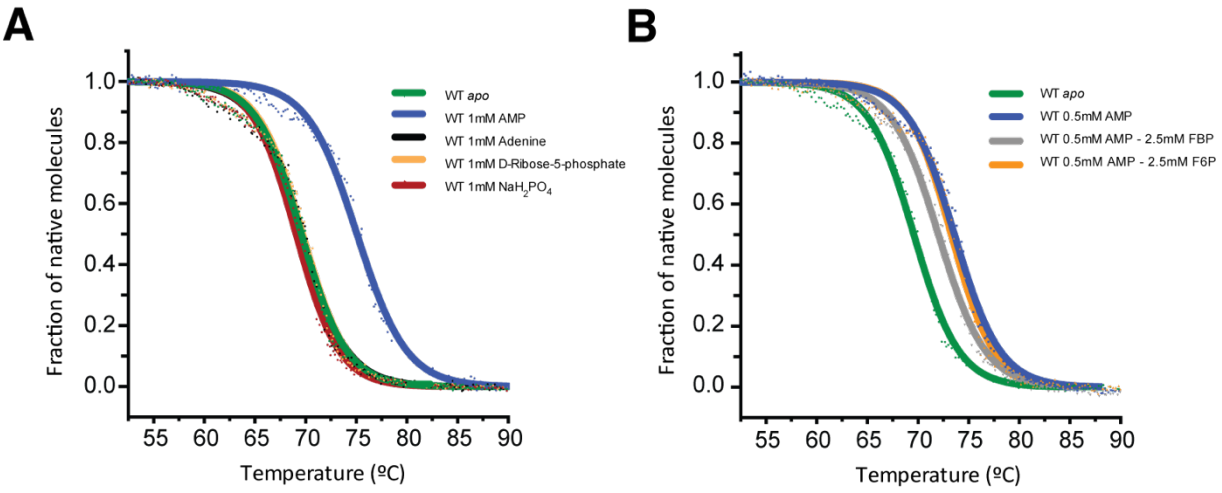


Figure 5.1. EcAGPase thermal unfolding transitions recorded at 222 between 20 ºC and 90 ºC. *(A) Thermal unfolding for the unliganded form of EcAGPase (green) and the EcAGPase•AMP (blue), EcAGPase•Adenine (black), EcAGPase•5RP (yellow) and EcAGPase•PO_4 (red) complexes. (B) Thermal unfolding for the unliganded form of EcAGPase (green) and the EcAGPase•AMP (blue), EcAGPase•AMP•FBP (grey) and EcAGPase•AMP•F6P*

5.2.2. Design of *Ec*AGPase single point mutants in the regulatory cleft and active site

We carefully explored the consequences of single point mutations in the allosteric cleft of *Ec*AGPase, in key residues involved in AMP and FBP binding, looking for changes in the allosteric properties, stabilization and their impact in enzymatic activity (Figure 5.2). Seven residues facing the allosteric cleft, Lys39, Arg40, His46, Arg130, Arg386, Arg419 and Arg423, were selected and replaced by alanine. Residues Lys39, Arg40, His46, and Arg130 belong to the N-terminal GT-A-like domain, whereas Arg386, Arg419 and Arg423 are located in the LβH domain (Figure 5.2). Specifically, Lys39, Arg40, His46 are located in the SM motif, Arg386 in the β25-β26 loop of the LβH domain, and residues Arg419 and Arg423 in the C-terminal α15. The crystal structure of the *Ec*AGPase•AMP•SUC revealed that (i) the α-PO4 group of the negative regulator AMP interacts with Arg40, His46, and Arg386, and (ii) the ribose moiety is at van der Waals distance of Lys39. In addition, the adenine heterocycle is stabilized by a strong stacking interaction with Arg130 and additional van der Waals interactions with Arg419 and Arg386 (Figure 5.2B).

Based on the *Ec*AGPase•FBP crystal structure, residues Lys39, Arg419 and Arg423 of the same protomer interact with the PO4 group at position 6 of FBP, meanwhile Arg130 in the neighbour protomer appears at salt-bridge distance (Figure 5.2C). It is worth noting that Lys39 proved to be essential for the FBP mediated enzymatic activation to take place (Gardiol & Preiss, 1990).Two single point mutants of the predicted catalytic residues, Lys42 and Lys195, were constructed by replacing them by alanine (Führing et al. 2013; Hill et al. 1991; Ballicora et al. 2005). The predicted catalytic residue Lys42 is located in a key region of the SM motif, at very close distance of several residues participating in both AMP and FBP regulators interactions, in the regulatory cleft. Moreover, the side chain of Lys42, is in closed contact with two aspartic residues, Asp142 and Asp276 (Clarisa M. Bejar et al., 2006; Cupp-Vickery et al., 2008; Frueauf, Ballicora, & Preiss, 2001), which were suggested to participate in the interaction of the divalent metal cation Mg^{2+} during catalysis, as observed in other nucleotidyltranferases (Steitz, Smerdon, Jäger, & Joyce, 1994; Swift et al., 2012). Finally, Lys195 is located in the β12-β13 loop of the sugar binding pocket, in the active site of the GT-A like domain, far apart from the allosteric cleft. Lys195 was proposed to interact with the β-PO4 group of ADP-Glc (Cifuente et al., 2016; Jin et al., 2005). It is worth noting that all *Ec*AGPase constructs have no additional amino acids when compared with the native enzyme. They were expressed and purified to apparent homogeneity, following three main steps including anionic exchange, ammonium sulfate precipitation, and hydrophobic interaction criteria (see section 5.1. for details).

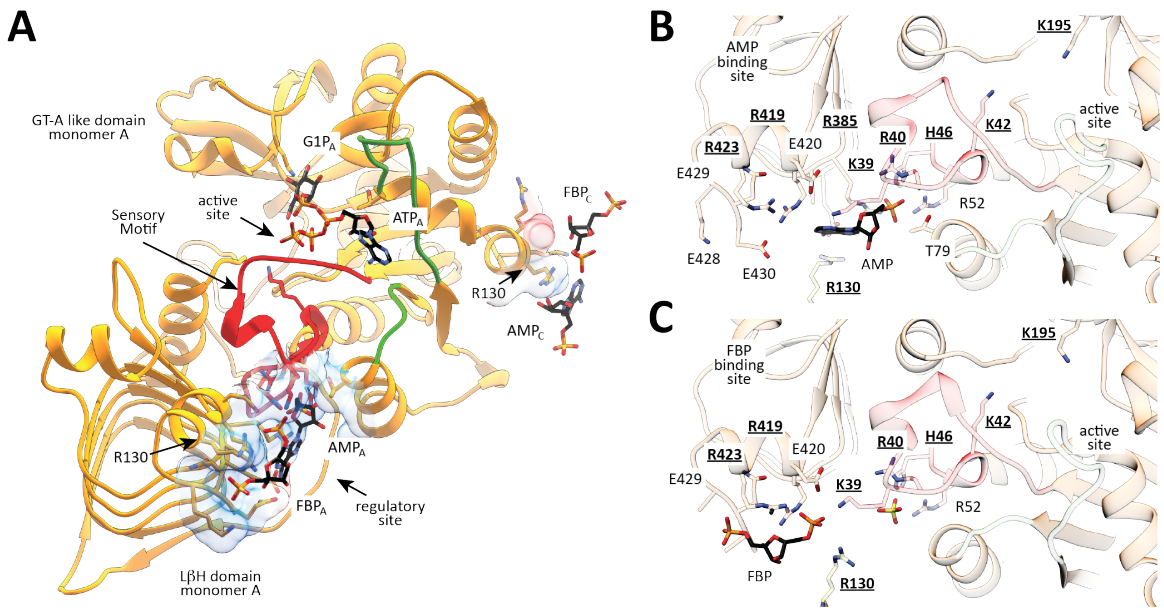


Figure 5.2. Localization of residues in the active and regulatory sites in EcAGPase. *(A) EcAGPase protomer showing the GT-A-like and LβH domains. (B) Close view of the AMP regulatory site, showing the location of key residues involved in AMP binding, and its communication with the active site of EcAGPase. Selected/mutated residues are underlined. (C) Close view of the FBP regulatory site, showing the location of key residues involved in FBP binding, and its communication with the active site of EcAGPase. Selected/mutated residues are underlined.*

5.2.3. Single point mutants localized in the regulatory cleft impact *Ec*AGPase stabilization

The addition of AMP did not modify the T_m values of the *Ec*AGPase•R40A, *Ec*AGPase•R46A, *Ec*AGPase•R130A and *Ec*AGPase•R386A mutants, strongly supporting a role of these residues in AMP binding, as visualized in the crystal structure of the *Ec*AGPase•AMP•SUC complex (Figures 5.2B and 5.3). In contrast, the addition of AMP to the *Ec*AGPase•K39A, *Ec*AGPase•R419A, and *Ec*AGPase•R423A mutants, which are involved in FBP recognition, according to the *Ec*AGPase•FBP crystal structure, triggered a clear increment in the T_m values as observed in the wild-type enzyme (Figure 5.2C, Figure 5.3). As expected, the stabilization of *Ec*AGPase•K42A and *Ec*AGPase•K195A mutants, the predicted catalytic residues, was not affected by the addition of the negative regulator.

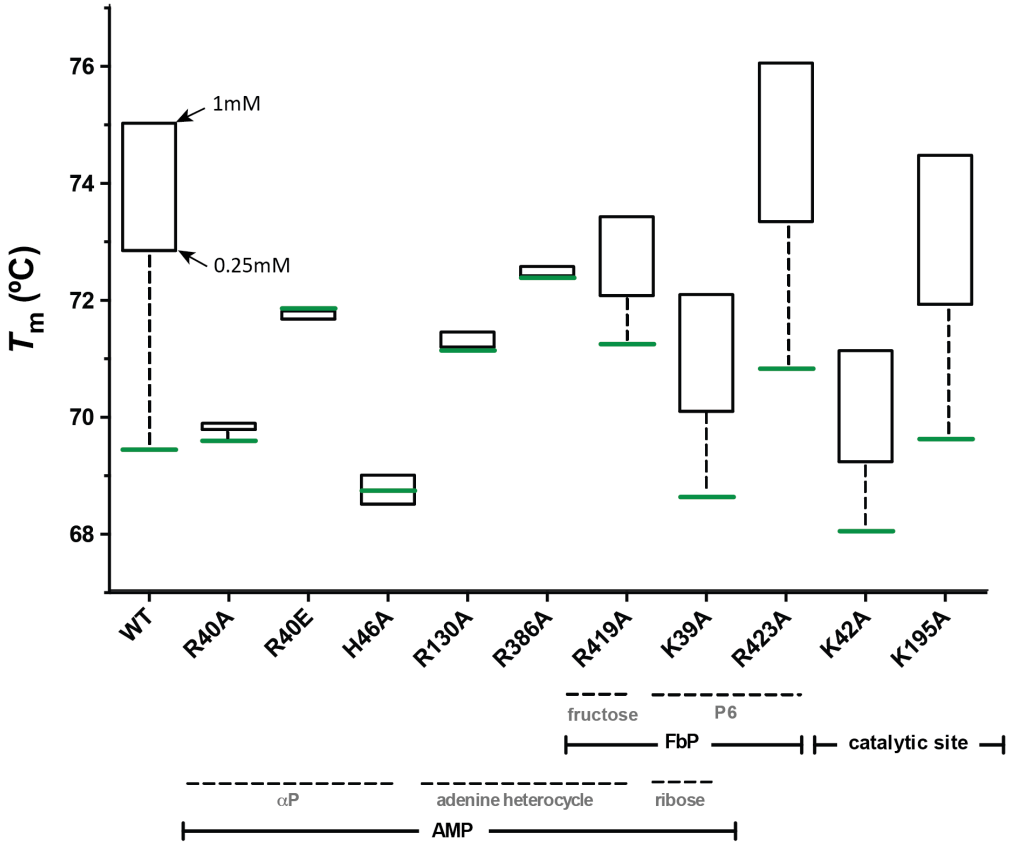


Figure 5.3. Thermal unfolding transitions of EcAGPase and selected EcAGPase variants in complex with AMP. *Thermal unfolding transitions were recorded at 222 nm between 20ºC and 90ºC. Tm value for the apo state of EcAGPase and selected EcAGPase variants are shown in green. AMP concentrations are shown in black.*

Interestingly, the addition of FBP to the *Ec*AGPase•K39A and *Ec*AGPase•R419A was unable to revert the stabilization of the enzyme variants mediated by AMP. In contrast, some reversion is observed in *Ec*AGPase•R423A (Figure 5.4). As expected, *Ec*AGPase•K195A displays a similar behavior. However, the T_m value of the *Ec*AGPase•K42A mutant in the presence of AMP or AMP/FBP regulators remained equally stable. Since the side chain of Lys42 is in closed contact with two aspartic residues, Asp142 and Asp276 (Clarisa M. Bejar et al., 2006; Cupp-Vickery et al., 2008; Frueauf et al., 2001) this phenomenon might reflect the communication between the allosteric and active sites.

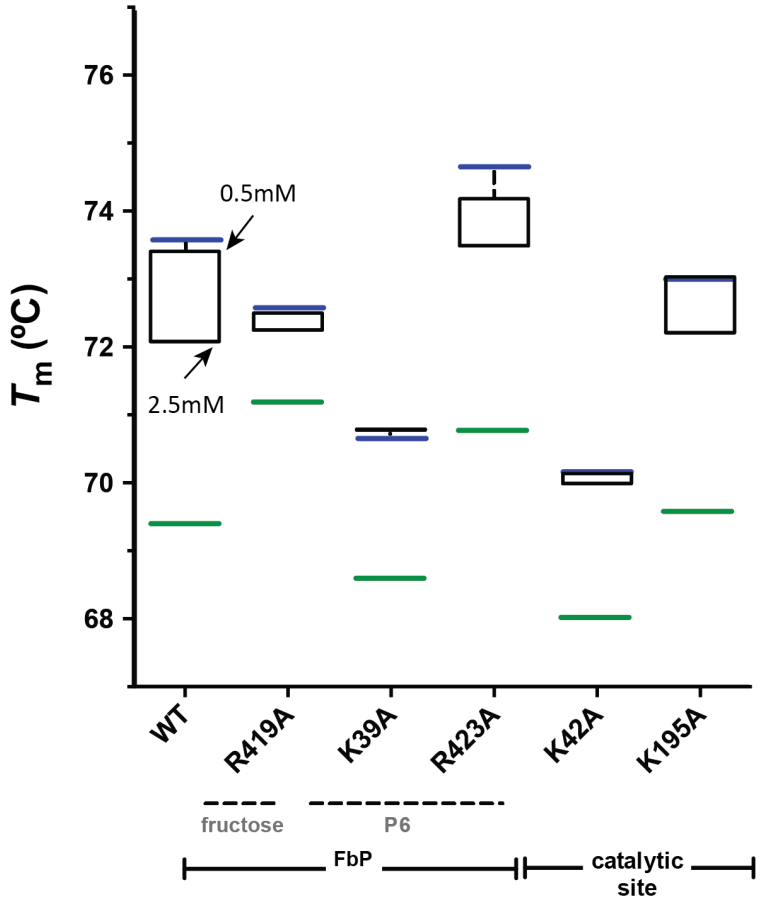


Figure 5.4. Thermal unfolding transitions of EcAGPase and selected EcAGPase variants in complex with AMP. *Thermal unfolding transitions were recorded at 222 nm between 20ºC and 90ºC. Tm value for the apo state of EcAGPase and selected EcAGPase variants are shown in green. AMP concentrations are shown in black.*

.

5.2.4. *Ec*AGPase•R130A deregulates AMP mediated inhibition of the enzymatic activity inducing the overproduction of glycogen *in vivo*

In order to characterize the impact of the single point mutations on the allosteric properties of *Ec*AGPase we measured their specific activities by the colorimetric end-point Malachite Green Phosphate assay. This activity assay showed that *Ec*AGPase•R130A doubles the specific activity compared to the wild-type *Ec*AGPase, in the absence of allosteric regulators. As expected, wild-type *Ec*AGPase is highly

activated by FBP but in contrast, most mutants failed to reach similar levels of activity compared to the *Ec*AGPase•FBP active state (Figure 5.5A). Interestingly, *Ec*AGPase•R130A was the only mutant activated by FBP, consistent with the requirement of an intra-protomeric signalling for the activation of the enzyme. Nevertheless, its activity does not reach the same levels of the wild type, which could be attributable to the lack of additional mechanism to reach full activation, the inter-protomeric crosstalk (Figure 5.5B).

