Filipa Alexandra Afonso Lopes

Licenciada em Bioquímica

Diabetes Type II Regulation – Inhibitory Potential of *hl*GP with Different Natural Compounds

Dissertação para obtenção do Grau de Mestre em Bioquímica

Orientador: Doutora Márcia Alexandra da Silva Correia, Investigadora, FCT/UNL

Co-orientador: Doutora Maria Angelina de Sá Palma, Investigadora Auxiliar, FCT/UNL

Júri:

Presidente: Doutora Sofia Rocha Pauleta
Arguente: Doutora Maria Leonor Carvalho Morgado
Vogal: Doutora Márcia Alexandra da Silva Correia

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Dezembro de 2020

Lombada

2020

Diabetes Type II Regulation – Inhibitory Potential of *hl*GP with Different Natural Compounds

Filipa Alexandra Afonso Lopes

Filipa Alexandra Afonso Lopes

Licenciada em Bioquímica

Diabetes Type II Regulation – Inhibitory Potential of *hl*GP with Different Natural Compounds

Dissertação para obtenção do Grau de Mestre em Bioquímica

Orientador: Doutora Márcia Alexandra da Silva Correia,
Investigadora, FCT/UNL

Co-orientador: Doutora Maria Angelina de Sá Palma,
Investigadora Auxiliar, FCT/UNL

Júri:

Presidente: Doutora Sofia Rocha Pauleta
Arguente: Doutora Maria Leonor Carvalho Morgado
Vogal: Doutora Márcia Alexandra da Silva Correia

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Dezembro de 2020

Copyright

Ao longo do desenvolvimento desta dissertação foi-me possível:

- Participar nas **Jornadas Intercalares dos Mestrados do DQ e do DCV**, onde apresentei o balanço inicial do trabalho, definindo objetivos e traçando perspetivas para o trabalho futuro.
- Obter uma **Bolsa de Investigação** de 3 meses financiada pela Fundação para a Ciência e a Tecnologia - Ministério da Ciência, Tecnologia e Ensino Superior no âmbito do programa Verão com Ciência.

Comunicações orais:

- Jornadas intercalares dos Mestrados do DQ e do DCV – "Diabetes type II regulation - inhibitory potential of *hl*GP with different natural compounds" – 14 de fevereiro de 2020.
- Simpósio do programa Verão com Ciência – "Diabetes type II regulation - inhibitory potential of *hl*GP with different natural compounds" – 21 de outubro de 2020.

Artigos Científicos:

- O trabalho desenvolvido foi crucial para a escrita de dois manuscritos, um deles aceite para publicação na revista Bioorganic Chemistry (Impact Factor: 4.831) e outro ainda em preparação.

Natércia F Brás, Rui P. P. Neves, **Filipa A.A. Lopes**, Márcia A. S. Correia, Angelina S. Palma, Sérgio F Sousa, Maria J. Ramos. – "Combined in Silico and in Vitro Studies to Identify Novel Antidiabetic Flavonoids Targeting Glycogen Phosphorylase", Bioorganic Chemistry (*in press*). Article reference: YBIOO_BIOORG-D-20-00737.

Filipa A.A. Lopes, Márcia A. S. Correia, Angelina S. Palma. – "A Novel Strategy for Heterologous Expression and Purification of Human Liver Glycogen Phosphorylase" (*in preparation*).

Agradecimentos

A realização desta dissertação não teria sido possível sem o apoio e esforço de diversas pessoas que me acompanharam e foram fundamentais para a concretização de mais um sonho. Manifesto assim toda a minha gratidão a todos aqueles que, direta ou indiretamente, contribuíram para que esta tarefa se tornasse realidade.

Em primeiro lugar, deixo um especial agradecimento à Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa por me ter recebido e proporcionado um excelente ambiente para o meu desenvolvimento em termos académicos e pessoais, e por ter sido a minha segunda casa nestes últimos 5 anos.

Às minhas orientadoras, Doutora Márcia Correia e Doutora Angelina Palma, manifesto uma imensa gratidão não só pela permanente orientação, apoio e preocupação, mas também pela abertura de horizontes de conhecimento que me fizeram crescer a nível pessoal e profissional.

À Professora Maria João Romão, agradeço por me ter dado a oportunidade de realizar o meu trabalho no seu laboratório de investigação.

À Elisabete Ferreira, técnica do laboratório de análises, sou grata por toda a ajuda e disponibilidade para a realização dos ensaios de MST.

A todos os membros dos grupos de Cristalografia de Macromoléculas e Glycolab, agradeço por sempre se disponibilizarem a ajudar e proporcionarem um excelente ambiente de trabalho.

Aos meus colegas de mestrado, André Favinha, Liliana Ringler e Raquel Gama, por todas as conversas, apoio e bons momentos que passámos juntos.

Aos colaboradores Doutora Natércia Brás e Doutor Rui Neves da Faculdade de Ciências da Universidade do Porto.

Ao Professor Ricardo Franco pela coordenação do Mestrado em Bioquímica

À Fundação para a Ciência e Tecnologia - Ministério da Ciência, Tecnologia e Ensino Superior pelo financiamento no âmbito da Unidade de Investigação UCIBIO (UID/Multi/04378/2019); e pela bolsa de investigação de 3 meses no âmbito da Escola de Verão UCIBIO, em colaboração com a Direção Geral do Ensino Superior (DGES).

À Joana Saraiva, Rita Baptista e Sara Brito por todos os almoços fantásticos, conversas, apoio e momentos que passámos juntas e por se mostrarem sempre disponíveis para ajudar quando mais precisei.

Aos meus pais por todo o apoio, dedicação e amor que sempre me dedicaram. Sem vocês, não teria chegado onde cheguei hoje e não seria a pessoa que sou.

À minha irmã Sofia, a minha alma gémea, por caminhar sempre ao meu lado, pela sua paciência, amor e amizade sem igual. Obrigada por todos os momentos que passamos juntas, pelas opiniões e força que sempre me deste.

Aos meus avós que, mesmo sem saberem muito bem o que faço, sempre me apoiaram incondicionalmente.

O importante de alcançar os nossos objetivos, é relembrar aqueles que nos ajudaram no caminho. A todos um grande obrigada.

Resumo

Os diabetes melitus são uma doença metabólica caracterizada pela diminuição da produção ou resistência celular à insulina que leva ao aumento dos níveis de glucose na corrente sanguínea (hiperglicemia). A Glicogénio Fosforilase (GP) é uma enzima-chave envolvida na cascata responsável pela degradação do glicogénio em glucose-1-fosfato. Esta enzima é considerada um alvo validado para o controlo dos diabetes, uma vez que alguns inibidores específicos para a GP diminuem os níveis de glucose no sangue *in vivo*.

Um estudo computacional, realizado pelos colaboradores da Faculdade de Ciências da Universidade do Porto, identificou possíveis compostos com elevada capacidade inibitória para a GP, sendo o principal objetivo da dissertação a validação experimental destes dados.

Assim, o primeiro passo foi a otimização da expressão heteróloga e purificação da GP do fígado humano para obter a proteína recombinante na sua forma menos ativa (GPb) e na sua forma mais ativa (GPa) através de fosforilação. Foram efetuados vários ensaios de TSA que foram determinantes na escolha das melhores condições para a estabilização de ambas as proteínas.

Os ensaios de Termoforese em Microescala (MST) permitiram quantificar as interações da GPb humana e vários compostos. As constantes de dissociação, K_d, foram determinadas para os flavonóides genisteina (1.2 ± 0.4 μM), 8-prenylnaringenin (4.1 ± 1.2 μM) e 8-prenylgenistein (7.4 ± 3.7 μM). Estes resultados mostraram-se concordantes com os dados teóricos calculados por Dinâmica Molecular e Perturbação de Energia Livre, validando a afinidade da ligação da GP para estes compostos.

Ensaios cinéticos para determinar a atividade enzimática ainda estão a decorrer, mas ensaios preliminares com as duas formas da enzima mostraram que a GPa apresenta uma elevada atividade quando comparada com a GPb que não apresentou atividade. Um ensaio com a GPa e a cafeina (inibidor) mostrou redução da atividade da enzima. Estes resultados revelaram-se bastante promissores e pretende-se de futuro determinar os parâmetros V0, Vm e Km da reação, assim como determinar as constates de inibição da proteína com os diferentes inibidores para acompanhar os dados de MST.

Os dados teóricos e experimentais sugerem que os flavonóides genistein, 8PN e 8PG são potenciais compostos inibitórios para a GP e que podem, no futuro, ser explorados para o tratamento dos diabetes.

Termos chave: Diabetes Tipo II, Glicogénio Fosforilase, Inibição, Termoforese em Microescala (MST)

Abstract

Diabetes melitus is a metabolic disease characterized by a lack of insulin production and/or increased cellular resistance to insulin, leading to high levels of glucose in bloodstream (hyperglycaemia). Glycogen phosphorylase (GP) is a key enzyme involved in the cascade responsible for the breakdown of glycogen to glucose-1-phosphate. GP is a validated target for the treatment of diabetes as some GP-specific inhibitors lower blood glucose levels *in vivo*.

A computational study carried out by our collaborators at the Faculty of Sciences, University of Porto, has identified possible compounds with high inhibitory capacity for GP. The main aim of this dissertation was to follow up these compounds and experimentally validate the theoretical data.

To achieve this, the first objective was the optimization of the heterologous expression and purification of human liver GP to obtain the recombinant protein in its less active form (GPb), and in its most active form (GPa) through phosphorylation. Several thermal shift assays (TSA) were carried out and were decisive in choosing the best conditions for the stabilization of both proteins.

Microscale Thermophoresis (MST) was applied to quantify the interactions of human GPb and several compounds. Dissociation constants, K_d, were determined for the flavonoids genistein (1.2 ± 0.4 µM), 8-prenylnaringenin (4.1 ± 1.2 µM) and 8-prenylgenistein (7.4 ± 3.7 µM). Overall, these results agree with the theoretical data calculated by Molecular Dynamics and Free-energy perturbations, thus validating the binding affinity of GP for these compounds.

Kinetic assays to determine the enzyme activity are ongoing, but preliminary assays with the two forms of the enzyme showed that GPa has a high activity while GPb did not show activity. An assay with GPa and caffeine (inhibitor) showed reduced enzyme activity. These results have shown to be quite promising and it is intended in the future to determine the parameters V0, Vm and Km of the reaction, and to determine the protein inhibition constants with the different inhibitors to follow up the MST data.

The combined theoretical and experimental data suggested the flavonoids genistein, 8PN and 8PG as potential inhibitory compounds for GP and may, in the future, be explored for the treatment of diabetes.

Keywords: Type II Diabetes, Glycogen Phosphorylase, Inhibition, MicroScale Thermophoresis

Index

Figure index

Table index

Abbreviations and symbols

8PG – 8-prenylgenistein
8PN – 8-prenylnaringenin
AMP – Adenosine monophosphate
ATP – Adenosine triphosphate
DHAP – Dihydroxyacetone phosphate
DTT – Dithiothreitol
DMSO – Dimethyl Sulfoxide
EDTA – Ethylenediamine tetraacetic acid
FT – Flow through
G1P – Glucose - 1 - phosphate
G6P – Glucose - 6 - phosphate
GP – Glycogen Phosphorylase
GS – Glycogen Synthase
HEPES – 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
***hl*GP** – Human Liver Glycogen Phosphorylase
IMAC – Immobilized Metal Affinity Chromatography
IPTG – Isopropyl β-d-1-thiogalactopyranoside
MAD – Multi-wavelength anomalous diffraction
MES – 2-(N-morpholino)ethanesulfonic acid
MIR – Multiple isomorphous replacement
MOPS – 3-(N-morpholino) propanesulfonic acid
MST – MicroScale Thermophoresis
MR – Molecular Replacement
OD – Optical Density
PDB – Protein Data Bank
PEG – Polyethylene glycol
Phk – Phosphorylase kinase
pI – Isoelectric point
PLP – Pyridoxal phosphate
PKA – Protein kinase A
Rpm – Rotation per minute
SAD – Single-wavelength anomalous diffraction
SDS-PAGE – Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
SEC – Size Exclusion Chromatography
T1DM – Type 1 Diabetes Mellitus
T2DM – Type 2 Diabetes Mellitus
T_M – Melting Temperature
TSA – Thermal Shift Assay
Tris – Tris(hydroxymethyl)aminomethane

1. Introduction

1.1. Diabetes mellitus – a global epidemic

Diabetes mellitus is a complex disease that occurs when the body cannot produce any or enough insulin, or cannot respond the insulin it produces. Insulin is an important hormone produced by the pancreas that allows glucose from the bloodstream to enter the body´s cells. The lack of insulin or the inability of cells to respond to it, leads to high levels of blood glucose, leading to hyperglycaemia, which is the clinical indicator of diabetes [1].

Hyperglycaemia is the most common effect of patients with uncontrolled diabetes and over time can affect numerous organs, such as the kidneys, heart, eyes, blood vessels and nerves. Adults with diabetes have a two-to three-fold increased risk of heart attacks, strokes and chance of foot ulcers that leads to limb amputation [2].

In Portugal, over 60 000 new cases are diagnosed every year. The International Diabetes Federation (IDF) estimated that 463 million adults aged 20-79 years were living with diabetes in 2019 (Figure 1.1.1). The total number is predicted to rise to 700 million by 2045 and the largest increases will come from the regions still developing. It was also estimated that approximately 4.2 million adults aged 20-79 years died as a result of diabetes and its complications in 2019. This is the equivalent to one death every eight seconds, making diabetes the fifth leading cause of death in the world [1].

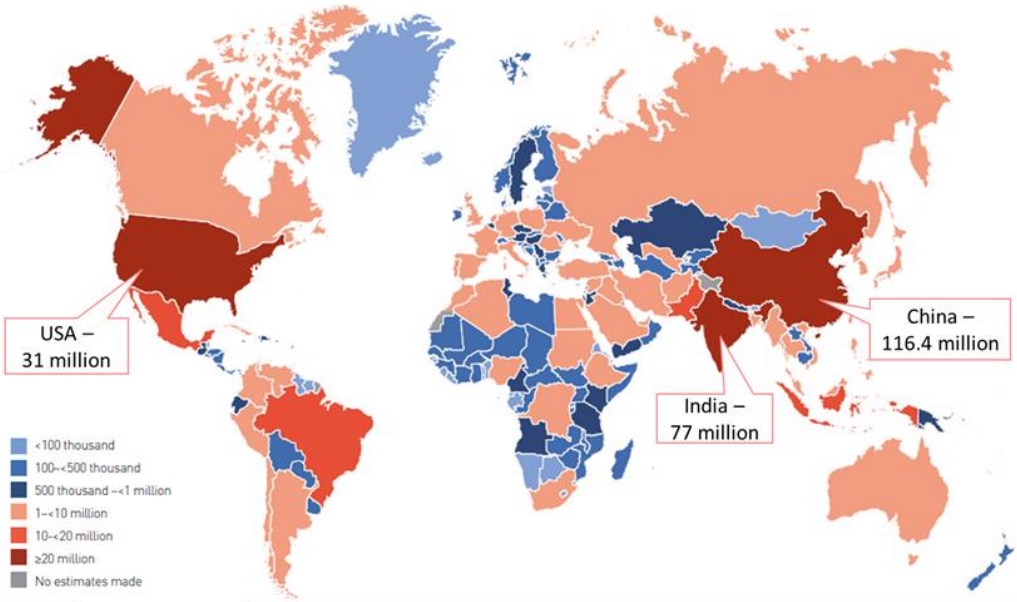


Figure 1.1.1 – Diabetes mellitus world map. Estimated total number of adults (20–79 years) living with diabetes mellitus, highlighting the top three countries or territories for number of adults with diabetes mellitus (20–79 years) in 2019. It was estimated that in 2019, 425 million adults aged 20–79 years had diabetes mellitus worldwide, being China, India and the USA the countries with major incidence. The colour of the country or territory in the map relates to the total number of adults aged 20–79 years living with diabetes mellitus in the area [1]. (Adapted from IDF Atlas)

There are two main categories of diabetes mellitus: type 1 and type 2.

Type 1 diabetes mellitus (T1DM) is a chronic condition in which the immune system attacks the glucose-responsive insulin-secreting β cells in the pancreas that would normally produce insulin. Due to this, glucose cannot enter the cells and remains in the bloodstream. In a way to control the high

blood sugar levels, the patients need to administrate insulin few times a day for the rest of their lives to avoid dangerous complications. Type 1 diabetes can occur at any age, however it is more common in children and young adults [1].

Type 2 diabetes mellitus (T2DM) is characterized by insulin resistance where the body cells do not fully respond to insulin. Because insulin cannot work properly, blood glucose levels continue to rise and more insulin is released. This type of diabetes is the most common, representing about 90% of cases worldwide [1], and is the one that will be study in this thesis.

1.2. Type 2 diabetes mellitus (T2DM)

1.2.1. Pathophysiology and molecular mechanism

When the body produces insulin in conditions of insulin resistance, the cells have less capacity for absorption and use of the hormone, leading to an accumulation in the bloodstream. Fat and muscle cells need to bind the hormone insulin to cell receptors so that glucose can be absorbed. When cells fail to respond to insulin, glucose levels in the blood stream rise.[3].

In states of insulin resistance, there is an increase in the production of insulin by the β-cells in the pancreas that leads to high blood insulin (hyperinsulinemia) to compensate for the high blood glucose. At this stage, blood glucose levels are able to be maintained due to high insulin levels. If compensatory insulin secretion fails, glucose concentrations increase and T2DM occurs. The inability of the β-cells to produce sufficient insulin in a condition of hyperglycemia is what characterizes the transition from insulin resistance to T2DM [3], [4].

In normal metabolism, in muscle (Figure 1.2.1.1.- A), insulin signalling is activated when insulin binds to the insulin receptor. The insulin receptor autophosphorylates and recruits the insulin receptor substrates 1. These proteins activate the P13K/AKT pathway, leading to glucose uptake, through translocation of GLUT4-containing storage vesicles (GSVs) to the plasma membrane, and to glycogen synthesis [3].

In liver (Figure 1.2.1.1- B), the process is similar, but glucose transport is not insulin regulated, and therefore, insulin exerts less control over glycogen synthetic rates than in skeletal muscle. In normal hyperglycemia, liver glycogen phosphorylase (GP) is inactivated by glucose allostery while glycogen synthase (GS) is activated. Hyperglycemia also causes the translocation of glucokinase (phosphorylates the glucose to glucose-6-phosphate) from the nucleus to the cytoplasm, allowing glucose-G6P flux [3].

However, hepatic insulin signalling is required for normal glycogen synthesis. Insulin facilitates the translocation of glucokinase to the cytoplasm and exerts a potent and rapid positive transcriptional control over Gck gene. Therefore, increased glucokinase expression is critical for

hepatic insulin action as it increases G6P allostery at GS and controls hepatic glucose utilization and storage [3], [5]. Interestingly, some studies shown that patients with T2DM display a decreased glucokinase expression [5].

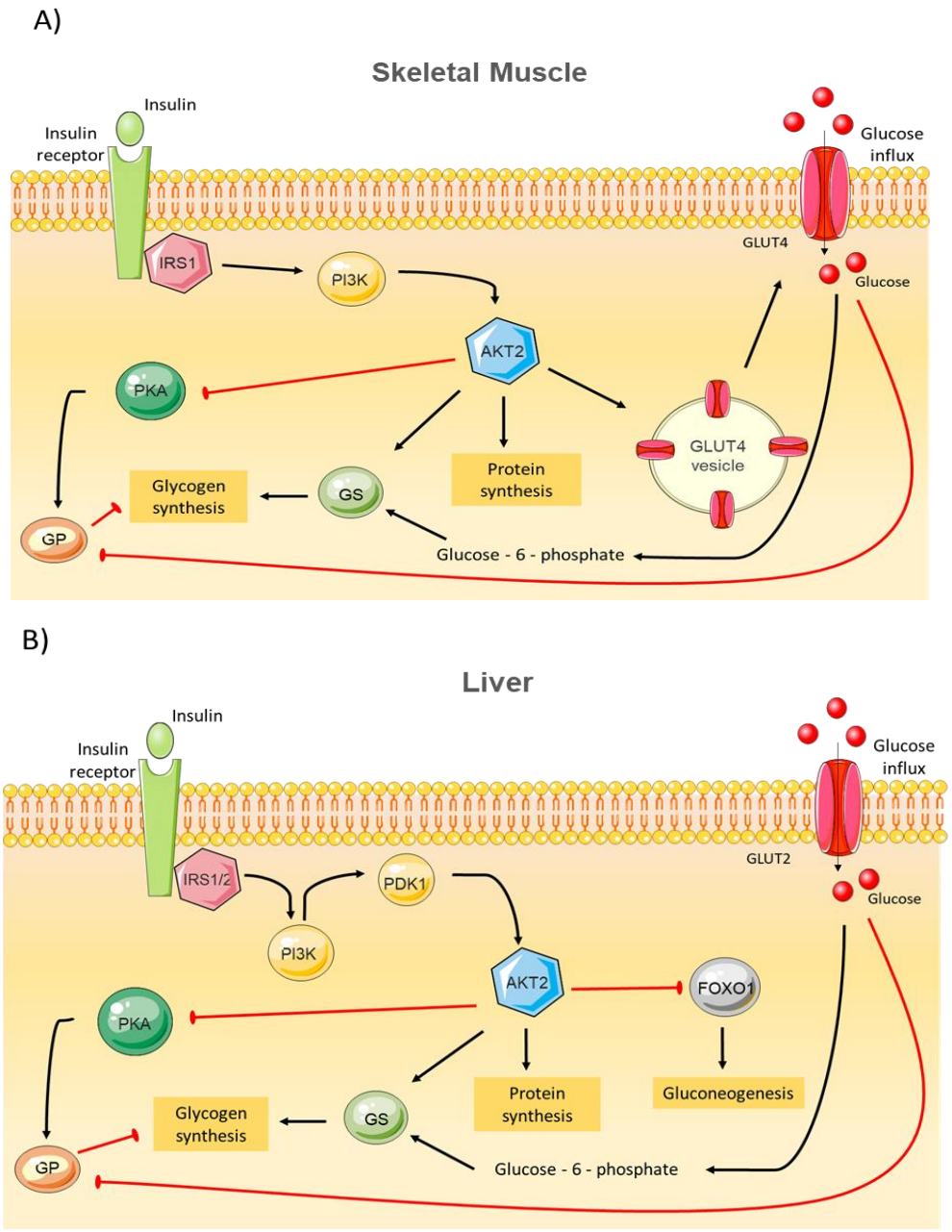


Figure 1.2.1.1 – A) Insulin signalling cascade in skeletal muscle. Insulin receptor activation leads to glucose uptake and glycogen storage. Insulin stimulation of glucose uptake occurs through translocation of GLUT4-containing storage vesicles (GSVs) to the plasma membrane. The resultant increase in intracellular glucose-6-phosphate production, together with a coordinated dephosphorylation of glycogen metabolic proteins, leads to the glycogen synthesis. B) Hepatic insulin signalling. The activation of insulin receptor leads to a fast activation of glycogen and protein synthetic machinery. Nomenclature: Black arrows represent activating events; red arrows represent inhibitory events. IRS, insulin receptor substrate; PI3K, phosphoinositide-3-kinase; PDK1, phosphoinositide-dependent protein kinase-1; AKT, serine/threonine kinase 1; PKA: Protein kinase A; GP, glycogen phosphorylase; GS, glycogen synthase; FOXO1, transcription factor forkhead box O1; GLUT2, glucose transporter type 2; GLUT4, Glucose transporter type 4.

1.2.2. Hepatic glucose production

The synthesis of glucose by the liver is made through two pathways: gluconeogenesis and glycogenolysis. (Figure 1.2.2.1)

Gluconeogenesis is the process of synthesizing glucose from other organic substrates. The pyruvic acid is the starting point but oxalic acid and dihydroxyacetone phosphate (DHAP) also provide entry points. Lactic acid, some amino acids from protein and glycerol from fat can be converted into glucose [6].

Glycogenolysis is the process by which glycogen, the primary carbohydrate stored in the muscle and liver cells, is broken down into glucose. The breakdown of glycogen polymer is catalysed by glycogen phosphorylase leading to the formation of glucose-1-phosphate. These is later converted in glucose-6-phosphate by phosphoglucomutase [6].

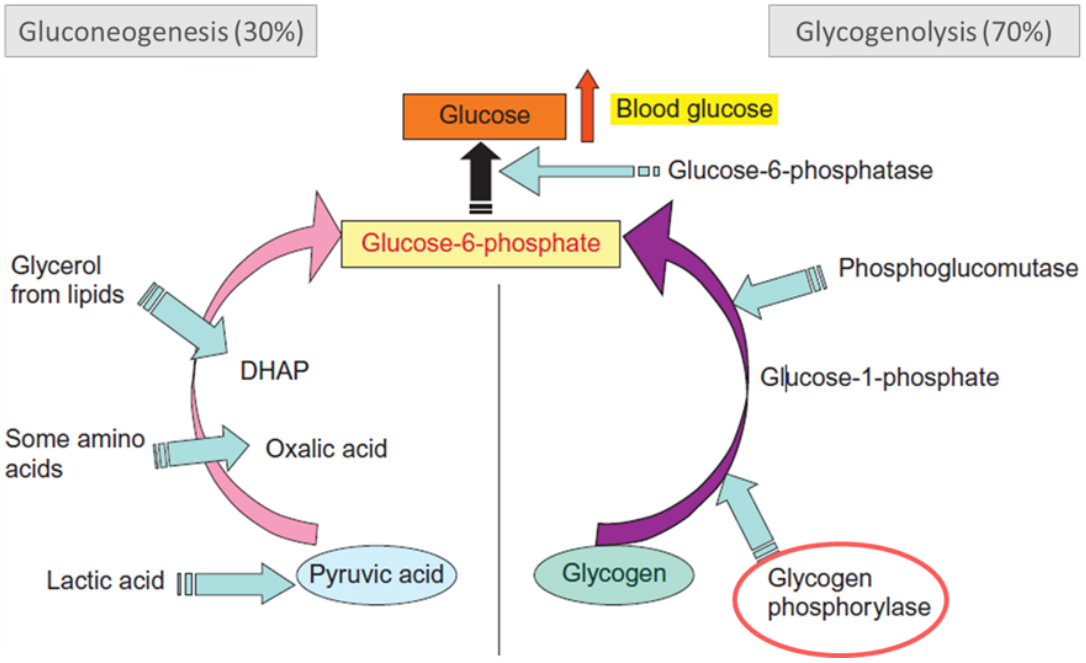


Figure 1.2.2.1 – Glucose production pathways by the liver. In diabetes mellitus type 2, glycogenolysis may contribute more than 70% of hepatic glucose production while gluconeogenesis may contribute 30% [6]. (Adapted from Naik *et al.*, 2013)

In patients with T2DM, higher hepatic glucose output is observed. Studies carried out by Krssak *et al*, 2004 [7] show that glycogen synthesis is decreased in mildly overweight patients and that this is accompanied by impaired inhibition of hepatic glucose production. Other studies have suggested that glycogenolysis is an important contributor to the abnormally high synthesis of glucose, contributing more than 70% of total hepatic glucose production [8].

1.2.3. Treatment approach

The treatment for T2DM focuses in maintaining blood glucose levels within normal parameters. Lifestyle changes, such as exercising and losing weight, can help reduce blood glucose levels and the risk of cardiovascular problems. However, this change in lifestyle is not always enough to reduce blood glucose and the use of medications is often necessary. The conventional medications in diabetes treatment focus on: increasing pancreatic insulin secretion (sulfonylureas), decreasing liver production of glucose (metformin), reducing glycids absorption (alpha-glucosidase inhibitors) and/or increasing peripheral glucose utilization (thiazolidinedione) [9], [10]. However, these have been shown to have several adverse side effects, such as hypoglycemia and/or gastrointestinal disorders, and are inappropriate for 30–40% of patients [11]. Other treatments such as gene therapy and induced β-cells regeneration have been studied but have not been widely introduced to manage diabetes. Due to the side effects and ineffectiveness of current therapies, other alternatives must be investigated [11]–[13].

The synthesis of glucose by the liver through glycogenolysis and gluconeogenesis pathways contributes greatly to the increase of glucose in blood. Therefore, the proteins involved in both enzymatic cascades are strategically important to therapeutic intervention and drug development [6], [11].

Therefore, the enzyme glycogen phosphorylase (GP) will be studied in this work since it is the main regulatory enzyme in the glycogenolysis pathway and its inhibition may help to reduce the glucose synthesis.

1.3. Glycogen Phosphorylase (GP)

1.3.1. Mechanism of action

Glycogen phosphorylase (GP; [EC 2.4.1.1]) is a key enzyme in the regulation of the glycogen breakdown. It belongs to the glycosyltransferase family and was the first phosphorylase enzyme to be discovered [14]. The first structure was obtained by X-ray crystallography by Barford *et al*, 1990 [15] (PDB entry – 8GPB), with a resolution of 2.20 Å, from the rabbit muscle [15].

Glycogen phosphorylase catalyzes the first step of glycogen degradation, a polymer of glucose subunits linked in (α1 → 4) with a high number of branches (α1 → 6). The enzyme binds to the glycogen macromolecule and catalyses the addition of phosphoric acid across the 1,4-anomeric linkage, leading to the release of one glucose-1-phosphate and a glycogen macromolecule shortened by one glucose molecule (Figure 1.3.1.1). This reaction is reversible but *in vivo* the enzyme acts only in the direction of glycogen phosphorolysis due to the concentration of phosphate in the cell being higher than the concentration of glucose-1-phosphate [16], [17].

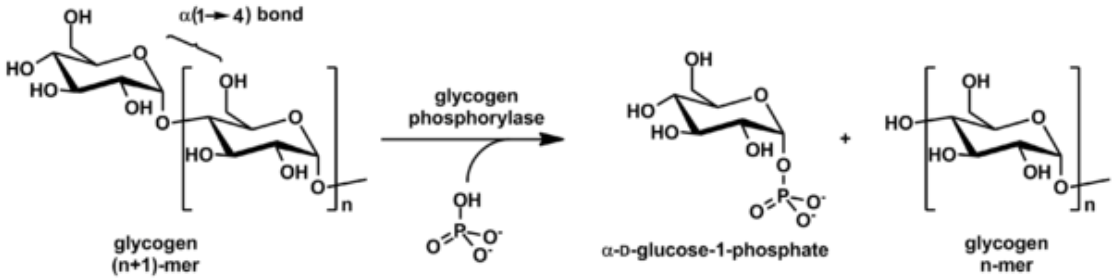


Figure 1.3.1.1 – Glycogen phosphorylase mechanism of action. Glycogen phosphorylase binds to the glycogen macromolecule and catalyses the addition of phosphoric acid across the 1,4-glycosidic bond, releasing glucose-1-phosphate and a glycogen macromolecule shortened by one sugar residue.

For GP to catalyse this reaction, it is necessary to have a cofactor, a pyridoxal phosphate (PLP), in each catalytic site that links with the Lys680 and covalently forms a Schiff base. After this, the phosphate group present in the PLP donates a proton to an inorganic phosphate molecule which in turn is deprotonated by oxygen, leading to the formation of the α-1,4 glycosidic linkage. The protonated oxygen is now a good leaving group and the glycogen chain is separated from the terminal glycogen in an S_N1 mechanism, forming a glucose molecule with a secondary carbocation. Then, the inorganic phosphate acts as a nucleophile and bonds with the carbocation leading to the formation of glucose 1-phosphate and shortened the glycogen chain by one glucose molecule [18]–[20].