In the presence of AMP, *Ec*AGPase•R40A, *Ec*AGPase•R40E and *Ec*AGPase•H46A mutants (the mutated residues interact with the AMP α-PO4), along with the *Ec*AGPase•R130A and *Ec*AGPase•R386A mutants (the mutated residues interact with the adenine heterocycle and ribose moieties of AMP, respectively), display a similar behavior than that observed for the wild-type enzyme. Specifically, a partial inhibition was shown in all cases, pointing to a compensatory effect of the mutated residues. No inhibition was detected in the *Ec*AGPase•K39A and *Ec*AGPase•R419A mutants (the mutated residues interact with the ribose and adenine heterocycle moieties of AMP, respectively). The *Ec*AGPase•R423A displays a behavior similar to the wild-type enzyme, which is consistent with the lack of interactions between this residue and AMP (Figure 5.5C). The effect of the negative regulator AMP was also addressed in the presence of FBP, as sensitivity to inhibition by AMP is modulated by the concentration of the activator FBP (Figure 5.5D; Gentner and Preiss 1967; J Preiss 1978; Gardiol and Preiss 1990). Strikingly, *Ec*AGPase•R130A displayed (i) a similar specific activity than in the presence of FBP and (ii) 7-fold higher than the wild-type enzyme in the presence of both AMP and FBP regulators, revealing a deregulation of the inactivation mediated by AMP for this mutant. Interestingly, the *Ec*AGPase•K39A and *Ec*AGPase•R419A mutants were unable to completely revert the slight activation induced by FBP in the presence of AMP, which is consistent with the results observed in the thermal unfolding transitions in the presence of both modulators (Figure 5.4). As expected, *Ec*AGPase•K42A and *Ec*AGPase•K195A mutants, the predicted catalytic residues, were inactive in all conditions tested.

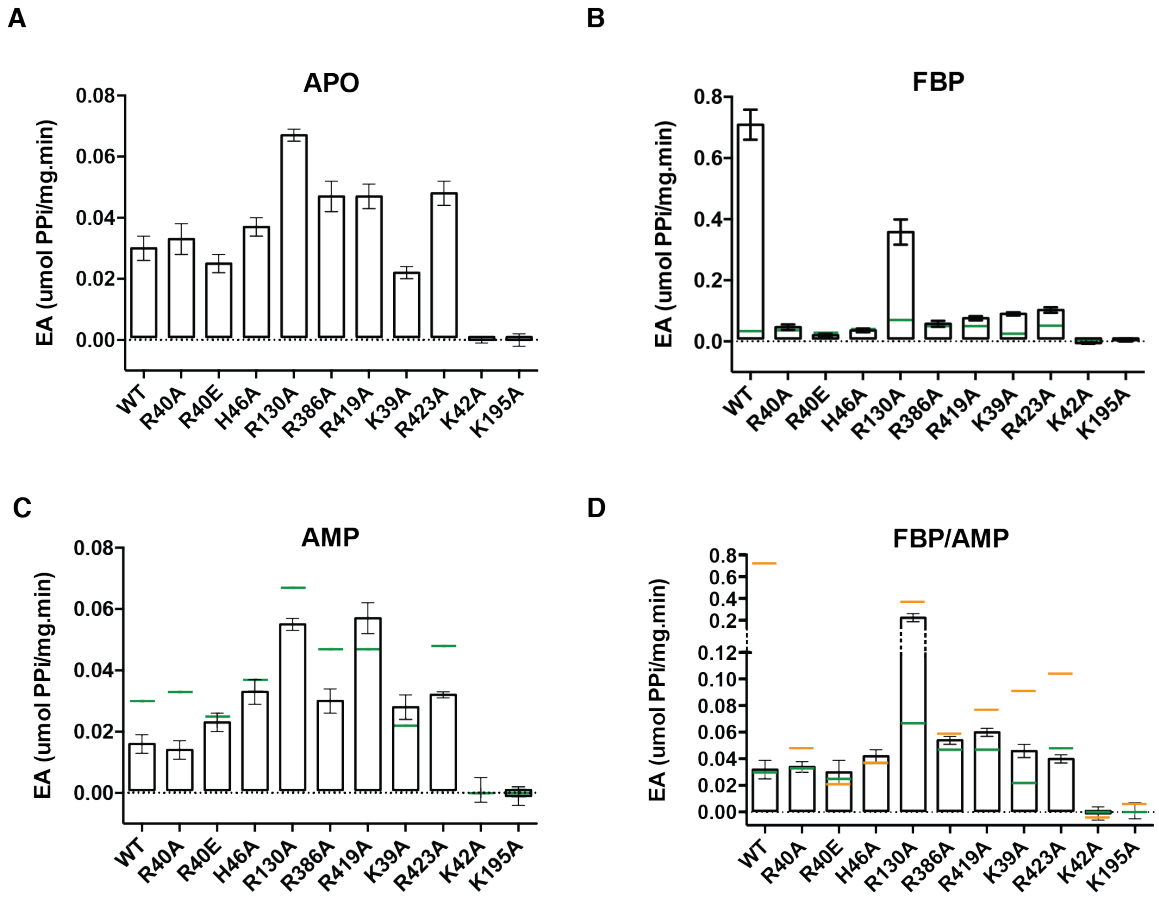
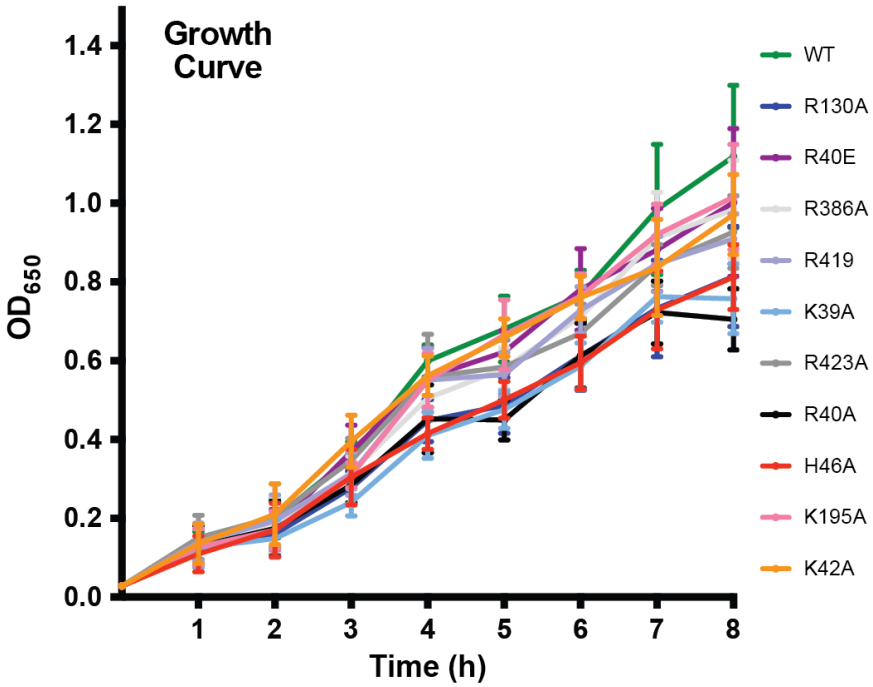


Figure 5.5. Activity measurements of EcAGPase and selected EcAGPase variants. *The enzymatic activity of EcAGPase and selected EcAGPase variants was measured in the absence of allosteric regulators (A), in the presence of FBP (B), in the presence of AMP (C), and in the presence of both positive and negative regulators FBP and AMP.*

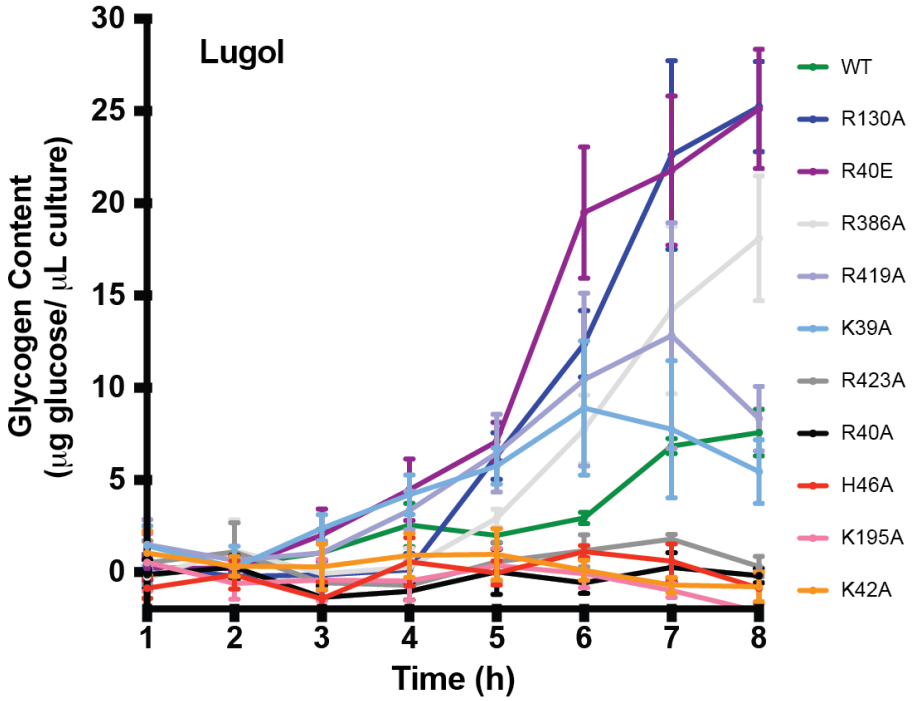
To study the impact of the single point mutations in the *in vivo* glycogen content of *E. coli*, *E. coli* K-12 Δ*glgC* pTARA strain was transformed with pET22b:*Ec*AGPase, or the corresponding mutants. The cultures were synchronized and aliquots were removed every hour before and after the induction for the determination of glycogen content. The growth curves of the different *Ec*AGPase mutants confirmed that they were in the exponential growth phase in all cases (Figure 5.5A). After glycogen extraction, a standard anthrone assay was used for the detection of sugars (Figure 5.5B). As a complementary approach, a less sensitive but more specific iodimetric method was also used to measure glycogen content (Figure 5.5C). As expected, the *Ec*AGPase•K42A and *Ec*AGPase•K195A mutants carrying out the predicted catalytic residues were unable

to accumulate glycogen. Strikingly, *Ec*AGPase•R130A mutant displayed an overproduction of glycogen *in vivo*, which correlates with the specific activities measured for this variant *in vitro*, supporting the idea of an hyperactive deregulated enzyme. Interestingly, *Ec*AGPase•R40E mutant displayed a glycogen content similar to that observed with *Ec*AGPase•R130A, although its specific activities did not show clear evidences of being a highly active enzyme. It is worth noting that *Ec*AGPase is mainly activated by FBP, however, it can be also activated by other glycolytic intermediates (Ball, 2011; Ballicora 2002). The change in the charge of *Ec*AGPase•R40E in the regulatory cleft might therefore modify the effect of other regulators *in vivo* highlighting the complexity of the allosteric mechanism of this key enzyme.

A



B



C

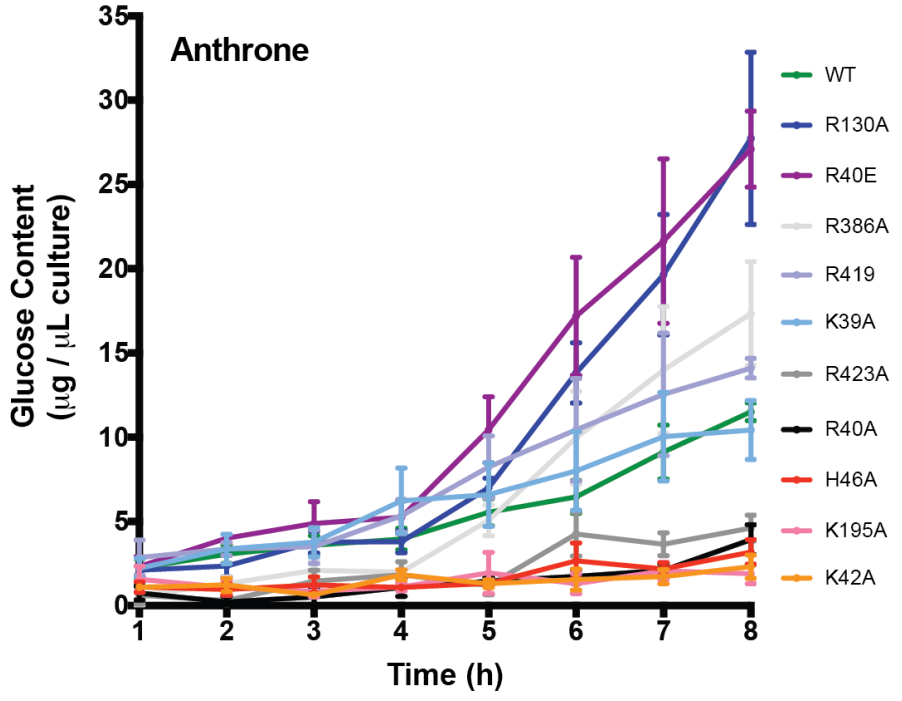


Figure 5.6. Glycogen content in vivo. *(A) Growth curve. (B) Glycogen content by using Lugol method* (Krisman, 1962)*. C. Glycogen content by using anthrone method* (Lodeiro, Di Lorenzo, Petruccelli, Molina-Ortiz, & Sorgentini, 1994).

5.2.5. The crystal structure of *Ec*AGPase•R130A

The crystal structure of *Ec*AGPase•R130A was solved in its unliganded form at 3.09 Å resolution (PDB: 5MNI). *Ec*AGPase•R130A crystallized in space group $P\,2_1$ with eight molecules in the asymmetric unit, representing two homotetramers of the enzyme. When compared with the *Ec*AGPase•AMP•SUC and *Ec*AGPase•FBP complexes, the crystal structure of *Ec*AGPase•R130A revealed significant conformational changes. Firstly, the two dimers of the homotetramer were reoriented, strongly suggesting conformational flexibility of the quaternary structure of *Ec*AGPase (Figure 5.7A and 5.7B). Moreover, conformational differences were also observed in key elements implicated in the proposed allosteric regulatory mechanism. Specifically, residues 26 to 41 of the SM motif displayed a new extended conformation, partially overlapping the α-PO_4 and D-ribose binding sites of AMP, observed in the *Ec*AGPase•AMP•SUC structure (Figure 5.7C). The side chain of Arg419 makes hydrogen-bonding interactions with the main chain carbonyl group of Leu34 and the side chain of Asn38. Interestingly, in the other seven protomers, this region of the SM was found partially disordered. The *Ec*AGPase•AMP•SUC and *Ec*AGPase•FBP crystal structures revealed that both AMP and FBP binding sites partially overlap pointing to an important role of dynamics and conformational changes of several structural elements located in the regulatory cleft, in the signal transduction mechanism. The replacement of Arg130 by alanine certainly affect the transduction of the negative regulatory signal to the active site, likely due to an inter-protomer cross-talk mechanism, as suggested by the *Ec*AGPase•AMP•SUC crystal structure. However, the absence of Arg130 might also modify the rearrangement/dynamics of other structural elements located in the regulatory site of *Ec*AGPase, most notably the SM motif, contributing to the activation/cooperativity of the enzyme.

Similarly, the RL1 loop, which is in close contact with the nucleotide-binding loop (NBL) (residues 26–33) including the GGxGxR consensus sequence involved in ATP binding, was also observed to adopt different structural arrangements. Interestingly, the replacement of the highly conserved Tyr75 per alanine resulted in an inactive enzyme, suggesting a functional role for this residue in the modulation of the enzymatic activity, as reported for the Gln74Ala mutant (Figueroa et al., 2011). Finally, the RL2, could only be observed/partially modelled, in a different conformation with respect to the *Ec*AGPase•AMP•SUC and *Ec*AGPase•FBP structures. Taken together, the new crystal structure of *Ec*AGPase•R130A provides evidence of conformational changes including

local flexibility, in key elements proposed to be involved in allosteric communication. It is worth noting that the transformation of plants with *E. coli* allosteric mutants on the *glgC* gene significantly increased starch content (Tuncel and Okita, 2013). Therefore, the identification of a *Ec*AGPase hyperactive deregulated mutant provides exciting possibilities for industrial/biotechnological applications.