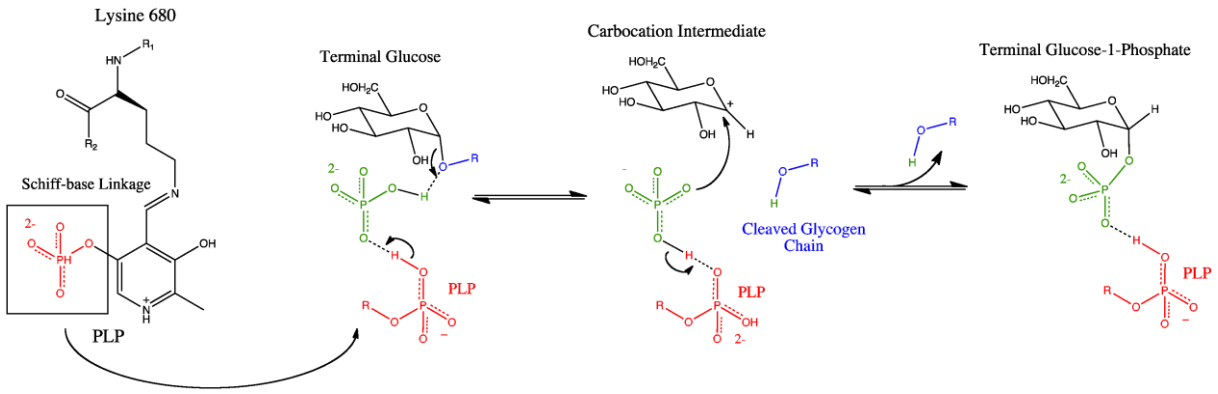


Figure 1.3.1.2 – Mechanism of PLP participation in the enzymatic reaction catalysed by glycogen phosphorylase [18]. (Adapted from Palm *et all*, 1990)

1.3.2. Regulatory sites and conformations

Humans express three different isoforms of the enzyme: brain, liver and skeletal muscle encoded by the respective genes PYGB, PYGL and PYGM. These genes are located on human chromosomes 20, 14 and 11, respectively. Although distinct, the three human isoforms of GP share approximately 80% amino acid identity and have a molecular weight of ~ 97 kDa (for the monomeric form). The brain and muscle enzymes serve the tissues in which they are found while the liver enzyme tends to meet the glycemic demands of the body as a whole [21]–[23].

The enzyme can exist inactive as a monomer or tetramer but its biologically active in the

homodimer form. Each subunit consists of two domains.

GP is regulated through two ways: reversible phosphorylation in serine 14 and by allosteric effectors.

Ser14 is close to the subunits interface and can be reversible phosphorylated. The phosphorylation leads to the conversion of phosphorylase b (GPb) to phosphorylase a (GPa) and changes the conformation of residues 10 to 22 into α helices. This alteration increases the GP activity in almost 25% [20], [24].

The enzyme can exist in a less active tense state (T-GP) or in an active and relaxed state (R-GP). Phosphorylation of Ser 14 by phosphorylase kinase (Phk) leads to the conversion of GPb form to GPa and simultaneously promotes the R state (R-GPa, high substrate affinity and high activity), while dephosphorylation leads to the opposite (T-GPb, low substrate affinity and low activity). The phosphorylase activity is also regulated by allosteric inhibitors (glucose, glucose-6-phosphate, caffeine and ATP) and allosteric activators (glycogen and AMP) (Figure 1.3.2.1). The liver GP is very responsive to glucose concentration, leading to the transition from the relax form to the tense form when the glucose concentrations are higher. The liver GP, contrary to the other isozymes, shows only a slight increase in activity in the presence of AMP (10%-20%) [11], [23]–[25].

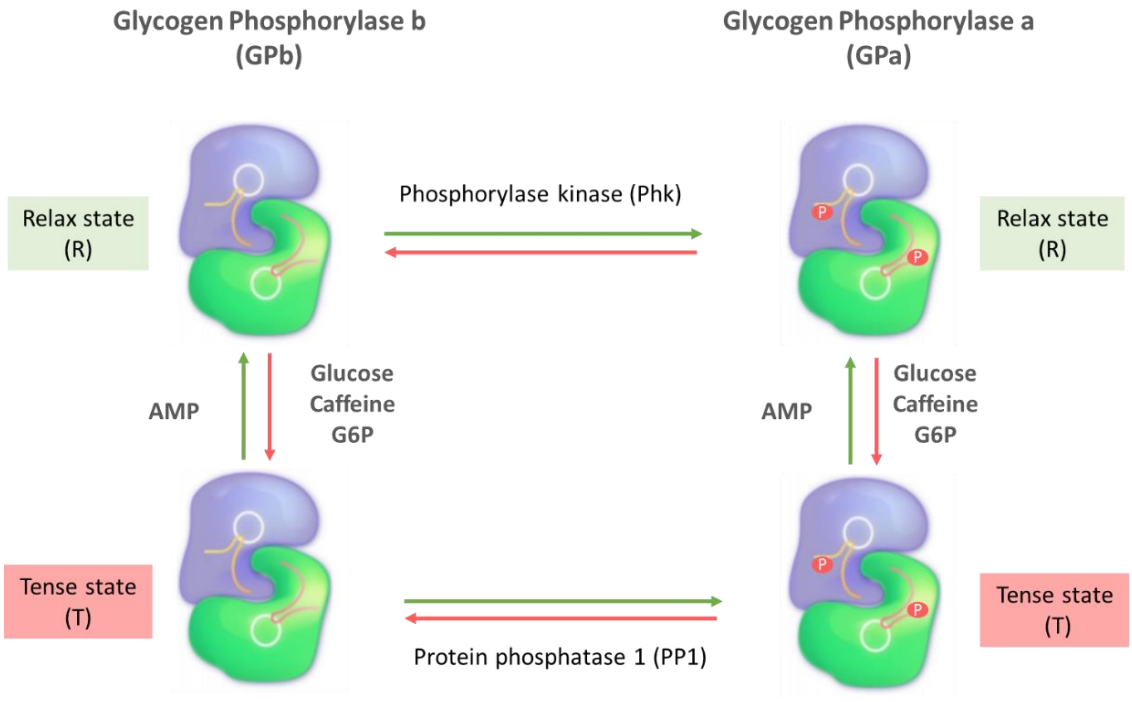


Figure 1.3.2.1 – Representation of the main mechanisms of glycogen phosphorylase regulation. Phosphorylase kinase promotes conversion of GPb to GPa by phosphorylation at Ser14. Allosteric ligands stabilize either a relaxed or a tense conformation. Glucose, glucose-6-P and ATP are allosteric inhibitors while AMP is an allosteric activator. Caffeine binds to the inhibitor site of GP [26]. (Adapted from Llavero *et al*, 2019)

1.3.3. Structure and binding sites

Six different GP ligand binding sites have been identified: the catalytic, allosteric, new allosteric, glycogen storage, inhibitor and quercetin-binding site [27].

The N-terminal contains several regulatory sites, including the serine-14, the allosteric binding site (binds G6P, fructose 1-phosphate, AMP and ATP) and the glycogen storage site (binds the glycogen molecule). The C-terminal contains the cofactor (PLP) binding site and together with residues form the N-terminal contains the catalytic site (Figure 1.3.3.1) [27].

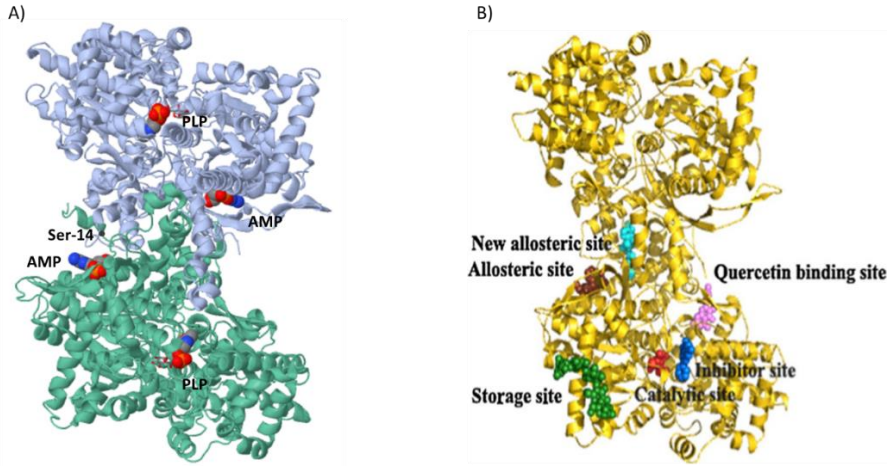


Figure 1.3.3.1 –3D structure of human liver GP and main mechanisms of regulation. A) Structure of hlGP dimer showing the positions of serine-14, AMP binding and pyridoxal phosphate (PLP) cofactor (PDB entry 1fa9). B) Highlighting of the binding sites of GP, including catalytic (red), allosteric (brown), new allosteric (cyan), inhibitor (blue), glycogen storage (green) and a very recently discovered new quercetin-binding site (pink) [27], [28].

The catalytic site includes the PLP and is buried at the centre of the subunit accessible to the bulk solvent through a 15 Å long channel. Therefore, access to the catalytic site becomes very difficult which makes the protein activity very susceptible to regulation by small allosteric effects. Glucose is the natural inhibitor of this site and it´s binding stabilizes the closed conformation of the 280s loop (residues 282-287), promoting the inactive T state conformation [20], [29], [30].

The allosteric site is composed by structural elements from the two subunits of the functional dimer and is located about 30 Å from the catalytic site. The allosteric site recognises a variety of phosphorylated compound such as AMP, ATP and G6P [20], [30].

The new allosteric site is inside a cavity created by the association of the two monomers located about 33 Å from the catalytic site, 37 Å from the inhibitor site and 15 Å from the allosteric site. The bonding of ligands in this site favours the tense state [31].

The glycogen binding site is where the enzyme binds covalently to glycogen macromolecule before initiating cleavage of terminal glucose molecules. The site is formed by residues 397 to 437 and is located 30 Å from the catalytic site [20].

The inhibitor (or nucleoside) site is a hydrophobic binding pocket located on the surface of the enzyme and 12 Å from the entrance to the catalytic site. The site binds purine analogs or fused-ring

molecules such as AMP, NADH, caffeine and adenosine when high concentrations are available. The heterocyclic rings of these compounds are able to bond and stabilize the inhibitor site and, therefore, block access to the catalytic site [20], [29].

Quercetin has recently been identified as a potent GP ligand of the so-called quercetin binding site, situated 15 Å from the active site. Quercetin stabilizes the T state through stabilization of the conformation involved in the structure of the binding site [27], [32].

1.3.4. Cascade signalling activation

Several hormones, such as epinephrine and glucagon, significantly affect glycogen metabolism by activate the breakdown of glycogen. Epinephrine, or adrenaline, is a catecholamine derived from tyrosine that is mainly secreted by the medulla of the adrenal glands. During muscle activity, epinephrine is released and stimulates glycogen breakdown in the liver and, especially, in the muscle. The liver shows a greater response to the glucagon hormone, which is secreted by the α cells of the pancreas when the blood-sugar level are low [33].

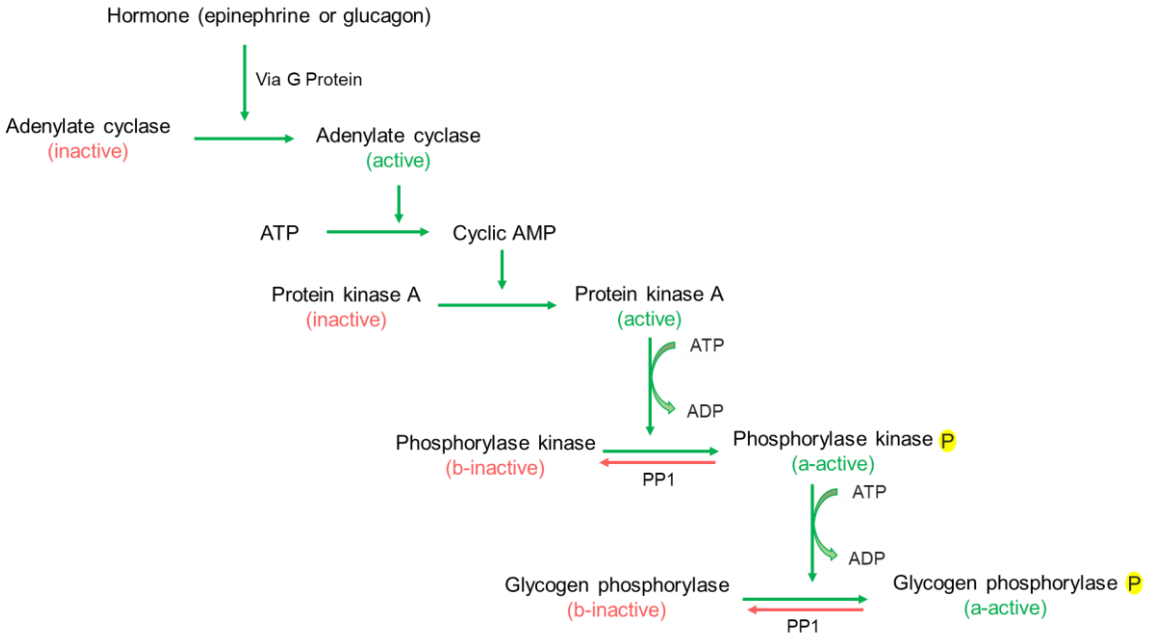


Figure 1.3.4.1 – Signal cascade by which human liver glycogen phosphorylase is regulate. Epinephrine and glucagon regulate glycogen phosphorylase using second messenger amplification systems linked to G proteins. Glucagon activates adenylate cyclase through a G protein leading to the increase of intracellular concentrations of cAMP. cAMP binds to and activates protein kinase A (PKA) which phosphorylates phosphorylase kinase, which in turn phosphorylates glycogen phosphorylase b at Ser14, converting it into the active glycogen phosphorylase a.

These hormones bind to specific receptors in the plasma membranes. Epinephrine binds to the β-adrenergic receptor in muscle, whereas glucagon binds to the glucagon receptor. The binding to the receptors leads to a change in their conformation that allows the binding and activation of the α subunit of the Gs protein. The via of G protein activates the adenylate cyclase, a transmembrane protein that catalyzes the formation of the secondary messenger cyclic AMP from ATP. This leads to an increase of

cyclic AMP in cytosol that activates protein kinase A through the binding of cyclic AMP to the regulatory subunits. These dissociate from the catalytic subunits, leaving them free and active. Protein kinase A phosphorylates the β subunit of phosphorylase kinase, which subsequently activates glycogen phosphorylase (Figure 1.3.4.1) [33].

1.3.5. Flavonoids – Natural inhibitors

As discussed above, the liver GP is the one responsible for the glycemic demands of the body as a whole, and it is already a validated target for the treatment of T2DM.

The synthesis of new molecules that mimic the structure of glucose has been very studied since G6P was recognized as an inhibitor of *hl*GP. However, some recent studies have focused preferential in natural compounds, in particular flavonoids [11], [34].

Flavonoids are family of polyphenolic plant compounds that naturally occurring in fruit, vegetables and beverages like wine and tea. Because of their presumed association with favourable health effects such as anti-inflammatory, anti-thrombogenic, antidiabetic, anticancer and neuroprotective, products rich in flavonoids become more and more popular. Jakobs *et al*, 2006 [34] showed that GP was strongly inhibited by some flavonoids, specially by compounds having a double bond ring C together with hydroxyl groups at positions 5 and 7 in ring A and vicinal hydroxyl groups in ring B. The flavonoids genistein, 8-prenylgenistein (8PG) and 8-prenylnaringenin (8PN) (Figure 1.3.5.1), show these characteristics. Genistein and 8PG belong to a class of flavonoids known as isoflavones while 8PN belongs to the flavanones class [34] .

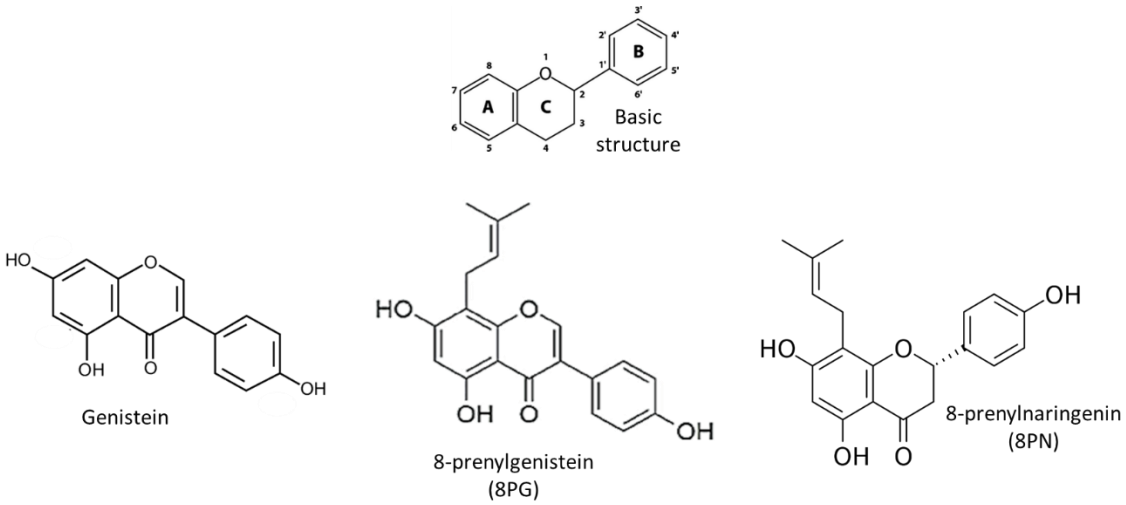


Figure 1.3.5.1 – Structures of flavonoids genistein, 8-prenylgenistein (8PG) and 8-prenylnaringenin (8PN).

Studies have shown that these compounds demonstrate a wide variety of pharmacological properties, including estrogenic, anti-inflammatory, anti-viral and anti-oxidative effects. Previous studies also demonstrate that the prenyl-C8 substituent on flavonoids improves their pharmacological profile and increases the bioactivities of its parent compound [35].

To gain deeper knowledge on potential effects of these compounds, the inhibitory capacity of *hl*GP by these three compounds will be studied.

2. Objectives

Previous computational studies, which include Molecular Dynamics and Free-energy perturbation calculations, carried out by our collaborators at Faculdade de Ciências da Universidade do Porto, have identified different natural compounds as potential inhibitors of human liver glycogen phosphorylase (*hl*GP). The present thesis aims to clarify the binding and inhibitory potential of *hl*GP with these different natural compounds that can assist in regulating blood glucose levels in diabetes. Therefore, this work had the following objectives:

- Optimize *E. coli* heterologous expression and purification protocols to improve expression yield and sample quality of human liver GP.

- Optimize purification and activation thought phosphorylation by phosphorylase kinase to obtain the protein in the most active state (*hl*GPa) as the purified protein is in the less active state (*hl*GPb).

- Perform Microscale Thermophoresis (MST) to quantify and characterize the biomolecular interactions of *hl*GPb with the inhibitors genistein, 8-prenylgenistein (8PG) and 8-prenylnaringenin (8PN) and validate the theoretical calculated data.

- Perform kinetic assays to determine the kinetic parameters (V0, Km and Vm) of *hl*GP with the inhibitors genistein, 8-prenylgenistein (8PG) and 8-prenylnaringenin (8PN).

- Characterize structurally, by X-ray crystallography, *hl*GP complexed with the selected compounds to clarify the binding and inhibition reaction mechanism of enzyme.

  All steps were performed at the UCIBIO Protein Crystallography Laboratory, FCT-UNL.

3. Materials and Methods

The purity and supplier of the reagents and solutions used throughout the experimental procedures described below are presented in Appendix 2.

3.1. Competent cells preparation

Frozen stocks of *Escherichia coli* (*E. coli*) BL21(DE3), Tuner (DE3) and Rosetta 2 (DE3) cells were used. The cells were grown in Luria Bertani (LB) medium (Appendix 2, Table 7.2.1) at a temperature of 37°C and 200 rpm. Once the optical density at 600 nm (OD_{600}) reached a value between 0.4 and 0.6 or, in other words, once the cell culture reached the mid-exponential phase, the cells were centrifuged at 2500 x g, at 8°C for 15 min. The cell pellet was resuspended in a cooled TBFI (60 mL) solution (Appendix 2, Table 7.2.2) and left to rest for about 90 min. Afterwards, the suspension of cells was centrifuged at 2500 x g for 10 min and the pellet was resuspended in a cooled solution of TBFII (8 mL) (Appendix 2, Table 7.2.3). The cells where then distributed into cooled Eppendorfs, each containing 200 µL of competent cells, and frozen at -80°C.

3.2. Recombinant protein expression

The gene coding for *hl*GP was previously cloned in plasmid pET-28a-c(+) vector (plasmid pHLGP28) that contains a kanamycin resistance marker (Appendix 1) and a hexa Histidine-tag in the N-terminal of the protein to facilitate the purification process through immobilized metal affinity chromatography (IMAC).

The recombinant protein presents 870 residues, a molecular weight of 99.601 kDa, a theoretical pI of 6.97 and a theoretical extinction coefficient of 118.150 M^{-1} cm^{-1} (data obtained from the ProtParam tool, ExPASy).

3.2.1. *HL*GP Recombinant Heterologous Overexpression Optimization

In order to obtain large amounts of protein, it is essential to determine which cells are the best for its production. Therefore, expression of the recombinant proteins requires a first transformation step using the previously competent *E. coli* cells from the different strains - Rosetta 2(DE3), BL21 (DE3) and Tuner (DE3). For each case, the cells were transformed with the plasmid pHLGP28 through the heat shock method [35] under sterile environment, where 1 µL of plasmid DNA (pDNA), was added to a previously cooled Eppendorf containing 100 µL of competent cells. The mixture was incubated on ice for 30 to 60 minutes followed by a thermal shock at 42°C for 50 seconds and placed on ice for 5 minutes. Subsequently, 1 mL of fresh LB medium was added and the cells were incubated at 37°C for 1h30. The cells were centrifuged (Centrifuge MiniSpin) for 1 minute at 5000 rpm, excess culture medium was removed, and the pellets resuspended. The resulting cell suspension was plated onto solid LB medium with kanamycin (NZYTech) and incubated overnight at a temperature of 37°C.

In the following day, a pre-inoculum was prepared, for each of the three strains, with 10 mL LB medium supplemented with 50 mg/L kanamycin, 100 mg/L pyridoxine (Sigma) and 600 mg/L $MnCl_2$ (Sigma), and incubated at 37°C, with stirring at 200 rpm (Orbital Shaker-Incubator ES-20, Grant.bio) overnight.

Subsequently, 2 mL of the pre-inoculum was transferred to 500 mL of fresh culture medium (in 2 L Erlenmeyers) supplemented with 50 mg/L kanamycin, 100 mg/L pyridoxine and 600 mg/L $MnCl_2$. The cultures were then incubated at 37°C and shaken at 200 rpm (Shaker Gallenkamp) until an optical density at 600 nm (OD_{600}) reached a value between 0.6 to 0.8.

After the period of incubation, different concentrations of IPTG (0.25 mM, 0.5 mM, 0.75 mM and 1 mM) were tested for the Tuner cells while for the BL21 and Rosetta 1 mM IPTG was added. The Tuner cells were incubated at 16°C and the other strains were incubated at 19°C. All the cells were shaken at 150 rpm (Shaker Gallenkamp) overnight.

The cells were then centrifuged for 15 minutes at 11325 x g (Rotor JA-10, Avanti J-26 XPI, Beckman Coulter) at a temperature of 8°C. The supernatant was discarded, the pellet was resuspended in biding buffer at 10 $mL.g^{-1}$ of wet weight cells and the cells were lysed by sonication (UP100H / Tip MS7, Hielscher Ultrasonics). Seven cycles of 2 minutes (cycle = 0.5, amplitude = 80%) were performed with a 2 minute rest interval. The lysate was clarified by centrifugation at 15557 x g (Centrifuge 5804 R, eppendorf) for 30 minutes at 8°C in order to separate the soluble fraction from the insoluble.

These fractions were analyzed by polyacrylamide gel electrophoresis under denaturing conditions (SDS-PAGE). A discontinuous Tris-Tricine polyacrylamide gel (4% stacking gel and 10% separating gel) was prepared and in each well was added 15 µL of sample with loading buffer previously heated at 100˚C during 5 minutes. In one well, 3 µL of the molecular weight marker (NZY Color Protein Marker II, NZYTech) was added and a current of 200 V and 35 mA was applied, for each gel, until the different bands of the molecular weight marker could be seen.

3.3. *HL*GP purification optimization

After determining the best expression conditions, more litres of cells were expressed as described above. The cells were then resuspended in binding buffer with DNase (PanReac), lysozyme (300 µg/mL) (Sigma) and protease inhibitors (Roche). Subsequently, the cells were lysed by sonication as previously and the supernatant was centrifuged at 15557 x g, at 8°C for 30 min. The insoluble fraction was discarded and the soluble fraction was filtered with a 0.45 µm filter (GE Healthcare). The sample went then to the first purification step.

3.3.1. Affinity chromatography

An immobilized metal affinity chromatography (IMAC) was made due to this technique being quite efficient in the separation of proteins that have tails of histidine residues (*hl*GP have a tail with 6 residues). Since the purification protocol for histidine-tailed GP has not yet been described, it was necessary to proceed to its optimization. Therefore, two different metal ions were tested: nickel and cobalt.

Chromatography with Ni^{2+}

The soluble fraction containing the protein was loaded into an HiTrap™ nickel column (GE-Healthcare) pre-equilibrated with binding buffer, using an AKTA START HPLC system (GE-Healthcare) and the UNICORN™ start 1.0 program to monitor the purification process. Initially, the system and the column were washed with binding buffer (50 mM HEPES pH 7.5, 1 M NaCl, 20 mM imidazol) and then the cell extract was loaded onto the column at a flow rate of 2 mL/min.

The column was washed with binding buffer at a flow rate of 2 mL/min, dragging all proteins with low affinity to the nickel ions immobilized on the column. The ionic strength was then increased by adding 10% of the elution buffer solution (50 mM HEPES pH 7.5, 1 M NaCl, 500 mM imidazol), eluting low affinity proteins that had remained on the column. Finally, the protein of interest was eluted by performing an elution solution gradient of up to 100%. The increase in the imidazol concentration allowed to elute the protein of interest, since it increases its competition with the protein for the metal ions immobilized in the column.

The collected fractions were subsequently analyzed by polyacrylamide gel electrophoresis under denaturing conditions (SDS-PAGE), as described above.

The collected fractions with the largest amount of protein were added to a concentrator (Centricon Vivaspin® Turbo 15, 30 kDa, Sartorius) and centrifuged (Centrifuge 5804 R, eppendorf) at 3000 x g, at a temperature of 8˚C in 10 minute cycles. After a volume of approximately 200 µL was obtained, the sample was stored at 4˚C.

Chromatography with Co^{2+}

With the cobalt ions, the process was similar with the exception of the following aspects:

- Type of column: it was used a HiTrap™ TALON® crude column (GE-Healthcare) with cobalt and a volume of 5 mL.
- Binding buffer: 50 mM sodium phosphate pH 7.5, 500 mM NaCl, 10% glycerol, 2 mM $MgCl_2$ and 5 mM b-mercaptoetanol.
- Elution buffer: 50 mM sodium phosphate pH 7.5, 500 mM NaCl, 10% glycerol, 2 mM $MgCl_2$, 5 mM b-mercaptoetanol and 500 mM imidazol.
- No gradient was made: instead two levels of 2% and 60% of elution buffer were made.

3.3.2. Anion exchange chromatography

An anion exchange chromatography was also tested. This chromatography allows to separate molecules based on their respective charged groups. The stationary phase (anionic resin) consists of a matrix where positively charged functional chemical groups are attached, which allows negatively charged molecules to be separated. The stronger the negative charge on the sample, the stronger it will be attracted to the positive charge on the stationary phase, and thus the longer it will take to elute. *Hl*GP has an isoelectric point of 6.9 and a buffer with pH 8 was used so that the total protein charge was negative allowing its binding to the column.

The process was similar to that of IMAC with the exception of the following aspects:

- Type of column: it was used a HiTrap™ Q HP anion exchange column (GE-Healthcare) with a volume of 5 mL.
- Binding buffer: 50 mM Tris-HCl pH 8, 2 mM $MgCl_2$, 10% glycerol and 1 mM DTT.
- Elution buffer: 50 mM Tris-HCl pH 8, 2 mM $MgCl_2$, 10% glycerol, 1 mM DTT and 1 M NaCl
- Imidazol was not used: the elution gradient was made with 0 to 100% of NaCl.

3.3.3. Protein phosphorylation

After the first purification step and the analysis of the fractions by SDS-PAGE, the samples with more protein were dialyzed overnight for the buffer 50 mM HEPES pH 7.5, 20% glycerol, 1 mM EDTA, 2 mM DTT, 2 mM $MgCl_2$ and 100 mM NaCl. The dialyzed fraction was divided into two samples. One sample was converted to *hl*GPa by adding 2 mM ATP, 1 mM AMP and phosphorylase kinase (5 µg per 1 mg of GP) and was incubated during 1 hour at room temperature. The samples were then purified by SEC.

3.3.4. Size exclusion chromatography

A size exclusion chromatography step was performed using a 24 mL Superdex™ 200 Increase 10/300 GL (GE Healthcare) connected to a shimadzu system and the program Lab Solutions was used to monitor the purification process. The column and the system were equilibrated with 50 mM HEPES pH 7.5, 10% glycerol, 1 mM EDTA, 1 mM DTT and 2 mM $MgCl_2$ before each injection of 1.5 mg/mL of protein. The flow rate was set to 0.7 mL/min, and 500 µL fractions were collected and analysed by polyacrylamide gel electrophoresis under denaturing conditions as previously described.

The fractions with the highest amount of protein were added to a concentrator (Centricon Vivaspin® Turbo 15 de 30 kDa, Sartorius) and centrifuged (Centrifuge 5804 R, eppendorf) at 3000 x g, at a temperature of 8°C until a volume between 150 and 300 µL was reached. After this, the sample was stored at 4°C.

3.4. Thermal Shift Assay (TSA)

Thermal shift assay (TSA) is a technique that allows to study the protein stability by determination of the protein melting temperature. The higher the melting temperature, the more stable the protein will be and this way, using different buffers and adding different additives, it is possible to determine the best conditions to stabilize the protein in solution.

For the *hl*GP, the best conditions were determined using a screen of buffers (Appendix 3, Table 7.3.1) and an additive screen (RUBIC, Molecular Dimensions). Different concentrations of DMSO, PEG 400, urea, imidazol and glucose alone or with AMP and $MgCl_2$ were also tested (Appendix 3, Table 7.3.2).

A 96-well PCR plate (Applied Biosystems) was used for the assay and was added to each reservoir:

- 2 µL of protein at a final concentration of 2 µM
- 8 µl of diluted TSA dye (ROX™): 2.4 µl of dye + 797.6 µL of H_2O/ protein buffer
- 10 µl of buffer screen/ additive solution

Negative controls were performed, replacing fluorophore or protein with distilled water.

The plate was sealed with a film (Excel Scientific) and centrifuged (Centrifuge 5804 R, eppendorf) for 1 minute at 1000 x g to eliminate air bubbles. The StepOnePlus™ Real-Time PCR System (Applied Biosystems™, ThermoFisher Scientific) was used to perform the TSA, with 2 minutes cycles of 1% increments between 25 and 95°C.

3.5. MicroScale Thermophoresis (MST)

MicroScale Thermophoresis (MST) is a biophysical technique that quantify biomolecular interactions. It is based on the movement of molecules during a temperature gradient (thermophoresis) that strongly depends on molecular properties such as size, charge and conformation. The temperature gradient is achieved using an infrared laser and the direct movement of particles results in a change of concentration that is detected by fluorescence. This fluorescence can be intrinsic to the molecule or due a covalently attached probe. By combining the precision of fluorescence detection with the sensitivity of thermophoresis, MST provides a flexible and strong way to study molecular interactions [36].

3.5.1. MST measurements

This technique was used to test the interactions of the *hl*GPb with the compounds genistein, 8-prenylnaringenin, 8-prenylgenistein, glucose and caffeine. As the *hl*GPb does not possess intrinsic fluorescence signal, it was necessary to label the protein with the fluorescent dye RED-tris-NTA (NanoTemper Technologies) that binds to the His-tag present in the N-terminus. Both the dye and the protein were diluted in 50 mM Tris-HCl pH 7.5 buffer, 1 mM $MgCl_2$ and 1 mM DTT, to the concentration of 100 nM and 200 nM, respectively, mixed together and incubated for 30 minutes at room temperature

in the dark, followed by centrifugation at 10000 x g for 10 min. After protein labelling, a pre-test was performed to check if the labelling had been successful.

MST was measured using a Monolith NT.115 instrument (NanoTemper Technologies) at 25°C under medium MST power, using standard capillaries that hold a reaction volume of around 10 µL. The excitation power using the Nano-RED filter set was 80% LED power. All dilutions were performed in 50 mM Tris-HCl pH 7.5 buffer, 1 mM $MgCl_2$ and 1 mM DTT and 0.05% Tween.