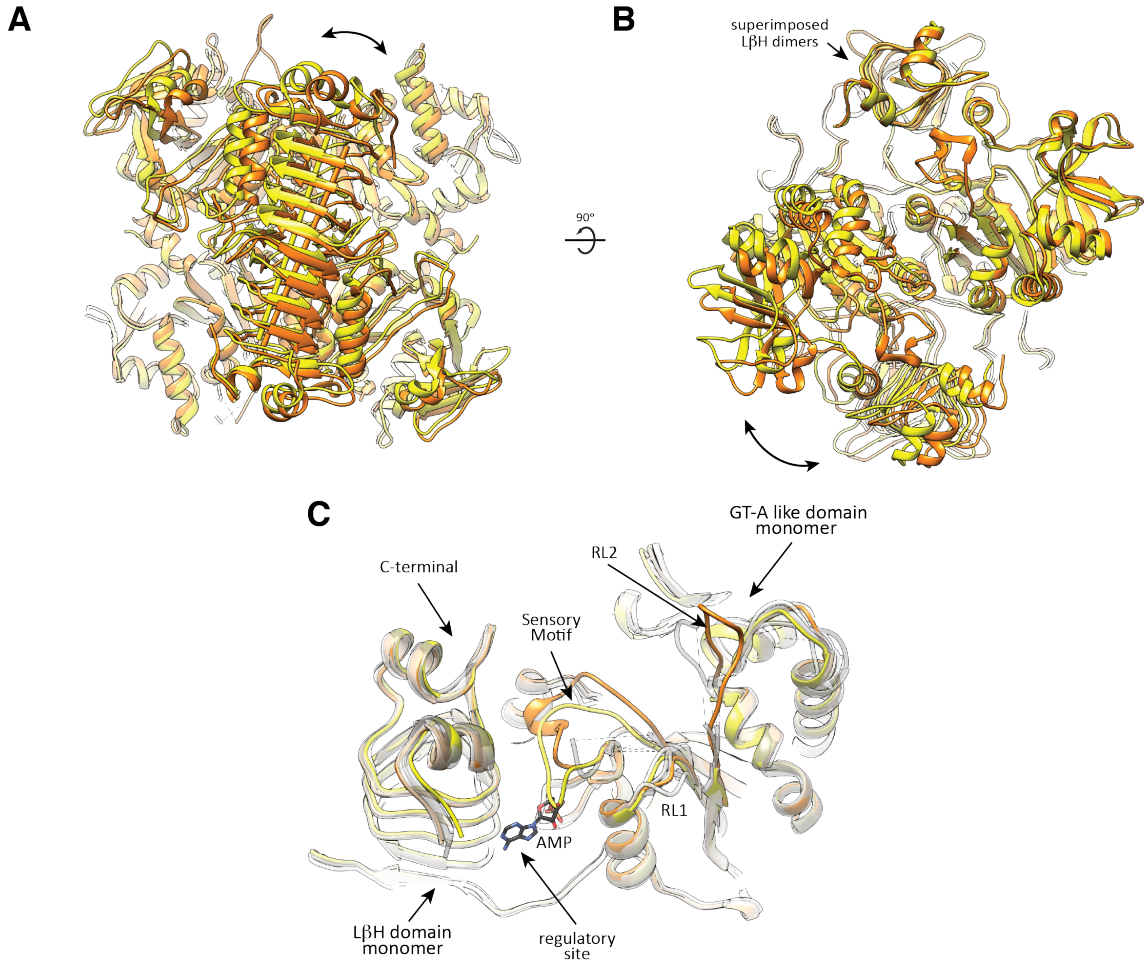


Figure 5.7. The crystal structure of EcAGPase•R130A. *(A-B) Two views of the structural superposition between the EcAGPase•R130A (yellow) and EcAGPase•FBP (orange) tetramers, based on the alignment of LβH domains pairs (arrow). The relative rotation of both domains is shown (curved arrow). C. EcAGPase•R130A structure (protomer D: yellow) centered at the sensory motif, superimposed with an EcAGPase•AMP•SUC protomer (orange), revealing the prominent conformational changes leading the SM motif to partially occupy the AMP binding site. The other EcAGPase•R130A protomers of the asymmetric unit appear superposed in grey scale, showing variability in the RL1, RL2 regions and the G-rich loop.*

5.2.6. *Ec*AGPase shares common sequence signatures for allosteric regulator binding with other enterobacterial AGPases

Multiple primary structure alignment among members of the enterobacteriaceae family of AGPases revealed Lys42 and Lys195, the proposed catalytic residues, along with Arg40, His46, Arg386 and Arg130, residues involved in AMP binding, highly preserved (Figure 5.8). Several experimental data indicated that the C-terminal region of *Ec*AGPase is involved in the recognition of the positive regulator and allosteric activation mechanism of the enzyme (Ballicora, 2002). Specifically, (i) the 419RxMLRKLxxKQER431 sequence, and (ii) the key residue Lys39 which were observed to interact with the positive regulator FBP in the *Ec*AGPase•FBP complex, are strictly conserved among enterobacterial AGPases known to use FBP as positive regulator (Cifuente et al. 2016; J Preiss 1978; Ballicora et al. 2002). Based on primary structure alignment, enterobacteria families, which have not been studied in terms of its allosteric activators and conserve this C-terminal region, are predicted to be also regulated by FBP (*Shigella flexneri, Shigella boydii, Salmonella choleraesuis, Escherichia fergusonii*; Figure 5.8 and 5.9). Interestingly, some enterobacterial AGPases not regulated by FBP (J Preiss 1978; Ballicora et al. 2002) displayed clear discrepancies in the C-terminal region (*Serratia marcences, Serratia liquefaciens, Hafnia alvei*), strongly suggesting that other enterobacteria AGPases lacking this region might have a different positive regulatory mechanism (*Yersinia pestis, Pantoea vagans, Erwinia sp., Pectobacterium atrosepticum, Dickeya dadantii, Proteus vulgaris*; Figure 5.8 and 5.9).

Cell energy metabolism is inherently linked to the evolution of organisms, thus, enterobacterial glycogen synthesis ought to be study in this context. Based on our observations, AGPase seems to follow the same phylogenetic history of the enterobacteriaceae family, as described in the Pathosystems Resource Integration Center (Wattam et al., 2014). In this regard, the conserved use of FBP as positive allosteric regulator in *Ec*AGPase seems to be a trait acquired by a sub-group of enterobacterial AGPases sharing the C-terminal 419RxMLRKLxxKQER431 sequence (Figure 5.9).

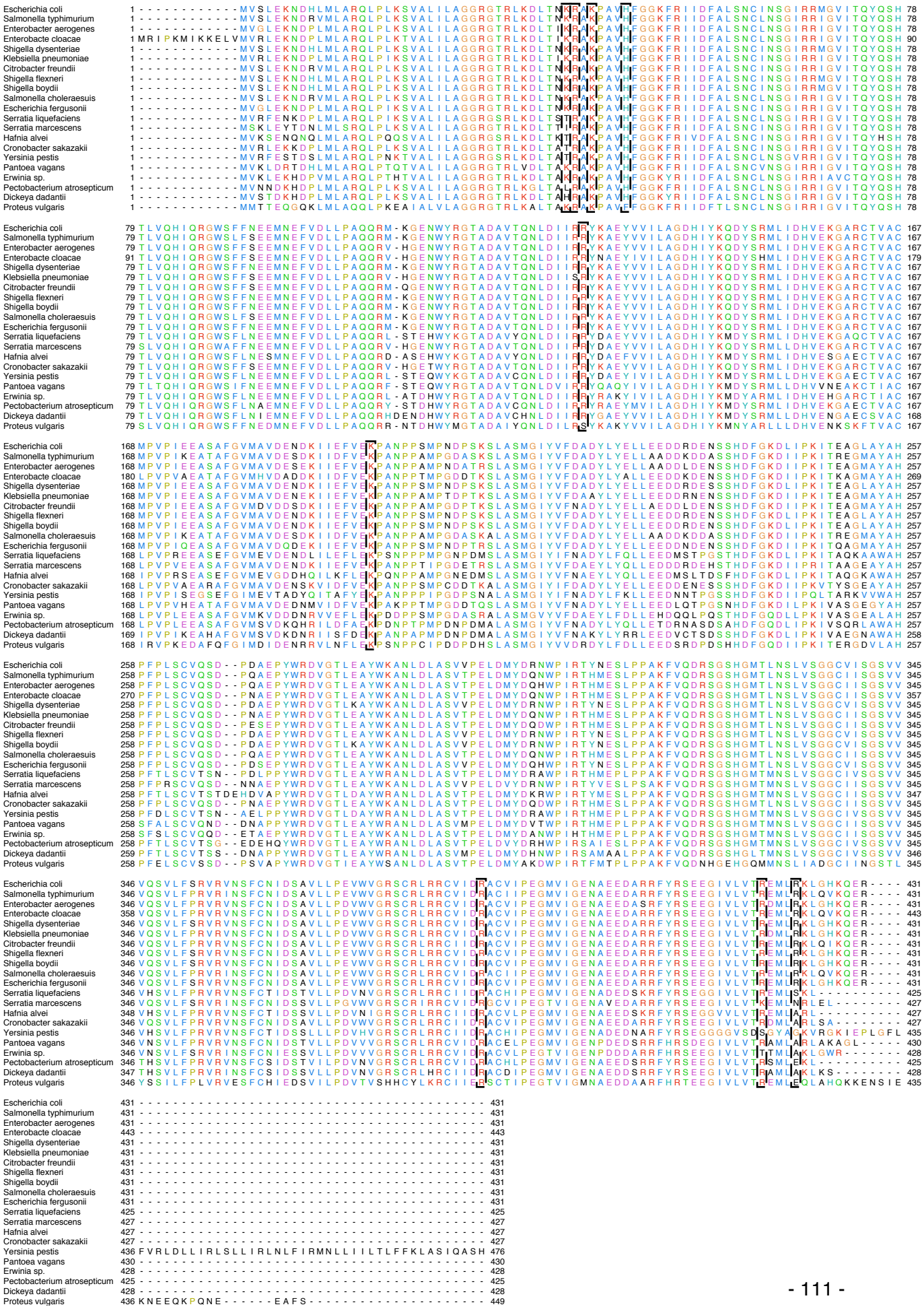


Figure 5.8. Structure-Weighted Sequence Alignment of EcAGPase with other AGPases.
Structural alignment between the crystal structures of EcAGPase (PDB: 5L6V; UniProt: P0A6V1), AtAGPase (PDB: 3BRK; UniProt: P39669), and StAGPase (PDB: 1YP3; UniProt: P23509). The secondary structure elements corresponding to the GT-A-like domain are shown in yellow (a helices) and orange (b helices); and to the LbH domain in green (a helices) and blue (b helices). Residues with poor electron density are highlighted as full boxes. The SM and the RL1 and RL2 loops are highlighted in yellow. Catalytic residues are highlighted as dotted boxes. The RMSD value is shown for each residue. Amino acid sequences of selected AGPases were aligned to the structure alignment: Mycobacterium smegmatis (class II; UniProt: A0R2E1), Serratia marcescens (class II; UniProt: A0A0U6P844), Rhodo- bacter sphaeroides (class V; UniProt: Q9RNH7), Rhodospirillium rubrum (class VI; UniProt: Q9ZFN4), Bacillus subtilis (class VII; UniProt: P39122), Synechococcus sp. (class VIII; UniProt: Q2JU94), Ostreococcus tauri (class VIII; UniProt: Q6PYZ7), Spinacia oleracea (class VIII; UniProt: Q43152), and Triticum aestivum (class IX; UniProt: P30523)

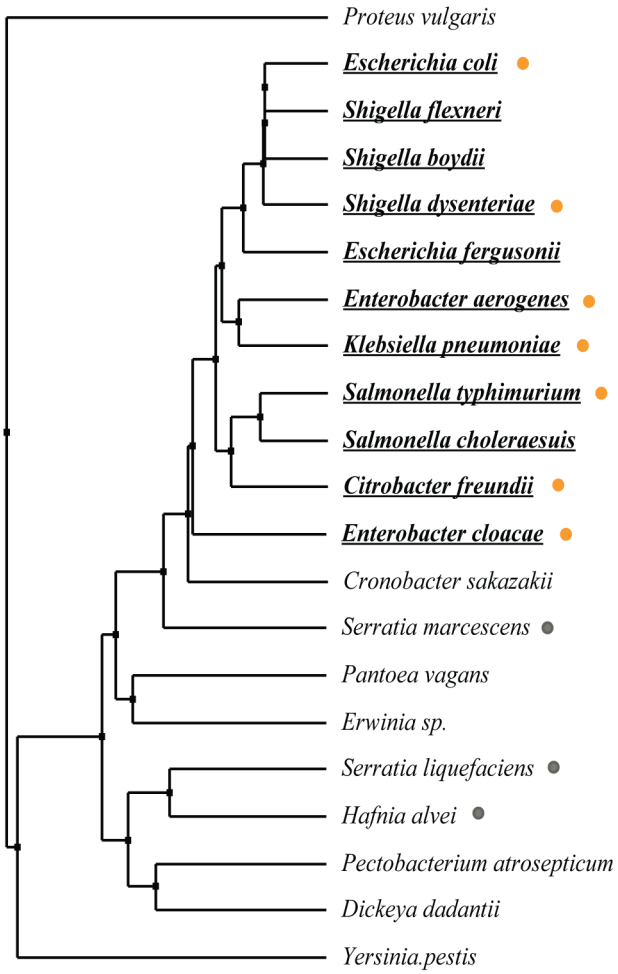


Figure 5.9. Evolution of the FBP regulatory site within the enterobacteriaceae family of AGPases. *Middle distance BLOSUM62 tree. The species known to use FBP are marked with an orange spot and those proposed to use it, based on the conservation of the C-terminus, with an empty orange spot. The species reported not to be regulated by this metabolite are marked with a grey spot and those proposed not to be with an empty grey spot.*

GLOBAL STRUCTURAL QUATERNARY REARRANGEMENTS DERIVED FROM *Ec*AGPase REGULATION

6. Global structural quaternary rearrangements derived from *Ec*AGPase allosteric regulation

As it has been shown in chapter 4, no significant differences were observed in the quaternary structure of *Ec*AGPase•AMP and *Ec*AGPase•FBP crystal structures. However, the crosstalk event promoted by the residue Arg130 between protomers from the same and different dimers in *Ec*AGPase•AMP structure suggests that AMP interactions might lead to the stabilization of the quaternary structure of *Ec*AGPase in solution. Moreover, the differences on the T_m values between *Ec*AGPase in its *apo* state (71.2°C) and *Ec*AGPase•AMP complex (75.8°C) might point to the same hypothesis. In the case of *Ec*AGPase•FBP complex, Arg130 side chain from the neighboring protomer changes its orientation, suggesting that FBP interactions might not lead to the stabilization of the quaternary structure. Interestingly, the addition of FBP to the *Ec*AGPase•AMP complex triggered a clear reduction in the T_m values indicating that FBP not only is able to compete with AMP but also to modify the structural arrangement of the *Ec*AGPase•AMP complex, leading to the occurrence of a less-stable structure.

Changes in the quaternary structure of allosterically modulated enzymes have already been observed (Filippova et al., 2015; Fischer, Olsen, Nam, & Karplus, 2011; Lipscomb & Kantrowitz, 2012). To analyse whether the allosteric regulation of *Ec*AGPase is associated with quaternary structural changes in solution we performed Small Angle X-ray Scattering (SAXS) and cryo-EM studies on *Ec*AGPase.