The compounds were dissolved in the MST buffer supplemented with 5% DMSO to increase their solubility (genistein, 8PN and 8PG), and a series of 16 dilutions (1:1) were prepared with the highest initial concentration of 0.25 mM for genistein (Sigma, G6649), 0.5 mM for 8PN (Cayman, 17462), 0.5 mM for 8PG (Musechem, M135845), 20 mM for caffeine (Fluka, 27600) and 5 mM for glucose (Sigma, G8270). The concentration of DMSO was kept constant at 5% for all the samples analyzed to standardize any possible interference. For the measurements, each ligand dilution was mixed with one volume of labeled *hl*GPb, which led to a final *hl*GPb concentration of 50 nM diluting to half the concentrations of each ligand and to a DMSO final concentration of 2.5%. For each case, data of three independently measurements were collected and analyzed using the MO. Affinity Analysis software version 2.3 (NanoTemper Technologies) using the signal from an MST-on time of 1.5 s and 2.5 s. As a negative control, the compound 8PN was titrated against the His6 control peptide from the Monolith-His-Tag-Labeling-Kit RED-tris-NTA.

3.6. Activity assays – malachite green phosphate assay

Malachite green assay is a simple and sensitive colorimetric method based on the formation of a complex between malachite green, ammonium molybdate and free orthophosphate (inorganic phosphate, P_i) under acidic conditions. Orthophosphate forms a complex with ammonium molybdate in a solution of sulphuric acid and the phosphomolybdate complex is measured at 650 nm in order to quantify phosphorylation and phosphate release from the reaction [37], [38].

The protocol used to measure *hl*GPa and *hl*GPb activities were adapted from Lanzetta *et al*, 1979 [38]. Briefly *hl*GP (30 ng) activity was measured in the direction of glycogen synthesis by the release of phosphate from glucose-1-phosphate at 22°C in 36.4 µL of a buffer containing 50 mM HEPES pH 7.5, 100 mM KCl, 2.5 mM EDTA, 2.5 mM $MgCl_2$, 0.5 mM AMP, 1 mg/mL glycogen and 0.5 mM glucose-1-phosphate. Several incubation times of 20, 40, 60, 100 and 120 seconds were performed and the reaction was stopped by the addiction of 145.4 µL of colour reagent (Appendix 4) followed immediately by sample vortex again. After 1 minute, it was added 18.2 µL of 34% (w/v) sodium citrate and the sample was vortex. After 5 minutes, phosphate was measured at 650 nm using the UV-Vis SpectraMax 190 Microplate Reader.

3.7. X-ray crystallography

X-ray crystallography is an essential technique for structural analysis, giving three-dimensional insight and information into the structural functions and mechanisms of important biological molecules [39]. Structural crystallography relies on the interactions between the electrons present in the molecule with the X-ray radiation. X-rays are a high energy electromagnetic radiation with wavelengths ranging between 0.1 and 100 Å, corresponding to the same range of the interatomic distances in molecules (1.0 Å).

To determine the three-dimensional structure of a molecule by X-ray crystallography it is necessary to obtain well-ordered single crystals. By adopting this crystalline structure, the molecules take one or few identical orientations and form an orderly three-dimensional network joined by hydrogen bridges [40], [41].

The most widely used crystallization method is vapor diffusion. This method allows to establish a balance between a drop of protein solution and precipitating solution, and a reservoir that contains a higher concentration of precipitating solution. Due to these differences in concentrations and because it is a closed system, vapor diffusion occurs with the transfer of water from the protein solution to the reservoir. After equilibrium is reached, the transfer of water from the drop to the well ceases and the protein solution is maintained at a constant precipitant concentration. This method includes two techniques, the hanging drop and the sitting drop. In the hanging drop technique, the drop is placed on a cover slip that rests on the reservoir while in the sitting, the drop is placed on a micro-bridge over the reservoir [41]–[43].

After obtaining well-ordered crystals, the diffraction experiment and data collection are carried out and the crystal is placed between an X-ray source and a detector. When the X-ray beam reaches the crystal, the radiation interacts with its atoms electrons, leading to the formation of a diffraction pattern according to the internal organization of the crystal, which is then registered in the detector in the form of reflections. However, the crystals can be damaged by these high-intensity beams. In order to minimize the effects caused by radiation damage, the crystals are frozen in the presence of cryoprotective agents, such as glycerol and paratone, before the diffraction experiment. Cryoprotective agents minimize the dispersion of X-rays by water molecules, as they prevent the formation of ice crystals on the surface and inside the protein crystal [40], [41].

After the data processing, it is necessary to determine the phases of all reflections in order to determine the three-dimensional structure of a protein. To do this, it is possible to use direct methods (in the case of very high-resolution diffraction data) or, most commonly, indirect methods like molecular replacement (MR), multiple isomorphous replacement (MIR) or single and multiple wavelength anomalous dispersion (SAD and MAD) [41].

The model must then be refined to improve the phases and the electronic densities in order to

obtain a structure that matches the experimental data and adopts a more favourable thermodynamic conformation. This is an iterative process that stops when no further improvement is achieved [41].

Finally, the model of the molecule structure can be deposited in the Protein Data Bank (PDB) database.

3.7.1. Protein crystallization assays

The crystallization assays were performed with the *hl*GPb alone or in the presence of the ligands.

The protein was in a buffer solution containing 50 mM Tris-HCl pH 7.5, 10% glycerol, 1 mM EDTA and 1 mM DTT in all assays with the exception of 8-prenylnaringenin, in which the protein was in a buffer solution containing 50 mM HEPES pH 7.5, 10% glycerol, 1 mM EDTA, 1 mM DTT and 2 mM $MgCl_2$.

For crystallization of *hl*GPb alone, it was used a drop ratio of 1:1, a reservoir volume of 50 µL, a concentration of 13.4 mg/mL and the commercial screens JBS 5,6,7,8 (from Jena Biosciences) and JCSG+ (from Molecular Dynamics) were tested.

In order to obtain the crystal structures of protein-ligand complexes, co-crystallization experiments with the genistein and 8PN were attempted. For the assays in the presence of 1 mM of genistein, it was used a drop ratio of 1:2 and a reservoir volume of 30 µL. Two concentrations of protein were used, 6.3 mg/mL and 13.7 mg/mL, and the commercial screens JBS 1,2,3,4 (from Jena Biosciences), JBS 5,6,7,8 (from Jena Biosciences) and JCSG+ (from Molecular Dynamics) were tested. For crystallizations in the presence of 1 mM of 8-prenylnaringenin, it was used a drop ratio of 1:1 and a reservoir volume of 30 µl. It was used a protein concentration of 5.02 mg/mL and it was tested the commercial screen Structure (from Molecular Dynamics).

All the assays were carried out using the vapor diffusion method and the sitting drop technique, at 20°C, and using the Oryx 8 crystallization robot (Douglas Instruments).

4. Results and Discussion

4.1. Recombinant protein expression optimization

To obtain stable and large amounts of protein for the future assays, it was necessary to determine which strain cells are best to over-express the protein. Thus, three bacterial strains were tested: (i) BL21, that are chemically competent *Escherichia coli* cells with a T7 RNA polymerase gene controlled by the *lac* promoter and, therefore, IPTG can be used to induce expression; (ii) Rosetta, that are BL21 derivatives designed to enhance the expression of eukaryotic proteins that contain codons rarely used in *E. coli*; (iii) Tuner, that also derivate from BL21, but have a mutation in the lac permease (lacY) that eliminates the active transport of lactose, which allows uniform entry of IPTG into all cells in the population and regulation of the expression by adjusting the concentration of IPTG [44].

After transforming the plasmid in study and expressing the three bacterial strains, the cells were centrifuged and lysed to release the cellular content since the protein of interest was not excreted by the host organism. In this work, cell lysis was done through sonication, which consists of applying ultrasounds waves that disrupt the bacterial cell membranes and release cellular contents. The cells were centrifuged in order to separate the insoluble cellular remains from the soluble ones.

To determine under what conditions the protein was better expressed and whether it was present in the soluble or insoluble extract, an SDS-PAGE gel was made (Figure 4.1.1).

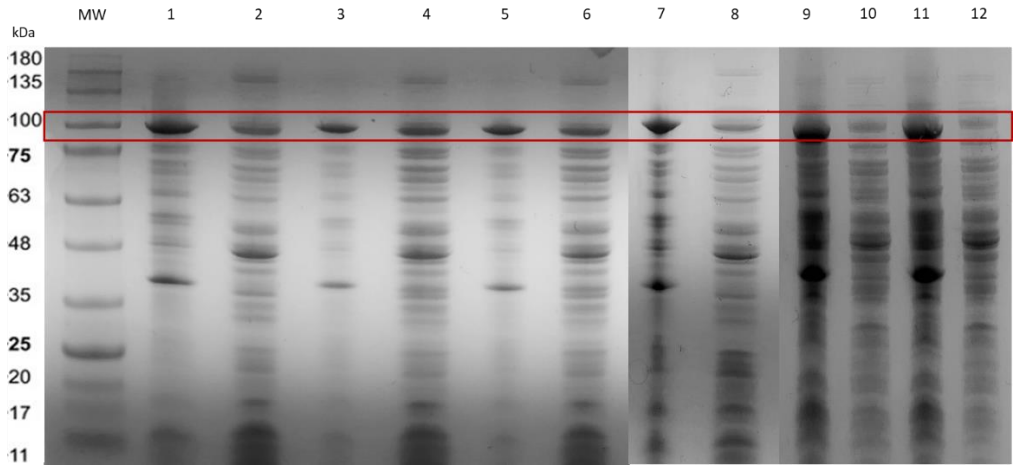


Figure 4.1.1 – SDS-PAGE of the insoluble and soluble extracts of the expression in BL21, Rosetta and Tuner. For the Tuner cells different concentration of IPTG were tested (0.25 mM, 0.5 mM, 0.75 mM and 1 mM) and the cells were induced at 16°C. The BL21 and Rosetta were induced at 19°C with 1 mM of IPTG. MW – NZY Colour Protein Marker II. 1-2 – Tuner cells with 0.25 mM of IPTG, insoluble and soluble extract respectively. 3-4 – Tuner cells with 0.5 mM of IPTG, insoluble and soluble extract. 5-6 – Tuner cells with 0.75 mM of IPTG, insoluble and soluble extract. 7-8 – Tuner cells with 1 mM of IPTG, insoluble and soluble extract. 9-10 – BL21 cells with 1 mM of IPTG, insoluble and soluble extract. 11-12 – Rosetta cells with 1 mM of IPTG, insoluble and soluble extract.

Through the gel it is possible to observe a band with approximately 100 kDa corresponding to the protein in all samples, indicating that the expression of the protein has been achieved. Comparing the solubility of the protein expressed by the different strains, it can be seen that, for the BL21 and

Rosetta, a much wider band is present in the insoluble extract when compared to the soluble, indicating a low solubility of the protein produced by these two strains. For Tuner cells, it is possible to verify that the expressed protein has a greater solubility and that the best concentration of IPTG is 0.5 mM. Thus, it is concluded that the protein is best expressed using the Tuner cells with a concentration of 0.5 mM IPTG and at an induction temperature of 16°C.

4.2. *HL*GP purification optimization

After expression, the cells were resuspended in ligation buffer and were lysed by sonication as described in section 3.2.1.

To help degrade the cell membranes and stabilize the released proteins, some enzymes were added: (i) lysozyme, to help in the degradation of the cell membrane of bacteria as it digests the peptidoglycan present in the membrane structure; (ii) Dnase, to break the DNA and allowing to reduce the viscosity of the solution and facilitating the purification process; (iii) protease inhibitors to inhibit enzymes responsible for protein breakdown.

To facilitate the purification process, the sample was centrifuged to separate the insoluble cellular remains from the soluble ones. The insoluble fraction was discarded and the following purification steps were carried out with the soluble fraction.

The first purification step consisted of an ion metal affinity chromatography (IMAC) since the *hl*GP was overexpressed with a six histidine tag (His-tag) and it has a great affinity for metallic ions, such as Ni^{2+} and Co^{2+}. Therefore, it was possible to separate the protein of interest from the others since it was attached to the stationary phase and the rest were eluted in the flow through.

To elute the proteins coupled to the column, the ionic strength was increased using a buffer solution with imidazol that has affinity for the column. Thus, there is a competition between the coupled protein and the imidazol which allowed the protein to elute. The obtained chromatogram is shown below (Figure 4.2.1).

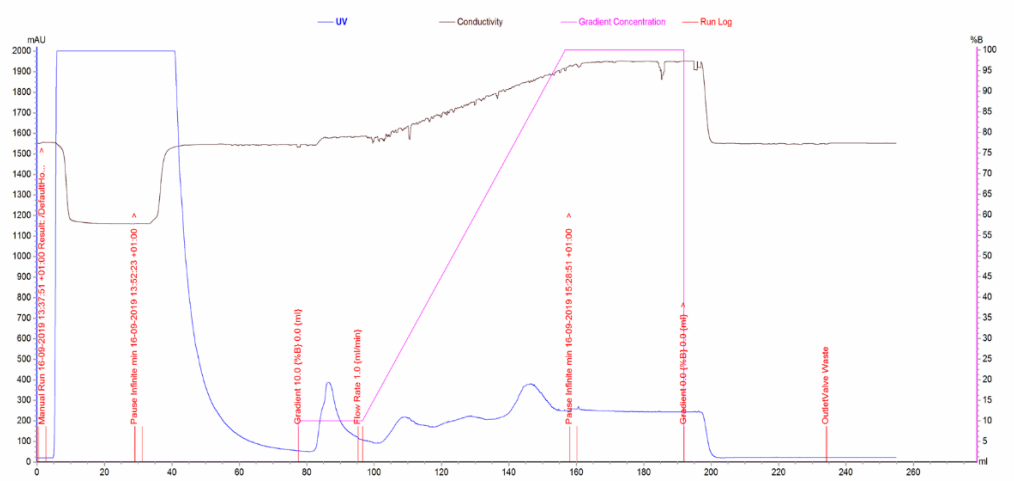


Figure 4.2.1 – Chromatogram obtained by IMAC for *hl*GP. The absorbance (mAU) at 280 nm and the percentage of elution buffer (% B) as a function of the elution volume (mL).

By analysing the chromatogram, it is possible to observe a peak between the 5-50 mL with a very strong intensity that exceeds the detection limit of the ultraviolet/visible detector and corresponds to the elution of proteins present in the soluble extract that had no affinity for the metal ion immobilized in the column (flow through - FT). Another peak is observed between the 80-100 mL that corresponds to the washing step with 10% of elution buffer (50 mM HEPES pH 7.5, 1 M NaCl, 500 mM imidazol) and the elution of proteins with low affinity for the column due to the increase in ionic strength. To collect the protein of interest, a gradient was made with 100% B, increasing the ionic strength. The increase in the concentration of imidazol in the buffer solution allowed the elution of the protein since it increased its competition with the metal ions immobilized in the column. It is possible to verify the appearance of different peaks during the elution gradient.

To confirm the purity of the samples collected, an SDS-PAGE gel was made (Figure 4.2.2).

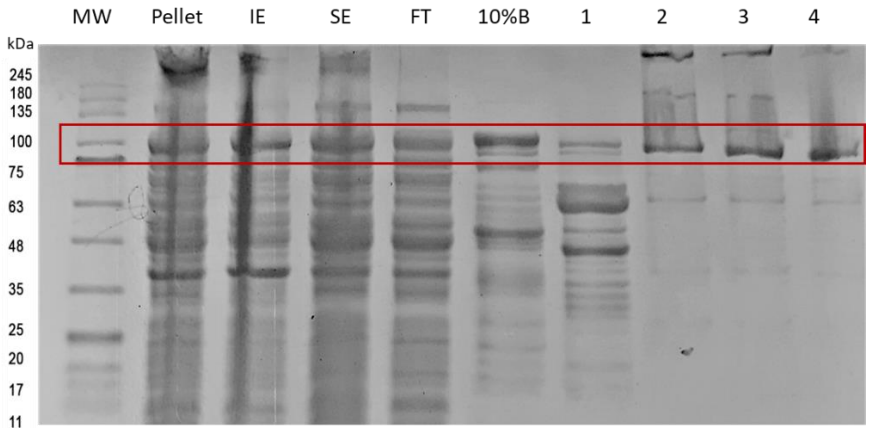


Figure 4.2.2 – SDS-PAGE of the fractions collected by IMAC. MW – NZY Colour Protein Marker II; Pellet – Sonicated extract; IE – Insoluble fraction; SE – Soluble fraction; FT – Flow through; 10%B – Eluted fraction with 50 mM of imidazol; 1 to 4 – Fractions collected from the gradient.

Through the gel it is possible to observe the broad band corresponding to the protein in the flow through and with 10% B, while in the fractions corresponding to the gradient (1-4) they present a thinner

band corresponding to the protein. This indicates that the protein is not fully coupling to the stationary phase of the column.

Trying to optimize the *hl*GP purification, the protein was expressed again and an anion exchange chromatography was performed. In this type of chromatography, the stationary phase consists of a matrix where positively charged functional chemical groups are attached, which allows negatively charged molecules to be separated. Since *hl*GP has an isoelectric point of 6.9, a buffer with pH 8 was used so that the total protein charge was negative. Proteins that have a positive charge are eluted in the flow through and proteins with a negative charge are coupled to the resin. To elute the protein of interest, it was necessary to gradually increase the ionic strength of the elution buffer so that the ions competed with the protein for the resin binding sites. In this case, a buffer with NaCl was used and the chromatogram is show below (Figure 4.2.3).

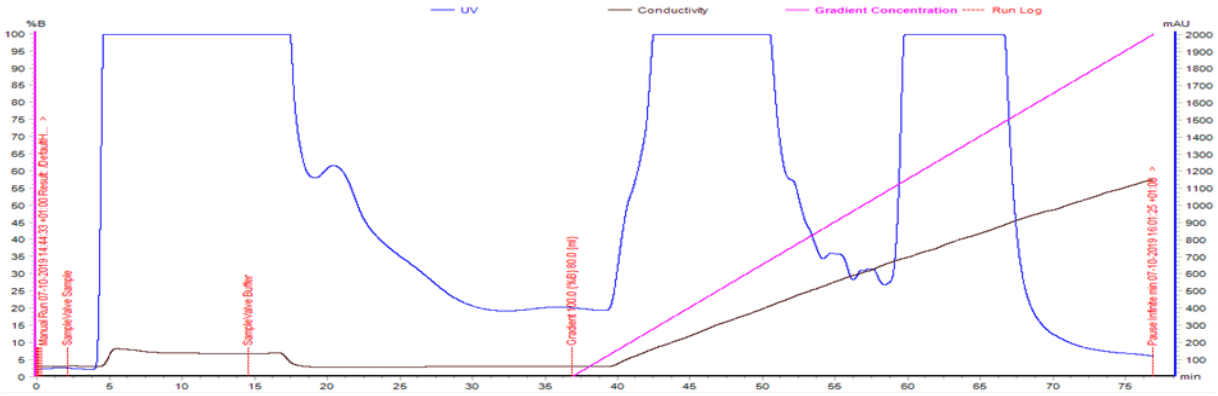


Figure 4.2.3 – Chromatogram obtained by anion change chromatography for *hl*GP. The absorbance (mAU) at 280 nm and the percentage of elution buffer (% B) as a function of the elution volume (mL).

Through the chromatogram it is possible to observe three peaks with high intensity and that exceed the detection limit of the detector. The first corresponds to the elution of proteins in the flow through that had a positive charge and, as such, are not coupled to the stationary phase. The second and third peaks correspond to the elution of negatively charged proteins when the ionic strength is increased with NaCl. In the second peak, proteins with a lower negative charge are eluted and, since the interaction they establish with the resin is weaker, the increase in NaCl causes them to be quickly eluted. The third peak corresponds to the proteins with the greatest negative charge and, as such, a greater ionic strength is required to elute these proteins.

Purity of the samples collected were checked by SDS-PAGE gel (Figure 4.2.4).

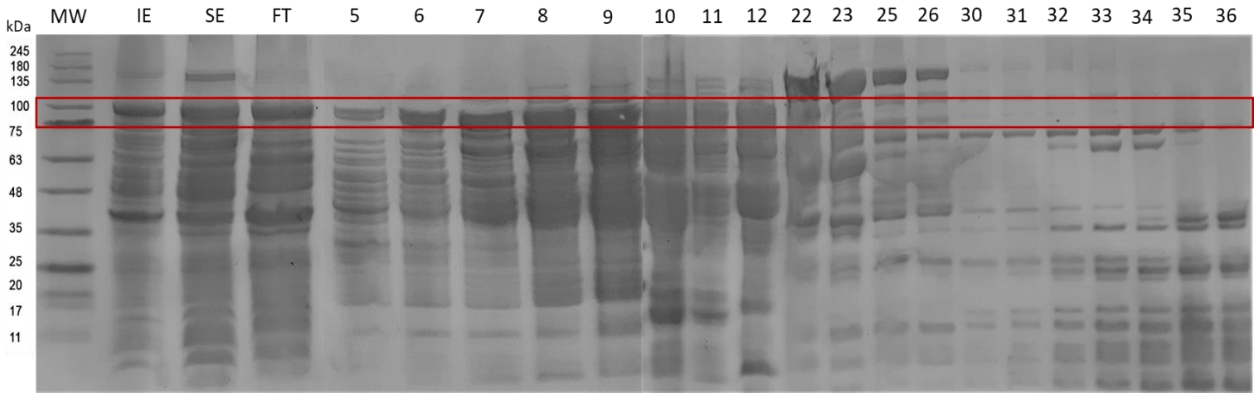


Figure 4.2.4 – SDS-PAGE of the fractions collected by anion change chromatography. MW – NZY Colour Protein Marker II; IE – Insoluble fraction; SE – Soluble fraction; FT – Flow through; 5 to 12 – Fractions from the second peak; 22 to 36 – Fractions from the third peak.

Through the gel it is possible to see that some of the protein was eluted during the flow through. It is also possible to verify that the protein of interest was eluted mainly in the second peak (fractions 5-12), although it is highly contaminated with other proteins.

Due to this, the fractions 5 to 12 were dialyzed overnight, for the buffer containing 50 mM sodium phosphate pH 7.5, 2 mM $MgCl_2$, 10% glycerol, 1 mM β-mercaptoetanol and 500 mM NaCl, in order to carry out an IMAC with nickel ions as performed previously, as a second purification step. However, the results were quite similar since the protein was eluted in the flow through or with other proteins.

Due to the necessity of further optimizing the *hl*GP purification, it was decided to express the protein again and perform a new IMAC, using a TALON crude column with coupled cobalt ions. These ions establish a more specific interaction with histidine tags, resulting in less nonspecific interaction and reduced affinity for host protein [45]. The chromatogram obtained is shown below (Figure 4.2.5).

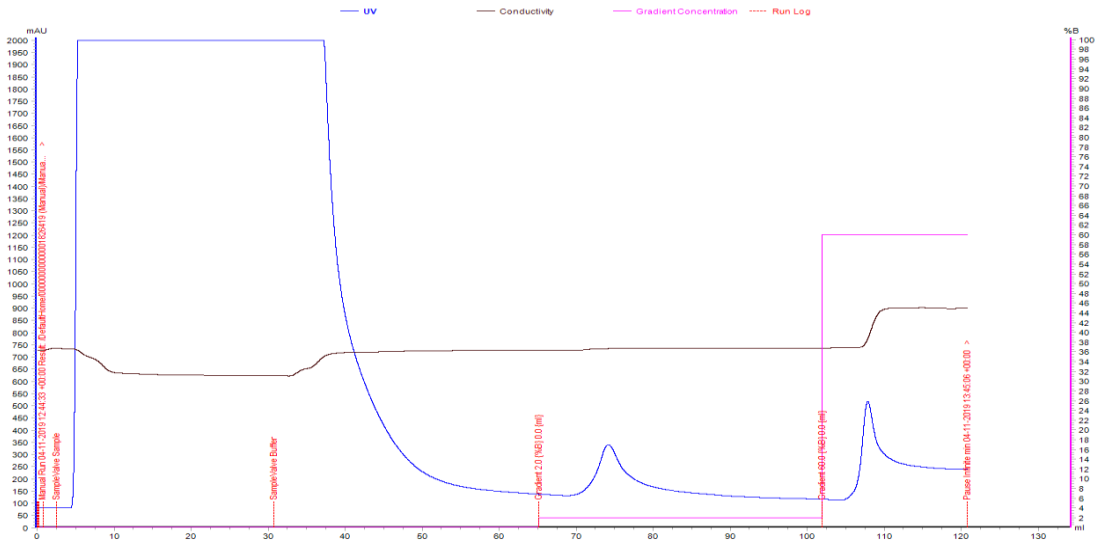


Figure 4.2.5 – Chromatogram obtained by IMAC, using a TALON crude column, for *hl*GP. The absorbance (mAU) at 280 nm and the percentage of elution buffer (% B) as a function of the elution volume (mL).

By analysing the chromatogram, it is possible to observe an initial peak corresponding to the flow through and, therefore, to the proteins present in the soluble extract that have no affinity for the metal ions immobilized in the column. The peak between 70-80 mL corresponds to the washing step with 2% of elution buffer (50 mM sodium phosphate pH 7.5, 500 mM NaCl, 10% glycerol, 2 mM $MgCl_2$, 5 mM b-mercaptoethanol and 500 mM imidazol) to elute proteins with low affinity for the column by increasing imidazol concentration. The peak between 105-110 mL corresponds to the elution of the remaining protein when it is added 60% of elution buffer.

An SDS-PAGE gel was performed to verify the purity of the samples (Figure 4.2.6).

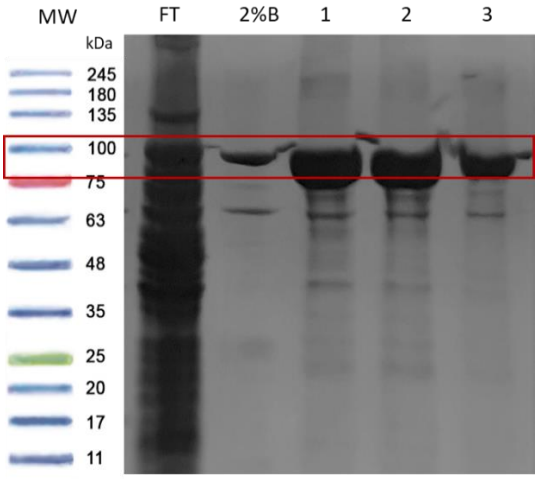


Figure 4.2.6 – SDS-PAGE of the fractions collected by IMAC with TALON crude column. MW – NZY Colour Protein Marker II; FT – Flow through; 2% B – Eluted fraction with 35 mM of imidazol; 1 to 3 – Eluted fractions with 300 mM of imidazol.

The gel shows that the protein of interest was mostly eluted at the last peak since a broad band corresponding to the molecular weight of *hl*GP is observed in these fractions. It is possible to observe the existence of other bands of molecular weights corresponding to other proteins, but with less intensity.

It can be concluded that IMAC using the TALON column with coupled cobalt ions is more efficient and specific than the chromatography techniques performed previously, since the protein of interest was much less present in the FT and the eluted fractions containing the *hl*GP have less contamination of other proteins. However, the high purity of the protein is important to carry out future assays and, therefore, it is necessary to remove these contaminating proteins.

A molecular exclusion chromatography (SEC) was performed as described in section 3.3.4 as second purification step. In this chromatography, the molecules are separated by their molecular weight, with those having a higher molecular weight being eluted first. The obtained chromatogram is show in figure 4.2.7.

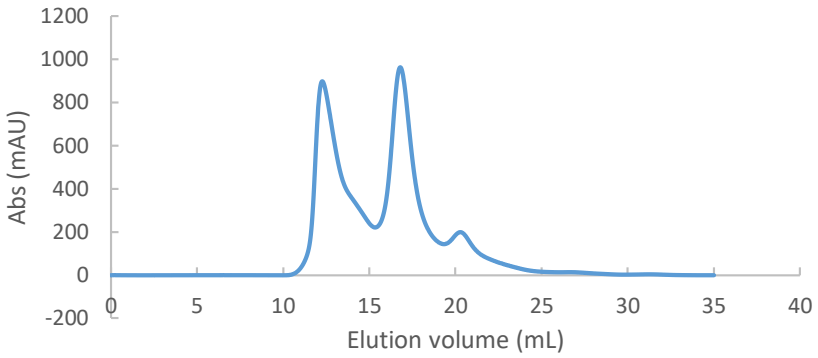


Figure 4.2.7 – Elution profile for the size exclusion chromatography process of *hl*GPb. The chromatography was performed with a Superdex^{TM} 200 Increase 10/300 GL column and in the buffer containing 50 mM HEPES pH 7.5, 10% glycerol, 1 mM DTT, 1 mM EDTA and 2 mM $MgCl_2$**.** The absorbance (mAU) as a function of the elution volume (mL).

Through the chromatogram analysis it is possible to observe three peaks corresponding to the elution of different proteins with different molecular weights.

To determine which fractions contain *hl*GP, an SDS-PAGE gel was made. (Figure 4.2.8)

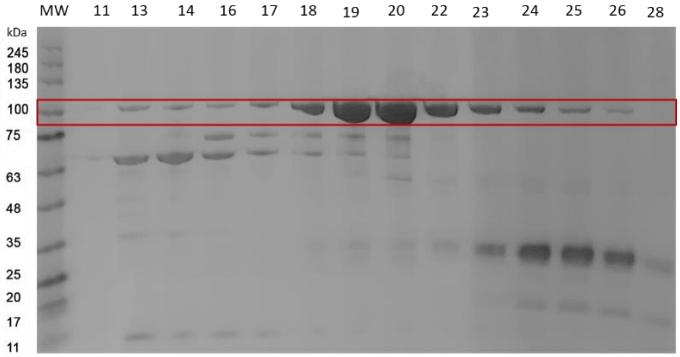


Figure 4.2.8 – SDS-PAGE of the samples collected by the size exclusion chromatography step. MW – NZY Colour Protein Marker II; lanes 11 to 28 – Eluted fractions.

By analysing the SDS-PAGE it is possible to verify that the protein of interest was mostly eluted in the second peak. However, the protein is also present in the remaining peaks, although in a much smaller amount and less pure. The second peak shows some contaminations, but due to these not being very significant, it is possible to perform structural and biophysical assays. These two purification steps allowed to obtain a yield of 0.17 mg/L.

Later, after performing the thermal shift assays (Section 4.3), it was decided to add 1.3 M urea to the sample, before the first purification step, to slightly destabilize the histidine tail and allow it to be more exposed. This way, more protein became coupled to the resin of the column and in the end it was possible to obtain a yield of 0.47 mg/L.

4.3. Thermal Shift Assay (TSA)

The stability of proteins in solution is a major requirement in functional studies involving native

and recombinant proteins. Because of this, the TSA technique was used to determine in which conditions the *hl*GP was more stable.

The TSA allows to study the stability by determining the melting temperature (T_M) of the protein. The higher the T_M value, the more stable the protein will be. In order to obtain the stability curve and the melting temperature value, it is necessary to add a fluorescent probe that binds non-specifically to the hydrophobic surfaces and water quenches its fluorescence. When the protein unfolds due to the increase of temperature, the exposed hydrophobic surfaces bind the probe and the fluorescence increases by excluding water. The stability curve and its midpoint value (T_M) are obtained by slowly increasing the temperature to unfold the protein and measuring the fluorescence at each point [46].

With the TSA, it was possible to test different buffer solutions, pH values and different sodium chloride concentrations (Appendix 7.3). The buffer solutions in which the protein had a higher melting temperature are shown in the figure 4.3.1 and table 4.3.1.