6.1. Materials and methods

***Ec*AGPase expression and purification** – The recombinant *Ec*AGPase has no additional amino acids when compared to the native enzyme. *E. coli* BL21(DE3) cells transformed with pET29a-*Ec*AGPase were grown in 3,000 ml of LB medium supplemented with 25 μg/ml of kanamycin at 37°C. When the culture reached an OD_{600} of 0.8, the *Ec*AGPase expression was induced by the addition of 1 mM isopropyl ☆ thiogalactopyranoside (IPTG) and further incubated at 18°C overnight. All *Ec*AGPase purification procedures were carried out at 4°C. Cells were harvested by centrifugation at 5,000 x g and resuspended in 40 ml of 50 mM Hepes pH 8.0, 5 mM MgCl2, 0.1 mM EDTA, 10% sucrose (w/v; solution A), containing protease inhibitors (Complete EDTA-free; Roche) and 10 mg/l of lysozyme (Sigma). Cells were then disrupted by sonication (five cycles of 1 min each) and centrifuged for 15 min at 20,000 x g. The supernatant was dialyzed twice against solution A by using an 100,000 Da molecular mass cutoff dialysis membrane. The solution was then applied to a DEAE FF column (5 ml; GE Healthcare) equilibrated with solution A. Elution was performed with a linear 0-0.5M NaCl gradient in 100 ml. Enzyme-positive fractions were pooled and precipitated with ammonium sulfate to 1.2 M (solution B). The resultant suspension was centrifuged for 20 min at 20,000 x g and the resulting supernatant applied into a Phenyl Shodex HIC PH-814 equilibrated in solution B. The enzyme was eluted with a linear gradient of 100% solution B to 100% solution A, in 50 ml. The most active fractions were pooled, concentrated to 10 mg/ml by an Amicon-Ultra spin concentrator (Merck Millipore) with a 100-kDa-molecular mass cutoff, and stored at -80°C.

Small Angle X-ray Scattering Measurements – HPLC coupled SAXS experiments were perform in beamline B21 at Diamond Light Source (Didcot, UK). Samples were injected in an in-line Agilent HPLC unit harboring a Shodex KW403-4F column, previously equilibrated in 50 mM Tris-HCl pH 7.5, 100 mM NaCl buffer for *Ec*AGPase in *apo* state, the same buffer plus 0.5 mM AMP for the *Ec*AGPase•AMP sample and 2.5 mM FBP for *Ec*AGPase•FBP sample. A sample volume of 45 μl at 4.5 mg/ml was injected in the column working at a flow rate of 0.16 ml/min at 15°C for each condition. X-ray scattering data were collected by illuminating 17 μl of sample at wavelength λ = 0.99 Å (12.4 keV) in different frames of 3 s exposure during the elution of the sample. Data was acquired using a Pilatus 2M detector at fixed distance configuration of 4.014 m, resulting

in a range of momentum transfer values of $0.0022 < q < 0.42$ Å^{-1} ($q = 4\pi \sin(\theta)/\lambda$ where 2θ is the scattering angle).

Diffraction data corresponding to each peak was integrated and buffer subtracted using SCATTER (Rambo & Tainer, 2011) and processed using PRIMUS (Konarev, Volkov, Sokolova, Koch, & Svergun, 2003). The forward scattering (I(0)) was evaluated using the Guinier approximation assuming the intensity is represented as $I(q) = I(0)\exp((qR_g)2/3)$, with R_g being the radii of gyration, for a very small range of momentum transfer values ($q < 1.3/R_g$). The maximum dimensions (D_{max}), the interatomic distance distribution functions ($P(r)$), and the R_g were computed using AUTOGNOM (Svergun, 1992).

***Ab Initio* shape determination and docking of the crystal structure**— The low-resolution structures of *Ec*AGPase *apo* and in complex with AMP and FBP were calculated *ab initio* by using the program GASBOR (Svergun, Petoukhov, & Koch, 2001) using P 2 2 2 symmetry. The rigid body fitting of *Ec*AGPase•AMP crystal structure into the SAXS *ab initio* model was performed using Chimera.

Negative stain EM sample preparation and data acquisition – *Ec*AGPase was diluted up to 0.05 mg/ml in 50 mM Tris pH 7.5, 100 mM NaCl to give good individual particle dispersion. CF200-Cu 200 mesh carbon grids were pre-treated by glow discharge for 40 s at 40 mA. First, the sample was added placing the grid on a 6 μl drop of *Ec*AGPase allowing the adsorption for ~2 min. Then, the grid was washed with ddH_2O and blotted with filter paper. Filtered uranyl formate at a concentration of 0.75% was used as the contrast agent leading the grid on the top of a 6 μl droplet for 40 s. After blotting, a fast stain step was performed leading the grid dry for 20 min. Micrographs were collected on a JEM-1230 TEM (JEOL,Japan) equipped with an Orius SC1000 (40008x2672 pixels) CCD camera (GATAN, UK) operating at 120 kV with a magnification of 50,000×.

Cryo-EM sample preparation and data acquisition – Quatifoil-Cu R1.2/1.3 300 mesh grids were pre-treated by glow discharge for 30 s at 4.0 mA and *Ec*AGPase samples were filtered using VIVASPIN 500 PES 1 MDa MWCO to avoid aggregates. Vitrified specimens were prepared by adding 4 μl of *Ec*AGPase 0.30 mg/ml in 50 mM Tris pH 7.5, 100 mM NaCl. Grids were blotted for 3 s after a 15 s pre-blotting time, then plunge-frozen in liquid ethane using VITROBOT FEI instrument, with the chamber maintained at 15°C and 80% humidity. Same conditions were used to prepare *Ec*AGPase samples

in complex with AMP (0.5 mM) and FBP (2.5 mM). Cryo-EM imaging was done on an JEM-2200FS/CR TEM (JEOL, Japan) operated at 200 kV, having Omega energy filter, in minimal dose mode at nominal magnification of 60.000 X. Images were acquired with a UltraScan 4000 SP (4008x4008 pixels) cooled slow-scan CCD camera (GATAN, UK) with a physical pixel sixe of 14 μm, resulting in image with a calibrated pixel size of 1.7 Å. Images were acquired using a defocus range between 2-5 μm with an exposition time of 35ms resulting in a dose of 30 e/ $Å^2$.

Cryo-EM image processing – Micrographs in raw DM3 format were converted to MRC format using the program e2proc2d.py from the package EMAN2 and imported to the program RELION 2.0.3. Determination of the CTF was assessed using CTFFIND (Rohou & Grigorieff, 2015), through the package RELION 2.0.3. In parallel, raw DM3 format micrographs were converted to TIF format using the program ImageJ and afterwards, median filtering for noise reduction using the function 'convert–median 8' from ImageMagick was apply in order to facilitate particle picking in supervised mode using Xmipp3 through the package SCIPION (de la Rosa-Trevín et al., 2016). Using the particle coordinates from picking, single particle images were extracted using RELION 2.0.3 with a final box size of 88x88 pixels, representing an image of 149x149 $Å^2$. In order to classify particles according to their quality, 2D class averages were perform in RELION 2.0.3, and best classes were selected. Particles corresponding to the best 2D classes were used to perform 3D classification imposing D2 symmetry versus an initial model obtained from the crystal structure of *Ec*AGPase tetramer calculated at 40 Å resolution using RELION 2.0.3. The best 3D class was subsequently refined and post-processed leading to a final map for *Ec*AGPase•AMP at 14 Å resolution and for *Ec*AGPase•FBP at 16 Å resolution

The final resolution was estimated using the Fourier Shell Correlation criterion with 0.5 cutoff. In all cases,

Docking of X-ray crystal structures – The docking of X-ray structures into the cryo-EM maps was made using Phenix real space refine (Afonine, Headd, Terwilliger, & Adams, 2013). First, a rigid body fitting was performed for the one tetramer of the *Ec*AGPase•AMP crystal structure. In addition, a second rigid body fitting for each of the four GTA-like domains and the two stacked LβH corresponding to the dimer was perform for the *Ec*AGPase•FBP complex.

6.2. Results and discussion

6.2.1. Allosteric regulators modulate *Ec*AGPase quaternary structure as visualized by SAXS

SAXS belongs to the family of X-ray scattering techniques, providing information about the average particle sizes and shapes in solution. HPLC coupled SAXS (HPLC-SAXS) incorporates an HPLC unit harboring an analytical size exclusion column so that the data can be collected during the elution of the sample (Boldon, Laliberte, & Liu, 2015; Vestergaard, 2016). The resulting chromatograms of *Ec*AGPase on its *apo* state and *Ec*AGPase•AMP complex indicate that *Ec*AGPase is most homogeneous/compact in the presence of AMP (peak Retention Time (T_R) of 15.5 min and Full Width at Half Maximum (FWHM) of 0.5 min; Figure 1.6A) than the *apo* form (T_R = 15.2 min and FWHM = 0.6 min; Figure 1.6A). Moreover, *Ec*AGPase•AMP maintains a constant and low radius of gyration (R_g = 36.8 Å) for each consecutive frame in the SAXS signal plot (Figure 6.1C), indicative of sample homogeneity and compactness. In the *apo* state, *Ec*AGPase is less compact (R_g = 39.2 Å) with some increase in heterogeneity, as indicated by the reduced values of R_g in consecutive frames in the SAXS signal plot (R_g range from 53.7 to 37.0) (Figure 6.1B). These results are consistent with different T_m values observed for the *apo* state and the *Ec*AGPase•AMP complex and the crosstalk event promoted by Arg130 in the presence of AMP (see chapter 4). Interestingly, in the case of transglutaminase type 2 (TG2), a multifunctional allosterically regulated enzyme, the binding of the negative allosteric regulator guanosine triphosphate (GTP) also promotes the adoption of a compact, catalytically inactive conformation (Begg et al., 2006). In contrast, a clear gain in heterogeneity and hydrodynamic radius can be observed in the presence of FBP as indicated by (i) the chromatographic profile (T_R=14.3 min and FWHM = 1.4 min; Figure 6.1A), and (ii) the increase in R_g values and their variability along the SAXS signal plot (R_g range from 52.5 to 41.3 Å; Figure 6.1D). For *Ec*AGPase•FBP peak, data were segmented in three parts for following data integration and analysis (FBP-1, -2, -3; Figure 6.1D). The dimensionless Kratky analysis of the integrated data (Figure 6.1E) also indicates *Ec*AGPase is in a less globular state in the presence of FBP when compared with the *apo* state or the AMP bound state, showing a large increase in particle asymmetry, which correlates very well with the increment in R_g values. Based on these

results, we propose that the competition of AMP and FBP regulators for the allosteric site modulates not only the dynamics of the SM element, and the RL1 and RL2 motifs, but also the compactness of the quaternary structure of the enzyme.

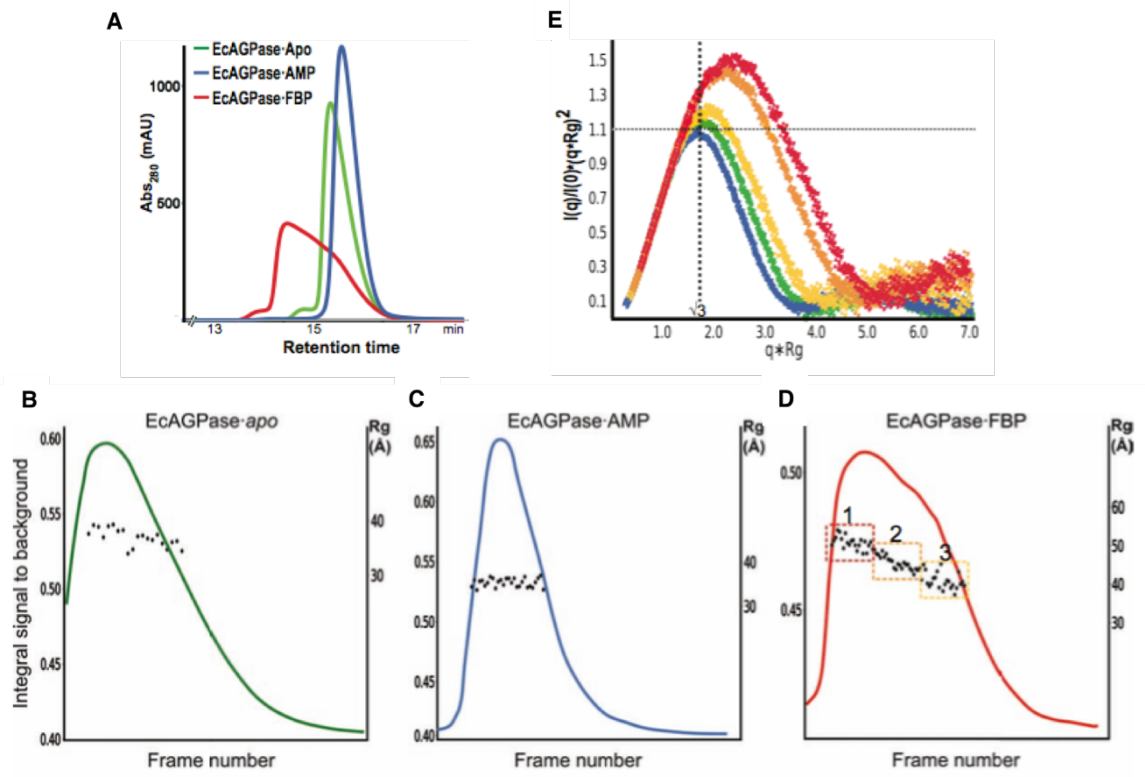


***Figure 6.1. SAXS biophysical analysis**. (A) Size exclusion chromatograms recorded by absorbance measurement at 280 nm corresponding to EcAGPase apo (green), EcAGPase•AMP (blue) and EcAGPase•FBP (red) displaying relative changes in retention times. (B-D) Signal plots of HPLC-SAXS experiment corresponding to apo condition (B), AMP condition (C) and FBP condition (D). Per frame R_g values from buffer subtracted SAXS data are displayed inside the peak as black dots. For EcAGPase•FBP peak data were segmented in three parts indicated by yellow, orange and red dashed boxes for following data integration and analisys (FBP-1, -2, -3). (E) Dimensionless Kratky plot base on volume of correlation (Vc) using Guinier's approximation in which the maximum theoretical height for a sphere is 0.82, showing apo EcAGPase (green) the closer peak maxima to the value. Compared with the apo state the plots for EcAGPase•AMP (blue) evidence AMP triggered compaction as the curve is displaced downwards. Also, FBP-1, -2, -3 data sets (red orange and yellow) displays a decrease in the surface-to-volume ratio as indicated by the curves shift upwards.*

The SAXS *ab initio* models were used to calculate the corresponding envelopes at 20 Å (Figure 6.2A). The shape of *Ec*AGPase•AMP resembles to the *Ec*AGPase•AMP crystal structure surface at low resolution, however, the *apo* state and *Ec*AGPase•FBP

complex display notable differences. In these cases, a less globular quaternary structure can be observed, been this effect more pronounced in the FBP complex. This conformational change is more drastic on the FBP-3 sample which correlates with the R_g and T_R values of this segment of the peak data (Figure 6.1D). Comparison with the *Ec*AGPase•AMP crystal structure envelope might suggest that this "enlargement" of the structure occurs between the GTA domains of *Ec*AGPase with a correlation of 0.58 (Figure 6.2B).

.