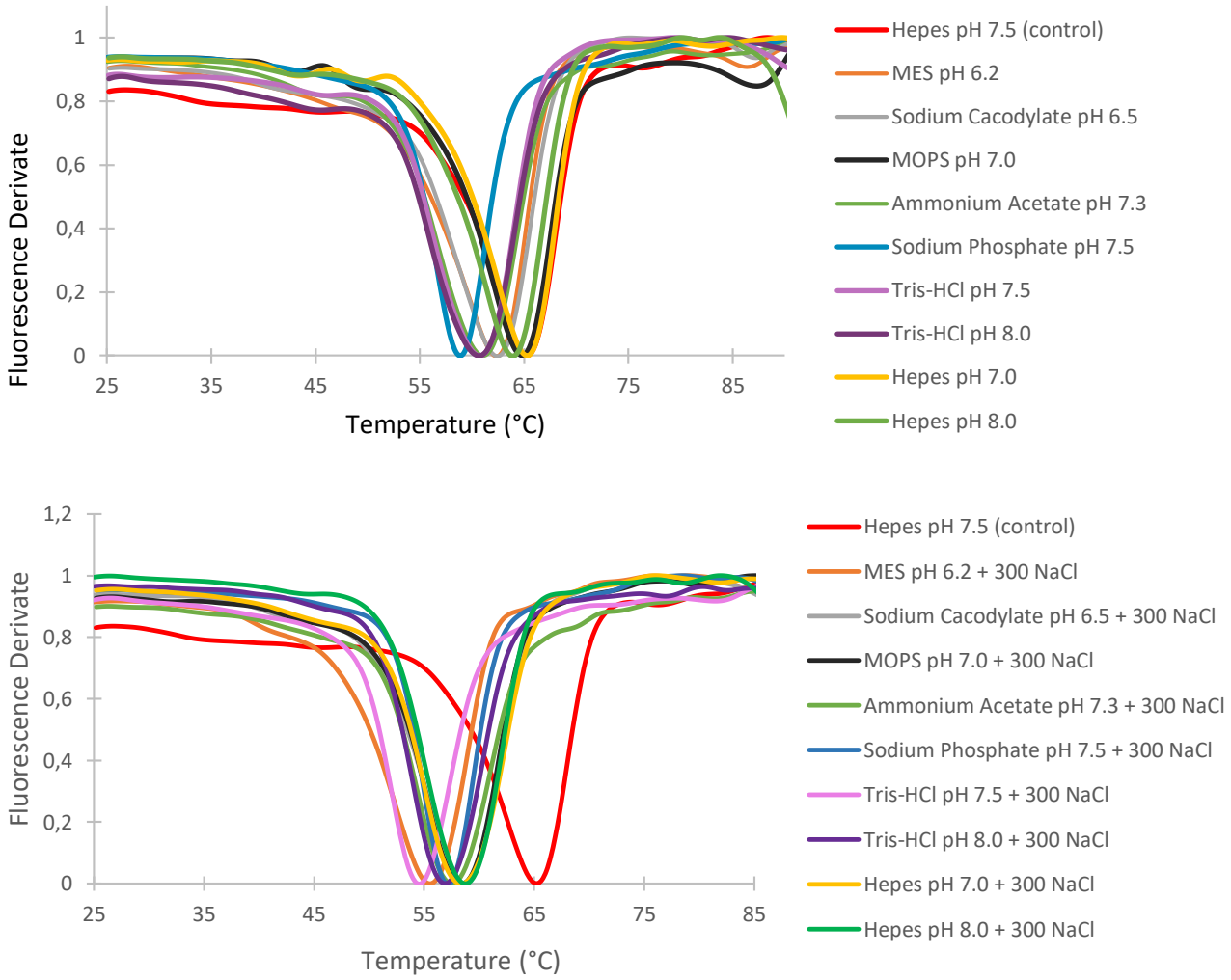


Figure 4.3.1 – TSA results for *hl*GP with the best buffers. The fluorescence derivative is normalized and it is represented as a function of temperature.

Table 4.3.1 – Melting temperatures for hIGP from the best buffer solutions.

Buffer solutions	Melting Temperature (°C)
HEPES pH 7.5	**65.34**
MES pH 6.2	62.17
Sodium Cacodylate pH 6.5	62.63
MOPS pH 7.0	64.76
Ammonium Acetate pH 7.3	61.22
Sodium Phosphate pH 7.5	58.85
Tris-HCl pH 7.5	60.90
Tris-HCl pH 8.0	60.57
HEPES pH 7.0	65.20
HEPES pH 8.0	63.75
MES pH 6.2 + 300 mM NaCl	55.59
Sodium Cacodylate pH 6.5 + 300 mM NaCl	58.45
MOPS pH 7.0 + 300 mM NaCl	58.54
Ammonium Acetate pH 7.3 + 300 mM NaCl	57.40
Sodium Phosphate pH 7.5 + 300 mM NaCl	57.02
Tris-HCl pH 7.5 + 300 mM NaCl	54.70
Tris-HCl pH 8.0 + 300 mM NaCl	56.79
HEPES pH 7.0 + 300 mM NaCl	58.16
HEPES pH 8.0 + 300 mM NaCl	58.77

Through the analysis of the figure 4.3.1 and the melting temperature values in table 4.3.1, it is possible to conclude that the best buffer that stabilizes the protein is HEPES pH 7.5, with a melting temperature of 65.34°C.

After determining the best buffers, new TSA assays were performed with different additives in order to increase the stability of the protein even more and, since the Tris-HCl pH 7.5 buffer is the one that will be used in the MST experiment (HEPES buffer showed some adherence to capillaries) (Section 4.4), the additives were be tested with the protein in this. The most promising additives are shown in the figure 4.3.2 and in the table 4.3.2.

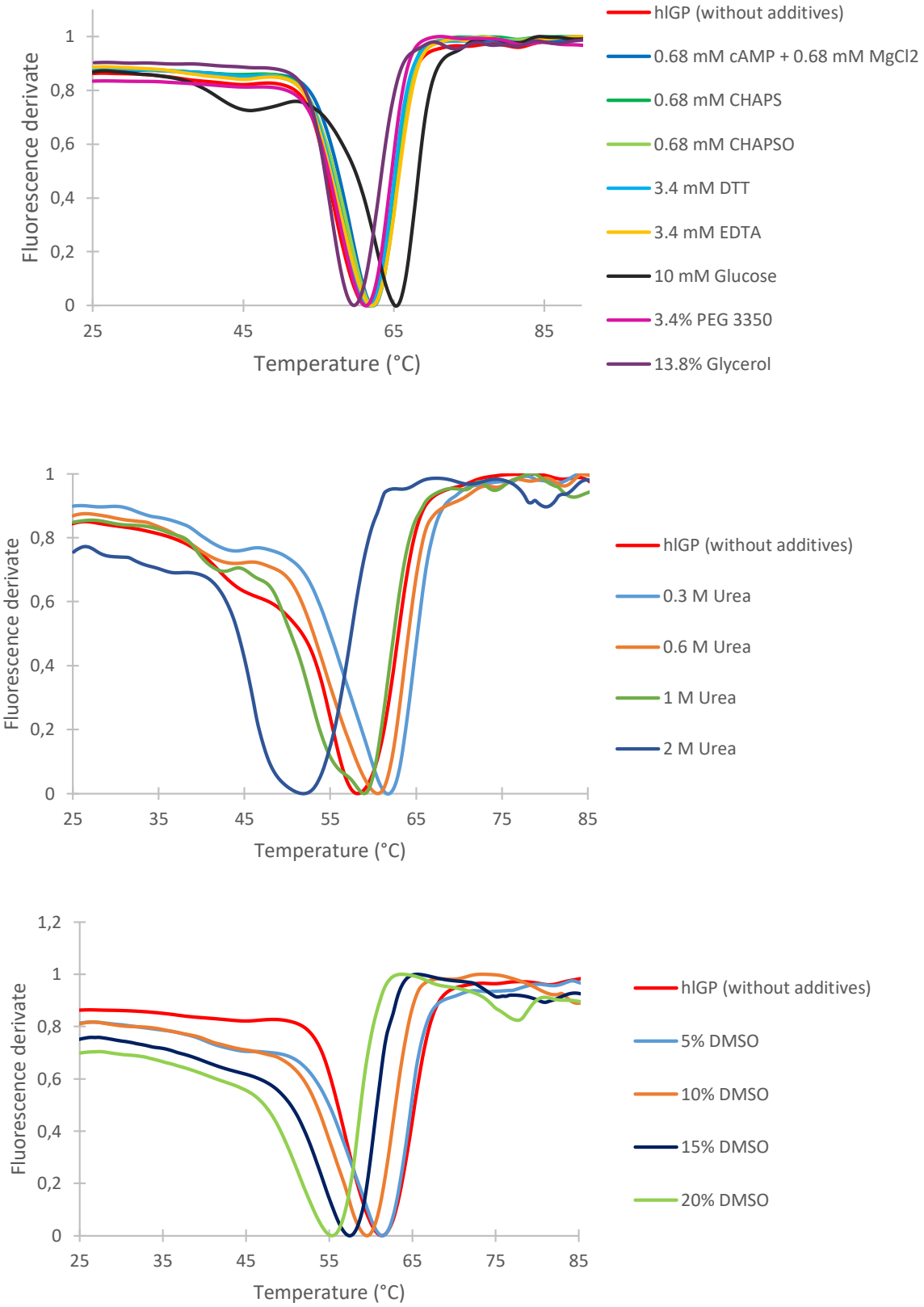


Figure 4.3.2 – TSA results for *hl*GP with the most promising additives. The fluorescence derivative is normalized and it is represented as a function of temperature. The assay was performed with 50 mM Tris-HCl pH 7.5.

Table 4.3.2 – Melting temperatures for *hl*GP with the most promising additives.

Additive solutions	Melting Temperature (°C)
*hl*GP (without additives)	61.48
0.68 mM cAMP + 0.68 mM $MgCl_2$	62.33
0.68 mM CHAPS	61.88
0.68 mM CHAPSO	62.30
3.4 mM DTT	61.54
3.4 mM EDTA	61.87
10 mM Glucose	65.17
3.4% PEG 3350	61.12
13.8% Glycerol	59.71
0.3 M Urea	61.56
0.6 M Urea	60.38
1 M Urea	58.83
2 M Urea	52.03
5% DMSO	61.43
10% DMSO	59.47
15% DMSO	57.48
20% DMSO	55.53

This TSA assay allowed to obtain very important results. First, the addition of DTT, EDTA and $MgCl_2$ help in the stabilization of the protein and, therefore, their addition in the purification process is an asset.

Second, it is possible to see that the addition of DMSO decreases the melting temperature of the protein but that for low concentrations (< 5%), this decrease is not significant. This is an important result since low concentrations of DMSO will be used to solubilize the ligands and to perform the MST assays.

Third, the melting temperature and *hl*GP stability increases with the addition of glucose and AMP. Since these compounds are physiological ligands of the enzyme, increased stability indicates that the *hl*GP is binding and that it is structurally functional.

Finally, the assay with different concentrations of urea shows that, between 1 M and 2 M, the protein stability decreases by several degrees without denaturing. Therefore, it was decided to add 1.3 M of urea to the protein sample before the first purification step, to slightly destabilize the protein and help to expose more the histidine tag so that more protein became coupled to the resin of the column.

4.4. MicroScale Thermophoresis (MST)

The MicroScale Thermophoresis (MST) is a powerful and sensitive technique based on the movement of molecules during a microscopic temperature gradient, known as thermophoresis. The temperature gradient is achieved by increasing the temperature 2-6°C, in a volume with a diameter of 50 µm, using an infrared laser. The direct movement of particles due to the gradient results in a change

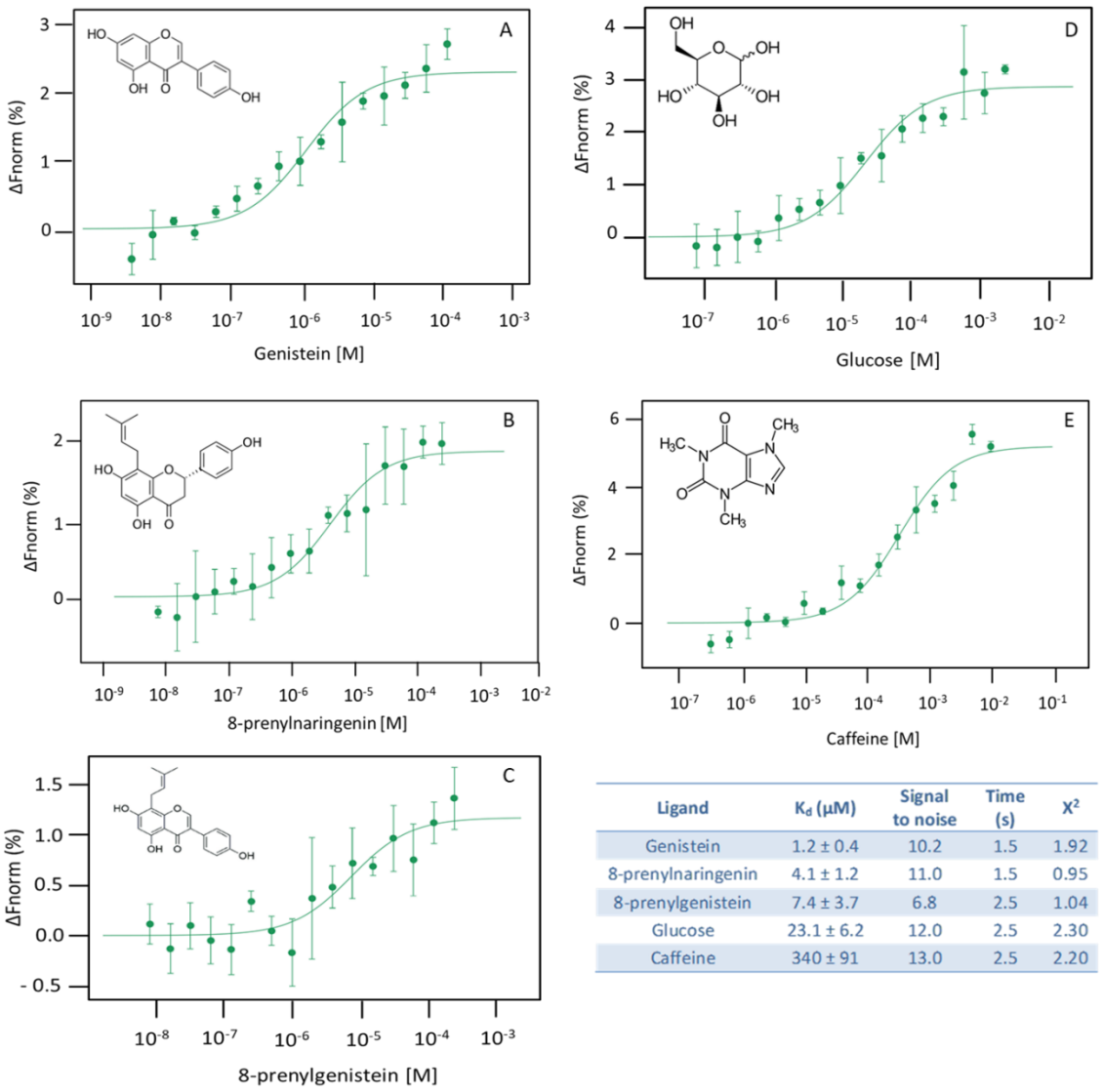
of concentration that is detected by fluorescence. This fluorescence may be intrinsic to the protein or it may be due to a probe that is covalently attached to the protein. MST measurements can be carried out in any kind of buffer, even in plasma, serum, cell lysate, mucus, urine, or other environmental matrices [36].

Both intensity change of fluorescent molecules and thermophoresis vary with three key parameters that are influenced by binding between the fluorescent target and the non-fluorescent ligand molecule: molecular size, molecular charge and hydration shell of the target molecule.

During the assay, the MST signal will detect the binding by a quantification of the change in the normalized fluorescence of the unbound and bound states. This way, through the signal to noise value, it is possible to determine if the binding is happening. According to the software used, the signal to noise should have more than 5 response amplitude units (amplitude is the difference between bound and unbound state) to conclude binding. The technique also allows to quantify high affinity interactions of proteins towards low molecular weight ligands with dissociation constants (K_d) in the order of picomolar. The data generation is fast and precise, and the data output is comparable to other biophysical methods [47].

The MST technique was used to validate the theoretical data obtained by our collaborators, work led by Dr Natércia Brás at Faculdade de Ciências da Universidade do Porto (FCUP). They employed Molecular Dynamics simulations and end-point free-energy MM-PBSA and MM-GBSA approaches to provide the atomistic and energetic details of the binding process of several natural flavonoids against the inhibitory site of GP.

For MST assays, genistein, 8PG and 8PN compounds were chosen due to their good inhibitory potency estimated by computational binding free-energies, as well as due to their easier solubilization in aqueous solution and commercially availability. Attempts to render the parent compound ellagic acid soluble at the concentrations and buffer conditions compatible with the MST experiment were not successful, and only a nonspecific MST signal was obtained (data not shown). Glucose (positive control), a known competitive inhibitor of GP, was chosen to assess *hl*GPb function, and caffeine was chosen as a low affinity inhibitor at the inhibitor site, for comparison purposes. A negative control with a peptide was also performed. In the MST experiments, the concentration of the His-tag-labeled *hl*GPb was kept constant (50 nM), while the concentration of the non-labeled binding partner was varied between 0.25 mM and 20 mM. The results are shown in figure 4.4.1.