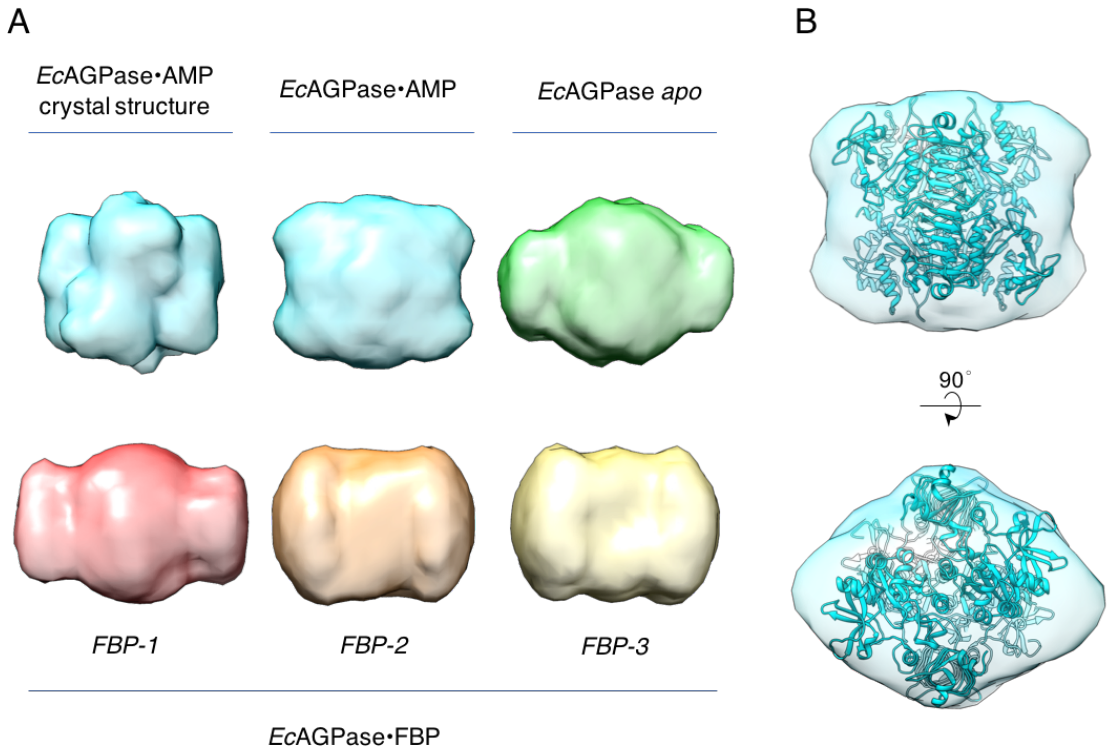


Figure 6.2. SAXS ab initio reconstructions of EcAGPase. *(A) Average low resolution structure of EcAGPase on its apo state (green), in complex with AMP (blue) and FBP (red, orange and yellow). In the case of FBP, the ab initio reconstructions were done for the individual FBP-1, -2, -3 data sets. The crystal structure surface of EcAGPase•AMP complex is incorporated as reference. (B) Average low resolution structure of EcAGPase•AMP complex with the high-resolution crystal structure of EcAGPase•AMP fitted by rigid body docking.*

6.2.2. Allosteric regulators modulate *Ec*AGPase quaternary structure as visualized by cryo-EM

To further advance in the comprehension of the conformational changes derived from the allosteric regulation of *Ec*AGPase, we decided to perform EM for *Ec*AGPase•AMP and *Ec*AGPase•FBP complexes. This technique allows us to determine the structure of *Ec*AGPase in a near-native state and corroborate the SAXS-based model in which we observe quaternary structural discrepancies between both complexes. The purity of the sample was confirmed by negative stain EM (Figure 6.3A). Cryo-EM structures were solved at 14 and 16 Å resolution for the *Ec*AGPase•AMP and *Ec*AGPase•FBP complexes, respectively. The crystal structure of *Ec*AGPase•AMP fits to the cryo-EM density map with a correlation of 0,93 (Figure 6.3B) In contrast, *Ec*AGPase•FBP cryo-EM model highlights the conformational changes previously suggested by SAXS (Figure 6.3C). Specifically, *Ec*AGPase•FBP cryo-EM structure shows a relative rotation of the dimer LβH of 31.2º which results in an increment on the distance between the GTA-like and LβH domains of neighboring protomers, where the allosteric cleft is located (Figure 6.3C).

Altogether, both SAXS and EM approaches support a model in which allosteric regulators modulate *Ec*AGPase quaternary structure. Interestingly, other enzymes of the central energy metabolism display similarities at structural and regulatory level to AGPase, as in the case of Pyruvate kinase (PK), Lactate dehydrogenase (LDH) and Glycerol kinase (GK; Fushinobu et al. 1996; Ormö, Bystrom, and Remington 1998; Valentini et al. 2000). These enzymes are homotetramers with D2 symmetry that binds FBP as allosteric effector in some organisms. In all these enzymes, the allosteric phenomena is proposed to occur by a Monod-Wyman-Changeaux model (MWC; Monod, Wyman, and Changeux 1965), where the symmetry of the tetrameric system must be preserved, and the catalytic properties are explained by conformational displacement between the so called “tense” low-activity/affinity state (T-state) and a "relaxed" high-activity/affinity state (R-state) in a concerted fashion among protomers. These structurally distinct conformational states had been structurally observed bound to modulators (Ikehara et al., 2014; Jurica et al., 1998). It is worth noting that whereas PK and LDH use FBP as positive allosteric modulators, GK is inhibited by this metabolite.

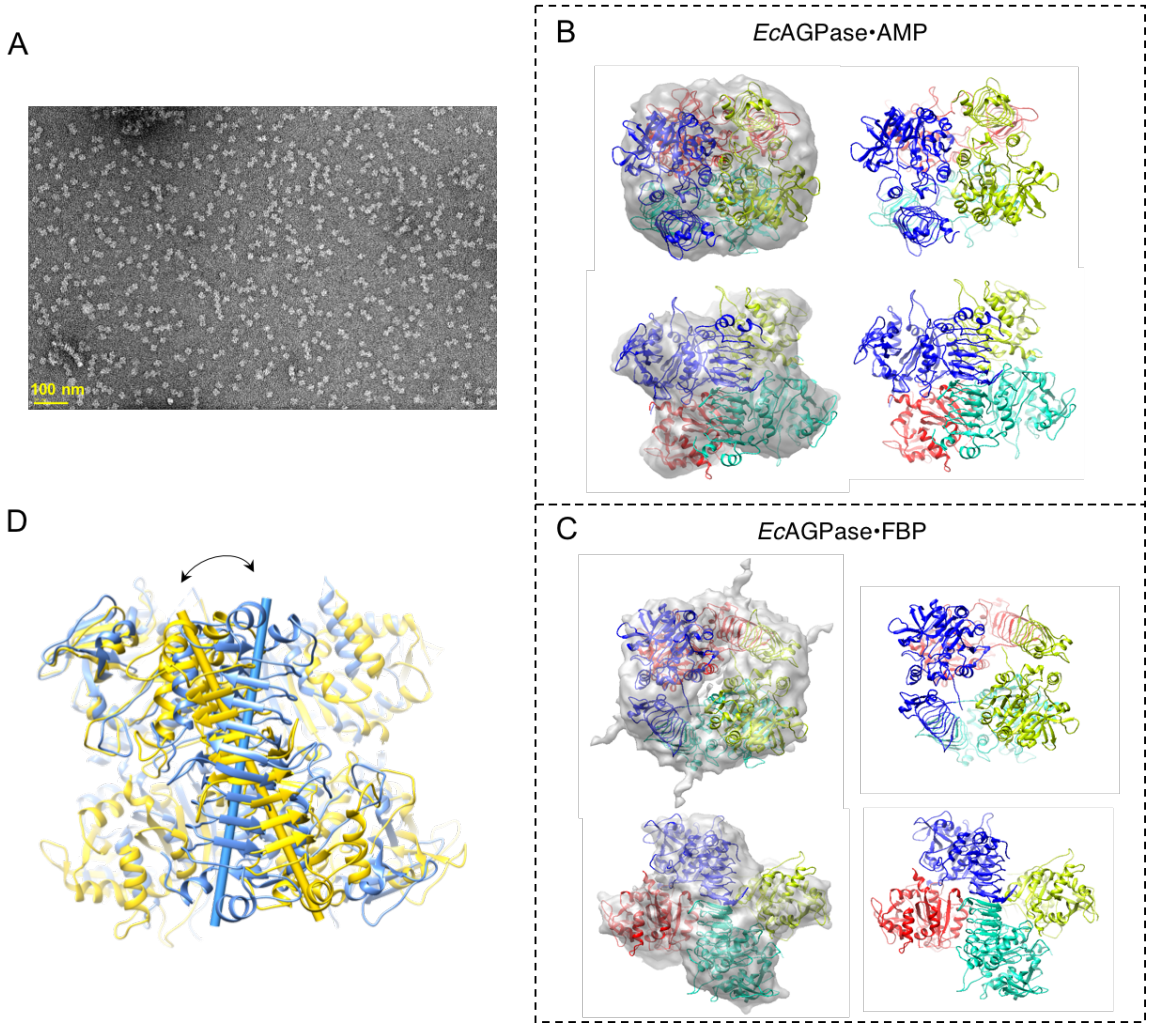


Figure 6.3. EcAGPase cryo-EM reconstructions. *(A) Negative staining of EcAGPase on its apo state. (B) Cryo-EM model of EcAGPase•AMP with the crystal structure of this complex fitted into the cryo-EM density. (C) Cryo-EM model of EcAGPase•FBP with the crystal structure of this complex fitted into the cryo-EM density. (D) Structural superposition of the EcAGPase•FBP crystal structure (blue) and cryo-EM based model (yellow) showing the relative rotation of the dimer LβH (curved arrow).*

The experimental data reported in this work, suggest that the crystal structure of *Ec*AGPase•AMP very likely represents the conformational state of the T-state, meanwhile *Ec*AGPase•FBP also represent this low activity state, but complexed with the positive regulator FBP bound. In this sense, we speculate the crystal structure might represent the first binding event of the positive allosteric regulator into the inactive state of the protein.

CONCLUSIONS

7. Conclusions

In this work, we focused in the study of the enzyme ADP-glucose pyrophosphorylase. This enzyme catalyzes the formation of the NDP-sugar donor ADP-Glc, the first committed reaction in the pathway of bacterial glycogen biosynthesis. Energy reporters within the cell allosterically regulate this key-limiting step modulating AGPase activity (Jack Preiss, Lammel, & Greenberg, 1976; Jack Preiss, Yung, & Baecker, 1983). Specifically, we studied the paradigmatic AGPase from *Escherichia coli*, and the interaction with its preferred allosteric negative regulator AMP and positive regulator FBP. Several experimental lines suggest that both N- and C-terminal domains are involved in *Ec*AGPase allosteric regulation and it was postulated that this phenomenon might be determined by a combined arrangement between both domains (M. A. Ballicora et al., 2002; Clarisa M. Bejar et al., 2006; Jack Preiss, 2014; Wu & Preiss, 1998). However, the lack of structural information regarding AGPase allosteric binding sites has suppose a difficulty in the understanding of the molecular mechanism that governs AGPase allosterism. The main aim of this work was to advance in the study of *Ec*AGPase allosteric regulation determining its structure in complex with AMP and FBP. The structural information derived from these complexes in combination with biophysical/biochemical evidences has placed us in an unprecedented situation to propose a suitable model for *Ec*AGPase regulation. It is worth noting that the transformation of plants with *E. coli* allosteric mutants on the *glgC* gene significantly increased (i) the rate of starch synthesis in tubers of transgenic potato (Sweetlove, Burrell, & ap Rees, 1996) and (ii) starch content (Stark, Timmerman, Barry, Preiss, & Kishore, 1992; Tuncel & Okita, 2013). Therefore, the information reported herein provides exciting possibilities for industrial/biotechnological applications.

Specifically, our new results concluded that:

1) Based on *Ec*AGPase•AMP•SUC and *Ec*AGPase•FBP crystal structures, AMP and FBP bind to partially overlapping sites located in a deep cleft between GT-A like and LβH domains of neighboring protomers. Each allosteric cleft is communicated with the corresponding active site of the same protomer through a region defined as the

'Sensory Motif' (SM), a complex structural element constituted by the nucleotide-binding loop (NBL), including a G-rich motif involved in ATP binding, and a segment rich in short secondary structure elements.

2) *Ec*AGPase•AMP complex is ca. 4.6 °C more stable than the unliganded form of the enzyme and the addition of FBP to this complex triggered a clear reduction in the apparent melting temperatures (T_m) values, indicating that FBP is able to compete with AMP and to modify the structural arrangement of the *Ec*AGPase•AMP complex, leading to a less stable structure.

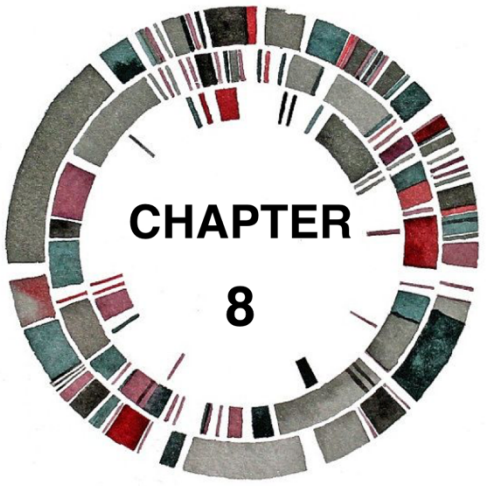
3) Taking into account these experimental evidences, we proposed a model, in which the binding of the positive and negative energy reporters regulates *Ec*AGPase catalytic activity through the SM and two critical regulatory loops RL1 and RL2 flanking the active binding site, via intra-protomer interactions and inter-protomer crosstalk.

4) Single point mutants on key residues of the AMP binding site decrease its inhibitory effect, but in addition, clearly abolish the overall stabilization effect mediated by this regulator in wild-type *Ec*AGPase. On the other hand, single point mutants on key residues of FBP binding were unable to revert the stabilization mediated by AMP.

5) *Ec*AGPase•R130A deregulated mutant display a dramatic increase in the activity when compared with wild-type *Ec*AGPase, which correlates with a significant increment of glycogen content *in vivo*. The crystal structure of EcAGPase•R130A revealed unprecedented conformational changes in structural elements proposed to be involved in the allosteric signal transmission.

6) The determination of *Ec*AGPase structure in solution by SAXS and cryo-EM indicate that its allosteric regulation also involves conformational changes in the quaternary structure, being the *Ec*AGPase•AMP complex more compact than the

*Ec*AGPase•FBP activated state. This experimental data suggests that *Ec*AGPase•AMP crystal structure corresponds to the conformational state of the MWC model based T-state whereas we speculate the *Ec*AGPase•FBP crystal structure might represent the first binding event of the positive allosteric regulator into the inactive state of the protein.



BIBLIOGRAPHY

A

Abad, M. C., Binderup, K., Rios-Steiner, J., Arni, R. K., Preiss, J., & Geiger, J. H. (2002). The x-ray crystallographic structure of Escherichia coli branching enzyme. *Journal of Biological Chemistry*, *277*(44), 42164–42170.

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., … Zwart, P. H. (2010). PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography*, *66*(Pt 2), 213–221.

Afonine, P. V, Headd, J. J., Terwilliger, T. C., & Adams, P. D. (2013). PHENIX News. *Computational Crystallography Newsletter*, *4*, 43–44.

Agirre, C. (2015). Validation of carbohydrate structures in CCP4 6.5.

Albesa-Jové, D., Giganti, D., Jackson, M., Alzari, P. M., & Guerin, M. E. (2014). Structure-function relationships of membrane-associated GT-B glycosyltransferases. *Glycobiology*, *24*(2), 108–124.

Albesa-Jové, D., & Guerin, M. E. (2016). The conformational plasticity of glycosyltransferases. *Current Opinion in Structural Biology*, *40*, 23–32.

Alonso-Casajús, N., Dauvillée, D., Viale, A. M., Muñoz, F. J., Baroja-Fernández, E., Morán-Zorzano, M. T., … Pozueta-Romero, J. (2006). Glycogen phosphorylase, the product of the glgP gene, catalyzes glycogen breakdown by removing glucose units from the nonreducing ends in Escherichia coli. *Journal of Bacteriology*, *188*(14), 5266–5272.

Archibald, A. R., Fleming, I. D., Liddle, A. M., Manners, D. J., Mercer, G. A., & Wright, A. (1961). 232. α-1,4-Glucosans. Part XI. The absorption spectra of glycogen– and amylopectin–iodine complexes. *J. Chem. Soc.*, 1183–1190.

Ardèvol, A., & Rovira, C. (2011). The molecular mechanism of enzymatic glycosyl transfer with retention of configuration: Evidence for a short-lived oxocarbenium-like species. *Angewandte Chemie - International Edition*, *50*(46), 10897–10901.

Asención Diez, M. D., Demonte, A. M., Syson, K., Arias, D. G., Gorelik, A., Guerrero, S. A., … Iglesias, A. A. (2015). Allosteric regulation of the partitioning of glucose-1-phosphate between glycogen and trehalose biosynthesis in Mycobacterium tuberculosis. *Biochimica et Biophysica Acta*, *1850*(1), 13–21.

B

Bai, X. C., Fernandez, I. S., McMullan, G., & Scheres, S. H. W. (2013). Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *eLife*, *2013*(2), 2–13.

Baker, T. S., Olson, N. H., & Fuller, S. D. (1999). Adding the Third Dimension to Virus Life Cycles: Three-Dimensional Reconstruction of Icosahedral Viruses from Cryo-Electron Micrographs.

Ball, S., Colleoni, C., Cenci, U., Raj, J. N., & Tirtiaux, C. (2011). The evolution of glycogen and starch metabolism in eukaryotes gives molecular clues to understand the establishment of plastid endosymbiosis. *Journal of Experimental Botany*, *62*(6), 1775–1801.

Ball, S. G., & Morell, M. K. (2003). FROM BACTERIAL GLYCOGEN TO STARCH: Understanding the Biogenesis of the Plant Starch Granule. *Annual Review of Plant Biology*, *54*(1), 207–233.

Ball, S., Guan, H. P., James, M., Myers, A., Keeling, P., Mouille, G., … Preiss, J. (1996). From glycogen to amylopectin: A model for the biogenesis of the plant starch granule. *Cell*, *86*(3), 349–352.

Ballicora, M. a., Dubay, J. R., Devillers, C. H., & Preiss, J. (2005). Resurrecting the ancestral enzymatic role of a modulatory subunit. *Journal of Biological Chemistry*, *280*(11), 10189–10195.

Ballicora, M. A., Dubay, J. R., Devillers, C. H., & Preiss, J. (2005). Resurrecting the ancestral enzymatic role of a modulatory subunit. *Journal of Biological Chemistry*, *280*(11), 10189–10195.

Ballicora, M. a., Erben, E. D., Yazaki, T., Bertolo, A. L., Demonte, A. M., Schmidt, J. R., … Preiss, J. (2007). Identification of regions critically affecting kinetics and allosteric regulation of the Escherichia coli ADP-glucose pyrophosphorylase by modeling and pentapeptide-scanning mutagenesis. *Journal of Bacteriology*, *189*(14), 5325–5333.

Ballicora, M. A., Sesma, J. I., Iglesias, A. A., & 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*(30), 9431–9437.

Ballicora, M. a, Iglesias, A. A., & Preiss, J. (2003). ADP-glucose pyrophosphorylase, a regulatory enzyme for bacterial glycogen synthesis. *Microbiology and Molecular Biology Reviews*, *67*(2), 213–225, table of contents.

Barengo, R., Flawia, M., & Krisman, C. R. (1975). The initiation of glycogen biosynthesis in Escherichia coli. *FEBS Letters*, *53*(3), 274–8.

Begg, G. E., Carrington, L., Stokes, P. H., Matthews, J. M., Wouters, M. a, Husain, A., … Graham, R. M. (2006). Mechanism of allosteric regulation of transglutaminase 2 by GTP. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(52), 19683–19688.

Bejar, C. M., Ballicora, M. A., Iglesias, A. A., & Preiss, J. (2006). ADPglucose pyrophosphorylase's N-terminus: Structural role in allosteric regulation. *Biochemical and Biophysical Research Communications*, *343*(1), 216–221.

Bejar, C. M., Jin, X., Ballicora, M. A., & Preiss, J. (2006). Molecular architecture of the glucose 1-phosphate site in ADP-glucose pyrophosphorylases. *Journal of Biological Chemistry*, *281*(52), 40473–40484.

Blankenfeldt, W., Asuncion, M., Lam, J. S., & Naismith, J. H. (2000). The structural basis of the catalytic mechanism and regulation of glucose-1-phosphate thymidylyltransferase (RmlA). *EMBO Journal*, *19*(24), 6652–6663.

Bobovská, A., Tvaroška, I., & Kóňa, J. (2015). Theoretical study of enzymatic catalysis explains why the trapped covalent intermediate in the E303C mutant of glycosyltransferase GTB was not detected in the wild-type enzyme. *Glycobiology*, *25*(1), 3–7.

Boldon, L., Laliberte, F., & Liu, L. (2015). Review of the fundamental theories behind small angle X-ray scattering, molecular dynamics simulations, and relevant integrated application. *Nano Reviews*, *6*, 25661.

Bragg, W. L. (1929). The Determination of Parameters in Crystal Structures by means of Fourier Series. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *123*(792), 537–559.

Breton, C., Šnajdrová, L., Jeanneau, C., Koča, J., & Imberty, A. (2006). Structures and mechanisms of glycosyltransferases. *Glycobiology*, *16*(2), 29–37.

Brown, K., Pompeo, F., Dixon, S., Mengin-Lecreulx, D., Cambillau, C., & Bourne, Y. (1999). Crystal structure of the bifunctional N-acetylglucosamine 1-phosphate uridyltransferase from Escherichia coli: A paradigm for the related pyrophosphorylase superfamily. *EMBO Journal*, *18*(15), 4096–4107.

Buchbinder, J. L., Rath, V. L., & Fletterick, R. J. (2001). Structural Relationships Among Regulated and Unregulated Phosphorylases. *Annual Review of Biophysics and Biomolecular Structure*, *30*(1), 191–209.

Buschiazzo, A., Ugalde, J. E., Guerin, M. E., Shepard, W., Ugalde, R. A., & Alzari, P. M. (2004). Crystal structure of glycogen synthase: homologous enzymes catalyze glycogen synthesis and degradation. *The EMBO Journal*, *23*(16), 3196–205.

C

Callaway, E. (2015). The Revolution Will Not Be Crystallized. *Nature*, *525*, 172–174.

Campagnolo, M., Campa, C., Zorzi, R. De, Wuerges, J., & Geremia, S. (2008). X-ray studies on ternary complexes of maltodextrin phosphorylase. *Archives of Biochemistry and Biophysics*, *471*(1), 11–19.

Caner, S., Nguyen, N., Aguda, A., Zhang, R., Pan, Y. T., Withers, S. G., & Brayer, G. D. (2013). The structure of the Mycobacterium smegmatis trehalose synthase reveals an unusual active site configuration and acarbose-binding mode. *Glycobiology*, *23*(9), 1075–1083.

Cantarel, B. I., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., & Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Research*, *37*(SUPPL. 1), 233–238.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., … Richardson, D. C. (2010). MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*, *66*(1), 12–21.

Chenal, A., Karst, J. C., Sotomayor Pérez, A. C., Wozniak, A. K., Baron, B., England, P., & Ladant, D. (2010). Calcium-induced folding and stabilization of the intrinsically disordered RTX domain of the CyaA roxin. *Biophysical Journal*, *99*(11), 3744–3753.

Cheng, Y., Grigorieff, N., Penczek, P. A., & Walz, T. (2015). A Primer to Single-Particle Cryo-Electron Microscopy. *Cell*, *161*(3), 438–449.

Cid, E., Gomis, R. R., Geremia, R. a, Guinovart, J. J., & Ferrer, J. C. (2000). Identification of two essential glutamic acid residues in glycogen synthase. *The Journal of Biological Chemistry*, *275*(43), 33614–21.

Cifuente, J. O., Comino, N., Madariaga-Marcos, J., López-Fernández, S., García-Alija, M., Agirre, J., … Guerin, M. E. (2016). Structural Basis of Glycogen Biosynthesis Regulation in Bacteria. *Structure*, *24*(9), 1613–1622.

Cockburn, D., Wilkens, C., Ruzanski, C., Andersen, S., Willum Nielsen, J., Smith, A. M., … Svensson, B. (2014). Analysis of surface binding sites (SBSs) in carbohydrate active enzymes with focus on glycoside hydrolase families 13 and 77 — a mini-review. *Biologia*, *69*(6), 705–712.

Comino, N., Cifuente, J. O., Marina, A., Orrantia, A., Eguskiza, A., & Guerin, M. E. (2017). Mechanistic insights into the allosteric regulation of bacterial ADP-glucose pyrophosphorylases. *Journal of Biological Chemistry*, *292*(15), 6255–6268.

Coutinho, P. M., Deleury, E., Davies, G. J., & Henrissat, B. (2003). An evolving hierarchical family classification for glycosyltransferases. *Journal of Molecular Biology*, *328*(2), 307–317.

Crevillén, P., Ballicora, M. a., Mérida, Á.́ , Preiss, J., & 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*(31), 28508–28515.

Cupp-Vickery, J. R., Igarashi, R. Y., Perez, M., Poland, M., & Meyer, C. R. (2008). Structural analysis of ADP-glucose pyrophosphorylase from the bacterium Agrobacterium tumefaciens. *Biochemistry*, *47*(15), 4439–4451.

D

Dauvillée, D., Kinderf, I. S., Li, Z., Samuel, M. S., Rampling, L., Ball, S., … Kosar-hashemi, B. (2005). Role of the Escherichia coli glgX Gene in Glycogen Metabolism. *Journal of Bacteriology*, *187*(4), 1465–1473.

de la Rosa-Trevín, J. M., Quintana, A., del Cano, L., Zaldivar, A., Foche, I., Gutierrez, J., … Carazo, J. M. (2016). Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy. *Journal of Structural Biology*, *195*(1), 93–99.

E

Egli, M. (2010). Diffraction Techniques in Structural Biology. In *Current Protocols in Nucleic Acid Chemistry* (Vol. 2009, p. 7.13.1-7.13.35). Hoboken, NJ, USA: John Wiley & Sons, Inc.

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography*, *66*(Pt 4), 486–501.

Espada, J. (1962). Enzymic Synthesis of Adenosine Diphosphate Glucose from Glucose 1-Phosphate and Adenosine Triphosphate. *J. Biol. Chem.*, *237*(12), 3577–3581.

F

Feng, L., Fawaz, R., Hovde, S., Gilbert, L., Chiou, J., & Geiger, J. H. (2015). Crystal Structures of Escherichia coli Branching Enzyme in Complex with Linear Oligosaccharides. *Biochemistry*, *54*(40), 6207–6218.

Figueroa, C. M., Esper, M. C., Bertolo, A., Demonte, A. M., Aleanzi, M., Iglesias, A. a., & Ballicora, M. a. (2011). Understanding the allosteric trigger for the fructose-1,6-bisphosphate regulation of the ADP-glucose pyrophosphorylase from Escherichia coli. *Biochimie*, *93*(10), 1816–1823.

Filippova, E. V., Weigand, S., Osipiuk, J., Kiryukhina, O., Joachimiak, A., & Anderson, W. F. (2015). Substrate-Induced Allosteric Change in the Quaternary Structure of the Spermidine N-Acetyltransferase SpeG. *Journal of Molecular Biology*, *427*(22), 3538–3553.

Fischer, S., Olsen, K. W., Nam, K., & Karplus, M. (2011). Unsuspected pathway of the allosteric transition in hemoglobin. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(14), 5608–5613.

Fox, J., Kennedy, L. D., Hawker, J. S., Ozbun, J. L., Greenberg, E., Lammel, C., & Preiss, J. (1973). De Novo synthesis of bacterial glycogen and plant starch by ADPG: α-

glucan 4-glucosyl transferase. *Annals of the New York Academy of Sciences*, *210*(1), 90–102.

Frueauf, J. B., Ballicora, M. a., & Preiss, J. (2001). Aspartate Residue 142 Is Important for Catalysis by ADP-glucose Pyrophosphorylase from Escherichia coli. *Journal of Biological Chemistry*, *276*(49), 46319–46325.

Führing, J., Cramer, J. T., Routier, F. H., Lamerz, A. C., Baruch, P., Gerardy-Schahn, R., & Fedorov, R. (2013). Catalytic mechanism and allosteric regulation of UDP-glucose pyrophosphorylase from Leishmania major. *ACS Catalysis*, *3*(12), 2976–2985.

Furukawa, K., Tagaya, M., Tanizawa, K., & Fukui, T. (1993). Role of the conserved Lys-X-Gly-Gly sequence at the ADP-glucose-binding site in *Escherichia coli* glycogen synthase. *Journal of Biological Chemistry*, *268*(32), 23837–23842.

Furukawa, K., Tagaya, M., Tanizawa, K., & Fukui, T. (1994). Identification of Lys277 at the active site of Escherichia coli glycogen synthase. Application of affinity labeling combined with site-directed mutagenesis. *The Journal of Biological Chemistry*, *269*(2), 868–71.

Fusari, C., Demonte, A. M., Figueroa, C. M., Aleanzi, M., & Iglesias, A. A. (2006). A colorimetric method for the assay of ADP-glucose pyrophosphorylase. *Analytical Biochemistry*, *352*(1), 145–147.

Fushinobu, S., Kamata, K., Iwata, S., Sakai, H., Ohta, T., & Matsuzawa, H. (1996). Allosteric activation of L-lactate dehydrogenase analyzed by hybrid enzymes with effector-sensitive and -insensitive subunits. *Journal of Biological Chemistry*, *271*(41), 25611–25616.

G

Gardiol, A., & Preiss, J. (1990). Escherichia coli E-39 ADP glucose synthetase has different activation kinetics from the wild-type allosteric enzyme. *Archives of Biochemistry and Biophysics*, *280*(1), 175–180.

Gentner, N., & Preiss, J. (1967). Activator-inhibitor interactions in the adenosine diphosphate glucose pyrophosphorylase of Escherichia coli B. *Biochemical and Biophysical Research Communications*, *27*(3), 417–423.

Geremia, S., Campagnolo, M., Schinzel, R., & Johnson, L. N. (2002). Enzymatic catalysis in crystals of Escherichia coli maltodextrin phosphorylase. *Journal of Molecular Biology*, *322*(2), 413–423.

Geurtsen, J., Chedammi, S., Mesters, J., Driessen, N. N., Sambou, T., Kakutani, R., … Daffe, M. (2017). Identification of Mycobacterial α -Glucan As a Novel Ligand for DC-SIGN: Involvement of Mycobacterial Capsular Polysaccharides in Host Immune Modulation.

Ghosh, H. P., & Preiss, J. (1966). Adenosine diphosphate glucose pyrophosphorylase. A regulatory enzyme in the biosynthesis of starch in spinach leaf chloroplasts. *Journal of Biological Chemistry*, *241*(19), 4491–4504.

Gibbons, B. J., Roach, P. J., & Hurley, T. D. (2002). Crystal structure of the autocatalytic initiator of glycogen biosynthesis, glycogenin. *Journal of Molecular Biology*, *319*(2), 463–477.

Gibson, R. P., Turkenburg, J. P., Charnock, S. J., Lloyd, R., & Davies, G. J. (2002). Insights into trehalose synthesis provided by the structure of the retaining glucosyltransferase OtsA. *Chemistry and Biology*, *9*(12), 1337–1346.

Giroux, M. J., Shaw, J., Barry, G., Cobb, B. G., Greene, T., Okita, T., & 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*(12), 5824–9.

Gómez, H., Lluch, J. M., & Masgrau, L. (2013). Substrate-Assisted and Nucleophilically Assisted Catalysis in Bovine α1,3-Galactosyltransferase. Mechanistic Implications for Retaining Glycosyltransferases. *Journal of the American Chemical Society*, *135*(18), 7053–7063.

Gore, M. G. (2000). *Spectrophotometry and Spectrofluorimetry: a practical approach*. Oxford University Press.

Govons, S., Gentner, N., Greenberg, E., & Preiss, J. (1973). Biosynthesis of bacterial glycogen. XI. Kinetic characterization of an altered adenosine diphosphate-glucose synthase from a “glycogen-excess” mutant of Escherichia coli B. *The Journal of Biological Chemistry*, *248*(5), 1731–40.

Green, O. M., McKenzie, A. R., Shapiro, A. B., Otterbein, L., Ni, H., Patten, A., … Breed, J. (2012). Inhibitors of acetyltransferase domain of N-acetylglucosamine-1-phosphate- uridyltransferase/glucosamine-1-phosphate-acetyltransferase (GlmU). Part 1: Hit to lead evaluation of a novel arylsulfonamide series. *Bioorganic and Medicinal Chemistry Letters*, *22*(4), 1510–1519.

Grigorieff, N. (2007). FREALIGN: High-resolution refinement of single particle structures. *Journal of Structural Biology*, *157*(1), 117–125.

Guerin, M. E., Kordulakova, J., Schaeffer, F., Svetlikova, Z., Buschiazzo, A., Giganti, D., … Alzari, P. M. (2007). Molecular recognition and interfacial catalysis by the essential phosphatidylinositol mannosyltransferase PimA from mycobacteria. *Journal of Biological Chemistry*, *282*(28), 20705–20714.

Guerin, M. E., Schaeffer, F., Chaffotte, A., Gest, P., Giganti, D., Korduláková, J., … Alzari, P. M. (2009). Substrate-induced conformational changes in the essential peripheral membrane-associated mannosyltransferase PimA from mycobacteria. Implications for catalysis. *Journal of Biological Chemistry*, *284*(32), 21613–21625.

H

Haugen, T., Ishaque, A., Chatterjee, A. K., & Preiss, J. (1974). Purification of Escherichia coli ADPglucose pyrophosphorylase by affinity chromatography. *FEBS Letters*, *42*(2), 205–208.

Hicks, J., Wartchow, E., & Mierau, G. (2011). Glycogen storage diseases: a brief review and update on clinical features, genetic abnormalities, pathologic features, and treatment. *Ultrastructural Pathology*, *35*(5), 183–96.

Hill, M. a., Kaufmann, K., Otero, J., & Preiss, J. (1991). Biosynthesis of bacterial glycogen: Mutagenesis of a catalytic site residue of ADP-glucose pyrophosphorylase from Escherichia coli. *Journal of Biological Chemistry*, *266*(19), 12455–12460.

Holme, T. (1957). Continuous Culture Studies on Glycogen Synthesis in Escherichia coli B. *Acta Chem. Scand*.

Hu, Y., Chen, L., Ha, S., Gross, B., Falcone, B., Walker, D., … Walker, S. (2003). Crystal structure of the MurG:UDP-GlcNAc complex reveals common structural principles of a superfamily of glycosyltransferases. *Proc Natl Acad Sci U S A*, *100*(3), 845–849.

I

Iglesias, A. A., & Preiss, J. (1992). Bacterial glycogen and plant starch biosynthesis. *Biochemical Education*, *20*(4), 196–203.

Ikehara, Y., Arai, K., Furukawa, N., Ohno, T., Miyake, T., Fushinobu, S., … Taguchi, H. (2014). The core of allosteric motion in Thermus caldophilus L-lactate dehydrogenase. *Journal of Biological Chemistry*, *289*(45), 31550–31564.

J

Jin, X., Ballicora, M. a, Preiss, J., & Geiger, J. H. (2005). Crystal structure of potato tuber ADP-glucose pyrophosphorylase. *The EMBO Journal*, *24*(4), 694–704.

Jurica, M. S., Mesecar, A., Heath, P. J., Shi, W., Nowak, T., & Stoddard, B. L. (1998). The allosteric regulation of pyruvate kinase by fructose-1,6-bisphosphate. *Structure*, *6*(2), 195–210.

K

Kalscheuer, R., Syson, K., Veeraraghavan, U., Weinrick, B., Biermann, K. E., Liu, Z., … Jacobs, W. R. (2010). Self-poisoning of Mycobacterium tuberculosis by targeting GlgE in an alpha-glucan pathway. *Nature Chemical Biology*, *6*(5), 376–84.

Kanamaru, S., Leiman, P. G., Kostyuchenko, V. A., Chipman, P. R., Mesyanzhinov, V. V., Arisaka, F., & Rossmann, M. G. (2002). Structure of the cell-puncturing device of bacteriophage T4. *Nature*, *415*(6871), 553–557.

Kawaguchi, K., Fox, J., Holmes, E., Boyer, C. D., & Preiss, J. (1978). De novo synthesis of *Escherichia coli* glycogen is due to primer associated with glycogen synthase and activation by branching enzyme. *Archives of Biochemistry and Biophysics*, *190*(2), 385–397.

Kelly, S. M., Jess, T. J., & Price, N. C. (2005). How to study proteins by circular dichroism. *Biochimica et Biophysica Acta - Proteins and Proteomics*, *1751*(2), 119–139.

Kirby, A. J. (2007). Enzymatic Reaction Mechanisms. Von Perry A. Frey und Adrian D. Hegeman. *Angewandte Chemie*, *119*(42), 8068–8070.

Ko, J. H., Kim, C. H., Lee, D. S., & Kim, Y. S. (1996). Purification and characterization of a thermostable ADP-glucose pyrophosphorylase from Thermus caldophilus GK-24. *The Biochemical Journal*, *319 (Pt 3*, 977–83.

Koliwer-Brandl, H., Syson, K., van de Weerd, R., Chandra, G., Appelmelk, B., Alber, M., … Kalscheuer, R. (2016). Metabolic Network for the Biosynthesis of Intra- and Extracellular ??-Glucans Required for Virulence of Mycobacterium tuberculosis. *PLoS Pathogens*, *12*(8), 1–26.

Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., & Svergun, D. I. (2003). *PRIMUS*: a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography*, *36*(5), 1277–1282.

Kornberg, A. (1962). On the metabolic significance of phosphorolytic and pyrophosphorolytic reactions. In H. Kasha and P. Pullman (Ed.), *Horizons in Biochemistry* (pp. 251–264). New York, NY: Academic Press.

Kostrewa, D., D'Arcy, A., Takacs, B., & Kamber, M. (2001). Crystal Structures of Streptococcus pneumoniaeN-Acetylglucosamine-1-phosphate Uridyltransferase, GlmU, in Apo Form at 2.33Å Resolution and in Complex with UDP-N-Acetylglucosamine and Mg2+ at 1.96Å Resolution. *Journal of Molecular Biology*, *305*(2), 279–289.

Krisman, C. R. (1962). A Method for the Calorimetric Estimation of Glycogen with Iodine1. *ANALYTICAL BIOCHEMISTRY*, *4*, 17–23.

L

Lairson, L. L., Henrissat, B., Davies, G. J., & Withers, S. G. (2008). Glycosyltransferases: structures, functions, and mechanisms. *Annual Review of Biochemistry*, *77*, 521–555.

Leadbetter, E. R., & Poindexter, J. (1989). *Bacteria in Nature*. (J. S. Poindexter & E. R. Leadbetter, Eds.). Boston, MA: Springer US.

Lew, C. R., Guin, S., & Theodorescu, D. (2015). Targeting glycogen metabolism in bladder cancer. *Nature Reviews Urology*, *12*(7), 383–391.

Light, S. H., Cahoon, L. A., Halavaty, A. S., Freitag, N. E., & Anderson, W. F. (2016). Structure to function of an α-glucan metabolic pathway that promotes Listeria monocytogenes pathogenesis. *Nature Microbiology*, *2*(November), 16202.

Lindenberger, J. J., Kumar Veleti, S., Wilson, B. N., Sucheck, S. J., & Ronning, D. R. (2015). Crystal structures of Mycobacterium tuberculosis GlgE and complexes with non-covalent inhibitors. *Scientific Reports*, *5*(August), 12830.

Lipscomb, W. N., & Kantrowitz, E. R. (2012). Structure and Mechanisms of Escherichia coli Aspartate Transcarbamoylase. *Accounts of Chemical Research*, *45*(3), 444–453.

Lira-Navarrete, E., Iglesias-Fernández, J., Zandberg, W. F., Compañón, I., Kong, Y., Corzana, F., … Hurtado-Guerrero, R. (2014). Substrate-guided front-face reaction revealed by combined structural snapshots and metadynamics for the polypeptide N-acetylgalactosaminyltransferase 2. *Angewandte Chemie - International Edition*, *53*(31), 8206–8210.

Lodeiro, A. R., Di Lorenzo, O., Petruccelli, S., Molina-Ortiz, S., & Sorgentini, D. (1994). An experiment on glycogen biosynthesis in Escherichia coli. *Biochemical Education*, *22*(4), 213–215.

Lomako, J., Lomako, W. M., & Whelan, W. J. (2004). Glycogenin: The primer for mammalian and yeast glycogen synthesis. *Biochimica et Biophysica Acta - General Subjects*, *1673*(1–2), 45–55.

Ludtke, S. J., Baldwin, P. R., & Chiu, W. (1999). EMAN: semiautomated software for high-resolution single-particle reconstructions. *Journal of Structural Biology*, *128*(1), 82–97.

Lütteke, T., & von der Lieth, C.-W. (2004). pdb-care (PDB carbohydrate residue check): a program to support annotation of complex carbohydrate structures in PDB files. *BMC Bioinformatics*, *5*, 69.

M

MacGregor, E. A., Janecek, S., & Svensson, B. (2001). Relationship of sequence and structure to specificity in the α-amylase family of enzymes. *Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology*, *1546*(1), 1–20.

Manners, D. J. (1991). Recent developments in our understanding of glycogen structure. *Carbohydrate Polymers*, *16*(1), 37–82.

Martinez-Fleites, C., Proctor, M., Roberts, S., Bolam, D. N., Gilbert, H. J., & Davies, G. J. (2006). Insights into the Synthesis of Lipopolysaccharide and Antibiotics through the Structures of Two Retaining Glycosyltransferases from Family GT4. *Chemistry and Biology*, *13*(11), 1143–1152.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007). Phaser crystallographic software. *Journal of Applied Crystallography*, *40*(Pt 4), 658–674.

Milne, J. L. S., Borgnia, M. J., Bartesaghi, A., Tran, E. E. H., Earl, L. A., Schauder, D. M., … Subramaniam, S. (2013). Cryo-electron microscopy - a primer for the non-microscopist. *FEBS Journal*, *280*(1), 28–45.

Mishra, A. K., Driessen, N. N., Appelmelk, B. J., & Besra, G. S. (2011). Lipoarabinomannan and related glycoconjugates: Structure, biogenesis and role in Mycobacterium tuberculosis physiology and host-pathogen interaction. *FEMS Microbiology Reviews*, *35*(6), 1126–1157.

Monod, J., Wyman, J., & Changeux, J. P. (1965). on the Nature of Allosteric Transitions: a Plausible Model. *Journal of Molecular Biology*, *12*(1), 88–118.

Moslemi, A.-R., Lindberg, C., Nilsson, J., Tajsharghi, H., Andersson, B., & Oldfors, A. (2010). Glycogenin-1 deficiency and inactivated priming of glycogen synthesis. *The New England Journal of Medicine*, *362*(13), 1203–1210.

Murshudov, G. N., Skubák, P., Lebedev, A. a., Pannu, N. S., Steiner, R. a., Nicholls, R. a., … Vagin, A. a. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D: Biological Crystallography*, *67*(Pt 4), 355–367.

N

Nakayama, A., Yamamoto, K., & Tabata, S. (2001). Identification of the Catalytic Residues of Bifunctional Glycogen Debranching Enzyme. *Journal of Biological Chemistry*, *276*(31), 28824–28828.

Newcomb, C. J., Moyer, T. J., Lee, S. S., & Stupp, S. I. (2012). Advances in cryogenic transmission electron microscopy for the characterization of dynamic self-assembling nanostructures. *Current Opinion in Colloid & Interface Science*, *17*(6), 350–359.

Newgard, C., & Hwang, P. (1989). The Family of Glycogen Phosphorylases: Structure and Functio. *Critical Reviews in*, *24*(1).

Noguchi, J., Chaen, K., Vu, N. T., Akasaka, T., Shimada, H., Nakashima, T., … Kimura, M. (2011). Crystal structure of the branching enzyme i (BEI) from Oryza sativa L with implications for catalysis and substrate binding. *Glycobiology*, *21*(8), 1108–1116.

O

O'Reilly, M., Watson, K. A., & Johnson, L. N. (1999). The Crystal Structure of the Escherichia coli Maltodextrin Phosphorylase–Acarbose Complex. *Biochemistry*, *38*(17), 5337–5345.

O'Reilly, M., Watson, K. A., Schinzel, R., Palm, D., & Johnson, L. N. (1997). Oligosaccharide substrate binding in Escherichia coli maltodextrin phosphorylase. *Nature Structural Biology*, *4*(5), 405–12.

Ohi, M., Li, Y., Cheng, Y., & Walz, T. (2004). Negative Staining and Image Classification - Powerful Tools in Modern Electron Microscopy. *Biological Procedures Online*, *6*(1), 23–34.

Oikonomakos, N. G., Chrysina, E. D., Kosmopoulou, M. N., & Leonidas, D. D. (2003). Crystal structure of rabbit muscle glycogen phosphorylase a in complex with a potential hypoglycaemic drug at 2.0 Å resolution. *Biochimica et Biophysica Acta - Proteins and Proteomics*, *1647*(1–2), 325–332.

Orlova, E. V, & Saibil, H. R. (2011). Structural analysis of macromolecular assemblies by electron microscopy. *Chemical Reviews*, *111*(12), 7710–48.

Ormö, M., Bystrom, C. E., & Remington, S. J. (1998). Crystal structure of a complex of Escherichia coli glycerol kinase and an allosteric effector fructose 1,6-bisphosphate. *Biochemistry*, *37*(47), 16565–16572.

P

Pain, R. (2005). Determining the CD Spectrum Of A Protein. *Handbook of Food Analytical Chemistry*, *1–2*(Cd), 219–243.

Pal, K., Kumar, S., Sharma, S., Garg, S. K., Alam, M. S., Xu, H. E., … Swaminathan, K. (2010). Crystal structure of full-length Mycobacterium tuberculosis H37Rv glycogen branching enzyme: Insights of N-terminal β-sandwich in substrate specificity and enzymatic activity. *Journal of Biological Chemistry*, *285*(27), 20897–20903.

Palomo, M., Pijning, T., Booiman, T., Dobruchowska, J. M., Van Der Vlist, J., Kralj, S., … Leemhuis, H. (2011). Thermus thermophilus glycoside hydrolase family 57 branching enzyme: Crystal structure, mechanism of action, and products formed. *Journal of Biological Chemistry*, *286*(5), 3520–3530.

Park, H.-S., Park, J.-T., Kang, H.-K., Cha, H., Kim, D.-S., Kim, J.-W., & Park, K.-H. (2007). TreX from Sulfolobus solfataricus ATCC 35092 displays isoamylase and 4-alpha-glucanotransferase activities. *Bioscience, Biotechnology, and Biochemistry*, *71*(5), 1348–1352.

Park, J.-T., Park, H.-S., Kang, H.-K., Hong, J.-S., Cha, H., Woo, E.-J., … Park, K.-H. (2008). Oligomeric and functional properties of a debranching enzyme (TreX) from the archaeon Sulfolobus solfataricus P2. *Biocatalysis and Biotransformation*, *26*(1–2), 76–85.

Park, J. T., Shim, J. H., Tran, P. L., Hong, I. H., Yong, H. U., Oktavina, E. F., … Park, K. H. (2011). Role of maltose enzymes in glycogen synthesis by Escherichia coli. *Journal of Bacteriology*, *193*(10), 2517–2526.

Parsons, T. F., & Preiss, J. (1978a). 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. *Journal of Biological Chemistry*, *253*(17), 6197–6202.

Parsons, T. F., & Preiss, J. (1978b). 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. *The Journal of Biological Chemistry*, *253*(17), 6197–202.

Patenaude, S. I., Seto, N. O. L., Borisova, S. N., Szpacenko, A., Marcus, S. L., Palcic, M. M., & Evans, S. V. (2002). The structural basis for specificity in human ABO(H) blood group biosynthesis. *Nature Structural Biology*, *9*(9), 685–690.

Paule, M. R., & Preiss, J. (1971). Biosynthesis of Bacterial Glycogen. X. THE KINETIC MECHANISM OF ADENOSINE DIPHOSPHOGLUCOSE PYROPHOSPHORYLASE FROM RHODOSPIRILLUM RUBRUM. *J. Biol. Chem.*, *246*(14), 4602–4609.

Pei, J., Kim, B. H., & Grishin, N. V. (2008). PROMALS3D: A tool for multiple protein sequence and structure alignments. *Nucleic Acids Research*, *36*(7), 2295–2300.

Pelissier, M. C., Lesley, S. a., Kuhn, P., & Bourne, Y. (2010). Structural insights into the catalytic mechanism of bacterial guanosine-diphospho-D-mannose pyrophosphorylase and its regulation by divalent ions. *Journal of Biological Chemistry*, *285*(35), 27468–27476.

Persson, K., Ly, H. D., Dieckelmann, M., Wakarchuk, W. W., Withers, S. G., & Strynadka, N. C. (2001). Crystal structure of the retaining galactosyltransferase LgtC from Neisseria meningitidis in complex with donor and acceptor sugar analogs. *Nature Structural Biology*, *8*(2), 166–175.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*, 1605–1612.

Pfister, B., & Zeeman, S. C. (2016). Formation of starch in plant cells. *Cellular and Molecular Life Sciences*, *73*(14), 2781–2807.

Preiss, J. (1978). Regulation of adenosine diphosphate glucose pyrophosphorylase. *Advances in Enzymology and Related Areas of Molecular Biology*, *46*, 317–381.

Preiss, J. (1984). Bacterial Glycogen Synthesis and its Regulation. *Ann. Rev. Microbiol.*, *38*, 419–458.

Preiss, J. (2006). Bacterial Glycogen Inclusions : Enzymology and Regulation of Synthesis. *Microbial Monogr*, *1*, 71–108.

Preiss, J. (2010). 6.15 - Biochemistry and Molecular Biology of Glycogen Synthesis in Bacteria and Mammals and Starch Synthesis in Plants A2 - Liu, Hung-Wen (Ben). *Comprehensive Natural Products II*, 429–491.

Preiss, J. (2014). Glycogen: Biosynthesis and Regulation. *EcoSal Plus*, *6*(1).

Preiss, J., Greenberg, E., & Sabraw, a. (1975). Biosynthesis of bacterial glycogen. Kinetic studies of a glucose-1-phosphate adenylyltransferase (EC 2.7.7.27) from a glycogen-deficient mutant of Escherichia coli B. *The Journal of Biological Chemistry*, *250*(19), 7631–7638.

Preiss, J., Lammel, C., & Greenberg, E. (1976). Biosynthesis of bacterial glycogen. *Archives of Biochemistry and Biophysics*, *174*(1), 105–119.

Preiss, J., Shen, L., & Partridger, M. (1965). THE ACTIVATION OF ESCHERICHIA COLI ADP-GLUCOSE PYROPHOSPHORYLASE. *Biochemical and Biophysical Research Communications*, *18*, 180–5.

Preiss, J., Yung, S., & Baecker, P. a. (1983). Regulation of bacterial glycogen synthesis. *Molecular and Cellular Biochemistry*, *57*(1), 61–80.

R

Raetz, C. R., & Roderick, S. L. (1995). A left-handed parallel beta helix in the structure of UDP-N-acetylglucosamine acyltransferase. *Science (New York, N.Y.)*, *270*(5238), 997–1000.

Rambo, R. P., & Tainer, J. A. (2011). Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers*, *95*(8), 559–571.

Rao, S. T., & Rossmann, M. G. (1973). Comparison of super-secondary structures in proteins. *Journal of Molecular Biology*, *76*(2), 241–256.

Rashid, A. M., Batey, S. F. D., Syson, K., Koliwer-Brandl, H., Miah, F., Barclay, J. E., … Bornemann, S. (2016). Assembly of α-Glucan by GlgE and GlgB in Mycobacteria and Streptomycetes. *Biochemistry*, *55*(23), 3270–3284.

Recondo, E., & Leloir, L. F. (1961). Adenosine diphosphate glucose and starch synthesis. *Biochemical and Biophysical Research Communications*, *6*(2), 85–88.

Ribéreau-Gayon, G., Sabraw, A., Lammel, C., & Preiss, J. (1971). Biosynthesis of bacterial glycogen IX: regulatory properties of the adenosine diphosphate glucose pyrophosphrylases of the Enterobacterieae. *Archives of Biochemistry and Biophysics*, *142*(2), 675–92.

Richard H Baltz, Julian Davies, A. L. D. (2010). *Manual of Industrial Microbiology and Biotechnology, Third Edition*. (L. R. Lynd, H. Zhao, L. Katz, R. H. Baltz, A. T. Bull, B. Junker, … A. L. Demain, Eds.) (Third edit). American Society of Microbiology.

Roach, P. J., Depaoli-Roach, A. A., Hurley, T. D., & Tagliabracci, V. S. (2012a). Glycogen and its metabolism: some new developments and old themes. *The Biochemical Journal*, *441*(3), 763–87.

Roach, P. J., Depaoli-Roach, A. A., Hurley, T. D., & Tagliabracci, V. S. (2012b). Glycogen and its metabolism: some new developments and old themes. *The Biochemical Journal*, *441*(3), 763–87.

Rohou, A., & Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of Structural Biology*, *192*(2), 216–221.

Rojas-Cervellera, V., Ardèvol, A., Boero, M., Planas, A., & Rovira, C. (2013). Formation of a covalent glycosyl-enzyme species in a retaining glycosyltransferase. *Chemistry (Weinheim an Der Bergstrasse, Germany)*, *19*(42), 14018–14023.

Romeo, T., Kumar, A., & Preiss, J. (1988). Analysis of the ~ sc ~ e ~ jc ~ ju cofi glyeogen gene cluster suggests that catabolic enzymes are encoded among the biosynthetic genes (Nucleotide sequencing ; computer search of protein sequence ; glycogen branching enzyme ; glucan transferase ; amylase, *70*, 363–376.

Roy, R., Usha, V., Kermani, A., Scott, D. J., Hyde, E. I., Besra, G. S., … Fütterer, K. (2013). Synthesis of α-glucan in mycobacteria involves a hetero-octameric complex of trehalose synthase tres and maltokinase Pep2. *ACS Chemical Biology*, *8*(10), 2245–2255.

S

Schinzel, R., & Nidetzky, B. (1999). Bacterial alpha-glucan phosphorylases. *FEMS Microbiology Letters*, *171*(2), 73–9.

Schüttelkopf, A. W., & Van Aalten, D. M. F. (2004). PRODRG: A tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallographica Section D: Biological Crystallography*, *60*(Pt 8), 1355–1363.

Sean Froese, D., Michaeli, A., McCorvie, T. J., Krojer, T., Sasi, M., Melaev, E., … Yue, W. W. (2015). Structural basis of glycogen branching enzyme deficiency and pharmacologic rescue by rational peptide design. *Human Molecular Genetics*, *24*(20), 5667–5676.

Seok, Y. J., Sondej, M., Badawi, P., Lewis, M. S., Briggs, M. C., Jaffe, H., & Peterkofsky, A. (1997). High affinity binding and allosteric regulation of Escherichia coli glycogen phosphorylase by the histidine phosphocarrier protein, HPr. *Journal of Biological Chemistry*, *272*(42), 26511–26521.

Shaikh, T. R., Gao, H., Baxter, W. T., Asturias, F. J., Leith, A., Frank, J., & Lourmel, R. De. (2009). Spider. *Nat Protoc*, *3*(12), 1941–1974.

Shearer, J., & Graham, T. E. (2002). New perspectives on the storage and organization of muscle glycogen. *Canadian Journal of Applied Physiology = Revue Canadienne de Physiologie Appliquee*, *27*(2), 179–203.

Sheng, F., Jia, X., Yep, A., Preiss, J., & Geiger, J. H. (2009). The crystal structures of the open and catalytically competent closed conformation of Escherichia coli glycogen synthase. *Journal of Biological Chemistry*, *284*(26), 17796–17807.

Sheng, F., Yep, A., Feng, L., Preiss, J., & Geiger, J. H. (2009). Oligosaccharide binding in Escherichia coli glycogen synthase. *Biochemistry*, *48*(42), 10089–10097.

Shimizu, R., Kataoka, Y., Kawai, S., & Tanaka, T. (1975). LaB6 single-crystal tips as an electron source of high brightness. *Applied Physics Letters*, *27*(3), 113–114.

Shively, J. M. (1974). Inclusion Bodies of Prokaryotes. *Annual Review of Microbiology*, *28*(1), 167–188.

Singh, S., Phillips, G. N., & Thorson, J. S. (2012). The structural biology of enzymes involved in natural product glycosylation. *Natural Product Reports*, *29*(10), 1201–37.

Sinnott, M. L., & Jencks, W. P. (1980). Solvolysis of D-glucopyranosyl derivatives in mixtures of ethanol and 2,2,2-trifluoroethanol. *Journal of the American Chemical Society*, *102*(6), 2026–2032.

Sivaraman, J., Sauvé, V., Matte, A., & Cygler, M. (2002). Crystal structure of Escherichia coli glucose-1-phosphate thymidylyltransferase (RffH) complexed with dTTP and Mg2+. *Journal of Biological Chemistry*, *277*(46), 44214–44219.

Song, H. N., Jung, T. Y., Park, J. T., Park, B. C., Myung, P. K., Boos, W., … Park, K. H. (2010). Structural rationale for the short branched substrate specificity of the glycogen debranching enzyme GlgX. *Proteins*, *78*(8), 1847–1855.

Sotomayor-Pérez, A. C., Subrini, O., Hessel, A., Ladant, D., & Chenal, A. (2013). Molecular crowding stabilizes both the intrinsically disordered calcium-free state and the folded calcium-bound state of a repeat in toxin (RTX) protein. *Journal of the American Chemical Society*, *135*(32), 11929–11934.

Sprang, S., Goldsmith, E., & Fletterick, R. (1987). Structure of the nucleotide activation switch in glycogen phosphorylase a. *Science (New York, N.Y.)*, *237*(4818), 1012–1019.

Stark, D. M., Timmerman, K. P., Barry, G. F., Preiss, J., & Kishore, G. M. (1992). Regulation of the Amount of Starch in Plant Tissues by ADP Glucose Pyrophosphorylase. *Science (New York, N.Y.)*, *258*(April), 287–292.

Steiner, K. E., & Preiss, J. (1977). Biosynthesis of bacterial glycogen: genetic and allosteric regulation of glycogen biosynthesis in Salmonella typhimurium LT-2. *Journal of Bacteriology*, *129*(1), 246–53.

Steitz, T. A., Smerdon, S. J., Jäger, J., & Joyce, C. M. (1994). A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. *Science (New York, N.Y.)*, *266*(5193), 2022–2025.

Strange, R. E. (1968). Bacterial “Glycogen” and Survival. *Nature*, *220*(5167), 606–607.

Suzuki, E., & Suzuki, R. (2016). Distribution of glucan-branching enzymes among prokaryotes. *Cellular and Molecular Life Sciences*, *73*(14), 2643–2660.

Svergun, D. I. (1992). Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria. *Journal of Applied Crystallography*, *25*, 495–503.

Svergun, D. I., Petoukhov, M. V, & Koch, M. H. (2001). Determination of domain structure of proteins from X-ray solution scattering. *Biophysical Journal*, *80*(6), 2946–2953.

Sweetlove, L. J., Burrell, M. M., & ap Rees, T. (1996). Starch metabolism in tubers of transgenic potato (Solanum tuberosum) with increased ADPglucose pyrophosphorylase. *Biochemical Journal*, *320*(Pt 2), 493–498.

Swift, R. V, Ong, C. D., & Amaro, R. E. (2012). Magnesium-induced nucleophile activation in the guanylyltransferase mRNA capping enzyme. *Biochemistry*, *51*(51), 10236–43.

T

Tao, B. Y., Reilly, P. J., & Robyt, J. F. (1989). Detection of a covalent intermediate in the mechanism of action of porcine pancreatic alpha-amylase by using 13C nuclear magnetic resonance. *Biochimica et Biophysica Acta*, *995*(3), 214–220.

Tetlow, I. J., & Emes, M. J. (2014). A review of starch-branching enzymes and their role in amylopectin biosynthesis. *IUBMB Life*, *66*(8), 546–558.

Thompson, R. F., Walker, M., Siebert, C. A., Muench, S. P., & Ranson, N. A. (2016). An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. *Methods*, *100*, 3–15.

Trivelloni, J. C., Recondo, E., & Cardini, C. E. (1962). Adenosine Diphosphate Glucose and Glucoside Biosynthesis. *Nature*, *195*(4847), 1202–1202.

Tuncel, A., & Okita, T. W. (2013). Improving starch yield in cereals by over-expression of ADPglucose pyrophosphorylase: Expectations and unanticipated outcomes. *Plant Science*, *211*, 52–60.

U

Ugalde, J. E., Parodi, A. J., & Ugalde, R. a. (2003). De novo synthesis of bacterial glycogen: Agrobacterium tumefaciens glycogen synthase is involved in glucan initiation and elongation. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(19), 10659–10663.

Uitdehaag, J. C. M., Mosi, R., Kalk, K. H., van der Veen, B. A., Dijkhuizen, L., Withers, S. G., & Dijkstra, B. W. (1999). X-ray structures along the reaction pathway of cyclodextrin glycosyltransferase elucidate catalysis in the α-amylase family. *Nat Struct Mol Biol*, *6*(5), 432–436.

Urresti, S., Albesa-Jove, D., Schaeffer, F., Pham, H. T., Kaur, D., Gest, P., … Guerin, M. E. (2012). Mechanistic insights into the retaining glucosyl-3-phosphoglycerate synthase from mycobacteria. *Journal of Biological Chemistry*, *287*(29), 24649–24661.

V

Valentini, G., Chiarelli, L., Fortini, R., Speranza, M. L., Galizzi, A., & Mattevi, A. (2000). The allosteric regulation of pyruvate kinase: A site-directed mutagenesis study. *Journal of Biological Chemistry*, *275*(24), 18145–18152.

Vestergaard, B. (2016). Analysis of biostructural changes, dynamics, and interactions - Small-angle X-ray scattering to the rescue. *Archives of Biochemistry and Biophysics*, *602*, 69–79.

Vetting, M. W., Frantom, P. A., & Blanchard, J. S. (2008). Structural and enzymatic analysis of MshA from Corynebacterium glutamicum: Substrate-assisted catalysis. *Journal of Biological Chemistry*, *283*(23), 15834–15844.

Vithani, N., Bais, V., & Prakash, B. (2014). GlmU (N -acetylglucosamine-1-phosphate uridyltransferase) bound to three magnesium ions and ATP at the active site. *Acta Crystallographica Section F Structural Biology Communications*, *70*(6), 703–708.

W

Watson, K. A., McCleverty, C., Geremia, S., Cottaz, S., Driguez, H., & Johnson, L. N. (1999). Phosphorylase recognition and phosphorolysis of its oligosaccharide substrate: Answers to a long outstanding question. *EMBO Journal*, *18*(17), 4619–4632.

Watson, K. A., Schinzel, R., Palm, D., & Johnson, L. N. (1997). The crystal structure of Escherichia coli maltodextrin phosphorylase provides an explanation for the activity without control in this basic archetype of a phosphorylase. *EMBO Journal*, *16*(1), 1–14.

Wattam, A. R., Abraham, D., Dalay, O., Disz, T. L., Driscoll, T., Gabbard, J. L., … Sobral, B. W. (2014). PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Research*, *42*(D1), 581–591.

Wilkinson, J. F. (1959). The problem of energy-storage compounds in bacteria. *Experimental Cell Research*, *7*, 111–130.

Wilkinson, J. F. (1963). Carbon and Energy Storage in Bacteria. *Journal of General Microbiology*, *32*(1963), 171–176.

Williams, D. B., & Carter, C. B. (2009). *Transmission Electron Microscopy*. *Methods in molecular biology (Clifton, N.J.)* (Vol. 816). Boston, MA: Springer US.

Wilson, W. A., Roach, P. J., Montero, M., Baroja-Fernández, E., Muñoz, F. J., Eydallin, G., … Pozueta-Romero, J. (2010). Regulation of glycogen metabolism in yeast and bacteria. *FEMS Microbiology Reviews*, *34*(6), 952–985.

Woo, E. J., Lee, S., Cha, H., Park, J. T., Yoon, S. M., Song, H. N., & Park, K. H. (2008). Structural insight into the bifunctional mechanism of the glycogen-debranching enzyme TreX from the archaeon Sulfolobus solfataricus. *Journal of Biological Chemistry*, *283*(42), 28641–28648.

Wu, M. X., & Preiss, J. (1998). The N-terminal region is important for the allosteric activation and inhibition of the Escherichia coli ADP-glucose pyrophosphorylase. *Archives of Biochemistry and Biophysics*, *358*(1), 182–188.

Wu, M. X., & 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. *Archives of Biochemistry and Biophysics*, *389*(2), 159–165.

Wycuff, D. R., & Matthews, K. S. (2000). Generation of an AraC-araBAD Promoter-Regulated T7 Expression System. *Analytical Biochemistry*, *277*(1), 67–73. http://doi.org/10.1006/abio.1999.4385

Y

Yep, A., Ballicora, M. A., & 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*(3), 960–966.

Yung, S. G., & Preiss, J. (1982). Biosynthesis of bacterial glycogen: Purification and structural and immunological properties of Rhodopseudomonas sphaeroides ADPGlucose synthetase. *Journal of Bacteriology*, *151*(2), 742–749.

Z

Zhai, L., Feng, L., Xia, L., Yin, H., & Xiang, S. (2016). Crystal structure of glycogen debranching enzyme and insights into its catalysis and disease-causing mutations. *Nature Communications*, *7*, 11229.

ANNEX I

ANNEX II