Ligand	K_d (µM)	Signal to noise	Time (s)	X^2
Genistein	1.2 ± 0.4	10.2	1.5	1.92
8-prenylnaringenin	4.1 ± 1.2	11.0	1.5	0.95
8-prenylgenistein	7.4 ± 3.7	6.8	2.5	1.04
Glucose	23.1 ± 6.2	12.0	2.5	2.30
Caffeine	340 ± 91	13.0	2.5	2.20

Figure 4.4.1 – MST quantification of the *hl*GPb interaction with the compounds: A – genistein, B – 8-prenylnaringenin, C – 8-prenylgenistein, D – glucose and E – caffeine. Error bars represent the standard deviation of each data point calculated form three independent thermophoresis measurements. Compound concentrations were varied from 3.8 nM to 125 µM, 7.6 nM to 250 µM, 7.6 nM to 250 µM, 76.3 nM to 2.5 mM and 0.3 µM to 10 mM, respectively. All assays were carried out in 2.5% DMSO with a constant concentration of 50 nM for the labelled *hl*GPb. K_d obtained from MST for the compounds are shown in the table on the right.

The MST results allowed to observe binding between the protein and the different compounds as the signal to noise value is higher than 5 for all assays. The results showed that *hl*GPb has a greater affinity for the compounds genistein, 8PN and 8PG, with K_d values of 1.2 ± 0.4 µM, 4.1 ± 1.2 µM and 7.4 ± 3.7 µM respectively. These results are quite promising as genistein, 8PN and 8PG showed to be good inhibitors for *hl*GP and may help in the treatment of diabetes.

The K_d value determined for glucose was 23.1 ± 6.2 µM which was in agreement with that reported for the related compound, glucose-6-P (K_d around 20 µM using electron spin resonance (ESR) spectroscopy) [48]. Caffeine was the compound with the lowest affinity, with a K_d of 340 ± 91 µM, showing that the enzyme binds 100 times less to caffeine when compared to the other inhibitors.

The affinity of caffeine determined by MST was also compared with data reported in the

literature. The K_d for *hl*GP measured by surface plasmon resonance (SPR) is around 100 µM [49], which is within the same order of magnitude of the K_d determined by MST in our analysis.

Thus, the results obtained using MST are within the same range of K_d values reported in the literature despite experimental set-up differences. This reinforces the confidence on the determined experimental affinity values by MST.

The experimental affinity values were then compared with the computational data obtained by our collaborators. (Table 4.4.1.)

Table 4.4.1 – Experimental K_d obtained by MST and calculated binding free energies (in kcal/mol) for the compounds genistein, 8PN and 8PG.

Ligand	K_d (µM)	$\Delta\Delta G_{binding}$ (MM-GBSA)
Genistein	1.2 ± 0.4	-4.4 ± 0.4
8-prenylnaringenin (8PN)	4.1 ± 1.2	0.4 ± 0.4
8-prenylgenistein (8PG)	7.4 ± 3.7	-5.9 ± 0.3

The experimental affinity values are in general agreement with our computational data, where genistein is pointed as a promising inhibitor. The confirmation of the great potency of genistein is highly relevant due to past data suggesting its non-significant inhibitory effect [34]. Since all the three compounds have a K_d in the same order of magnitude we can conclude that all three are promising compounds for inhibition of *hl*GP.

4.5. Activity assay – malachite green phosphate assay

Malachite green phosphate assay is a simple, sensitive and non-radioactive method for measuring inorganic phosphate in aqueous solutions. The kinetic analysis of *hl*GPb and *hl*GPa was done by quantifying the phosphate released in glycogen synthesis. Our goals are to verify the activity of the recombinant enzyme, determine the initial rate constants and perform tests with the inhibitors to determine the inhibition constants (K_i) of the compounds for *hl*GP.

Initially, a calibration curve was made using different concentrations of phosphate in the assay buffer (Appendix 7.4).

Then, to confirm that the phosphorylation of the protein had been achieved, an assay with the two conformations of the enzyme was carried out in order to be able to compare their activities (Figure 4.5.1). Controls without protein were also performed.

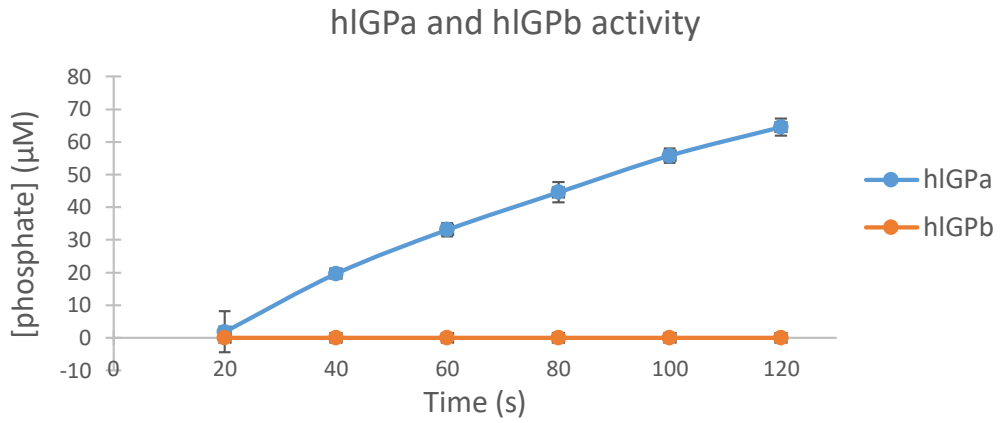


Figure 4.5.1. Activity of dephosphorylated (*hl*GPb) and phosphorylated (*hl*GPa) forms. It was used 30 ng of both proteins *hl*GP and activity was measured in the direction of glycogen synthesis by the release of phosphate from glucose-1-phosphate at 22°C. The buffer used contained 50 mM HEPES pH 7.5, 100 mM KCl, 2.5 mM EDTA, 2.5 mM $MgCl_2$, 0.5 mM AMP, 1 mg/ml glycogen and 0.5 mM glucose-1-phosphate. Error bars represent the standard deviation of each data point calculated form three independent measurements.

The activity assay allowed to verify that the phosphorylated form (*hl*GPa) showed activity and its functional since there is phosphate formation, and therefore, glycogen synthesis. On the contrary, the dephosphorylated form (*hl*GPb) reaction didn't show phosphate release, indicating that the enzyme is inactive. Since phosphorylation of serine 14 promotes the conversion of the enzyme to a more relaxed and active state, these results are in agreement with the expected and proved to be of great importance since they confirm that we were able to purified and obtain both forms of the functional protein.

A preliminary assay with caffeine was done to check if there was a decrease in the release of phosphate by *hl*GPa and a decrease in activity. This compound was chosen due to easier solubilization in aqueous solution and availability in laboratory (Figure 4.5.2).

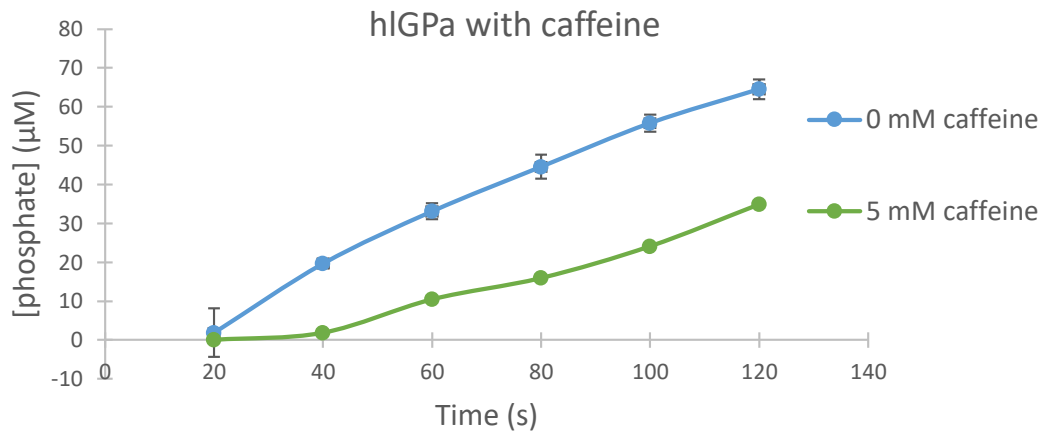


Figure 4.5.2. Activity assay with *hl*GPa without and with 5 mM of caffeine. It was used 30 ng of *hl*GPa and activity was measured in the direction of glycogen synthesis by the release of phosphate from glucose-1-phosphate at 22°C. The buffer used contained 50 mM HEPES pH 7.5, 100 mM KCl, 2.5 mM EDTA, 2.5 mM $MgCl_2$, 0.5 mM AMP, 1 mg/ml glycogen and 0.5 mM glucose-1-phosphate. Error bars represent the standard deviation of each data point calculated form three independent measurements. Only one measurement of caffeine was done.

As expected, the experiment showed that caffeine inhibited the enzyme since there was a decrease in the release of phosphate. However, the inhibition was not complete, which may have been due to the fact that caffeine is a weak inhibitor of *hl*GP. These assays have yet to be done in triplicate but seem to be promising. In future is intended to do the same assay with the compounds genistein, 8PN, 8PG and glucose. The experiments for determining the initial rate conditions are already being prepared and optimized.

4.6. Protein crystallization assays and diffraction experiment

To try to obtain the three-dimensional structure of the recombinant *hl*GPb, two screens were tested under the conditions described in section 3.7.1. After three weeks, the drops were observed and a multiple crystal had formed in a solution containing 1 M sodium phosphate and 0.1 M sodium acetate pH 4.5 (Figure 4.6.1).

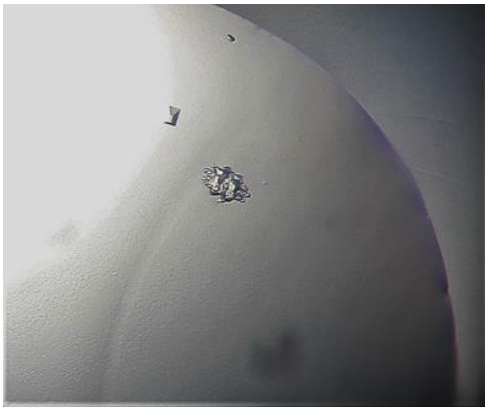


Figure 4.6.1 – Multiple crystal obtained for *hl*GPb. The assay was done at 20°C with sitting drop technique and a crystallization solution containing 1 M sodium phosphate and 0.1 M sodium acetate pH 4.5.

The crystal was frozen in liquid nitrogen in the presence of the cryoprotective solution and analyzed by synchrotron radiation (DLS). During the diffraction experiment it was possible to observe reflection spots far apart and with great intensity, indicative of salt crystals.

Other crystallization assays were carried to try to obtain the structure of the protein in complex with the inhibitors and, therefore, the method of co-crystallization was used. In this method the ligand is added to the protein to form a complex that is subsequently used in crystallization assays. This method is widely used when the compounds are quite insoluble or the protein aggregates easily [50].

In the co-crystallization assay with genistein, three different commercial screens were tested under the conditions described in section 3.7.1. After three weeks, it was observed the appearance of needle-shaped crystals in conditions that contained high concentrations of sodium acetate or lithium sulphate indicating that we were probably in the presence of salt crystals, which was then verified in the diffraction experiment.

For co-crystallization with 8PN, it was decided to try the Structure screen since no protein

crystals were observed on the previous screens. After three weeks, crystals appeared in three different crystallization solutions: 1 – 2 M ammonium sulphate (Figure 4.6.2 – A); 2 – 1 M lithium sulphate, 0.1 M tris pH 8.5 and 0.01 M nickel chloride (Figure 4.6.2 – B); 3 – 2 M sodium formate and 0.1 mM HEPES pH 7.5 (Figure 4.6.2 – C). At the time, crystalline growth in the control drops was not detected, which suggests that the crystals in the sample drop could be protein crystals. The crystals were frozen in the presence of the respective cryoprotective solution and analyzed by synchrotron radiation (DLS). Unfortunately, the diffraction pattern indicated that they were salt crystals.

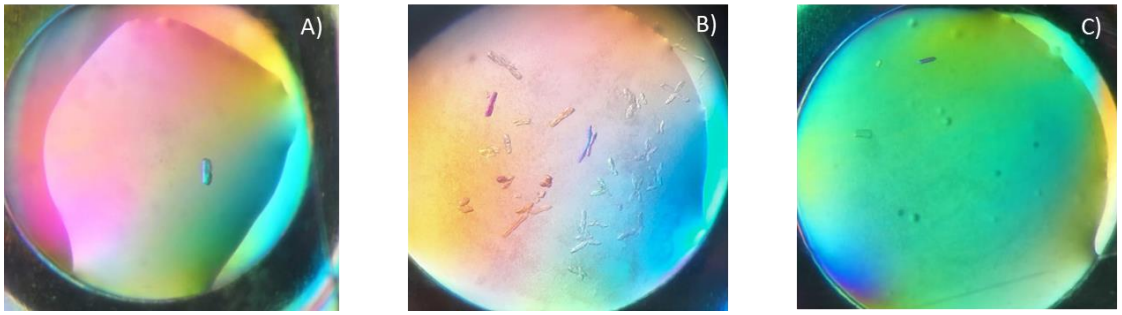


Figure 4.6.2 – Crystals obtained by co-crystallization of *hl*GPb with 1 mM of 8PG. The assays were done at 20°C with sitting drop technique and different crystallization solutions. A) 2 M ammonium sulphate; B) 1 M lithium sulphate, 0.1 M tris pH 8.5 and 0.01 M nickel chloride; C) 2 M sodium formate and 0.1 mM HEPES pH 7.5.

Therefore, it is necessary to optimize the crystallization conditions and try new crystallization solutions in order to obtain protein crystals that allow to determine the three-dimensional structure of *hl*GPb alone and in complex with the inhibitors.

5. Conclusions and future perspectives

Type 2 diabetes mellitus is a serious problem worldwide that continues to increase. The resistance of the cells to the hormone insulin prevents glucose from entering the cells, leading to higher levels of glucose in the bloodstream (hyperglycaemia). In Portugal are diagnosed more than 60 000 people per year and it is estimated, that in 2019, more than 4 million people have died worldwide due to diabetes complications. Due to these numbers, that continues to increase, and the side effects and ineffectiveness of current therapies, new drugs and treatments continue to be studied.

The inhibition of the enzyme glycogen phosphorylase (GP) appears to be a good approach since this enzyme is responsible for the breakdown of glycogen and can contribute greatly to the increase of glucose in blood. Our collaborators had already obtained computational data from the interaction of human liver GP with some possible natural inhibitors (flavonoids). Therefore, the main goal of this thesis was to clarify and validate the binding and inhibitory potential of hlGP with the compounds genistein, 8-prenylgenistein (8PG) and 8-prenylnaringenin (8PN) through MST experiments and kinetic assays.

To achieve this goal, it was necessary to optimize the expression and purification of recombinant *hl*GP. Three different bacteria strains from *E. coli*, Rosetta 2(DE3), BL21 (DE3) and Tuner (DE3), were tested. Tuner cells were the ones that produce more soluble protein with a temperature of induction of 16°C with 0.5 mM of IPTG.

For the optimization of the first purification step, an affinity chromatography with nickel ions and an ion exchange chromatography were attempted. However, the protein did not come out in a single peak and was highly contaminated. An affinity chromatography with cobalt ions was then attempted and it was possible to obtain a single peak with high purity corresponding to the *hl*GPb.

Part of the protein obtained was phosphorylated in order to also obtain the most active form (*hl*GPa), and both proteins were then purified by size exclusion chromatography. In literature, as far as we know, only protocols using ion-change chromatography or affinity chromatography with copper ions without histidine tail are published [51], [52] making it the first time that a histidine-tailed *hl*GP purification protocol has been described.

The protein buffer and additives were optimized by thermofluor assays and the best buffer for stabilize the *hl*GP protein was HEPES pH 7.5 with the additives DTT, EDTA and $MgCl_2$. This allowed to obtain a yield of 0.47 mg/L.

The MST experiments with *hl*GPb allowed to determine the dissociation constants, K_d, for the compounds genistein (1.2 ± 0.4 μM), 8PN (4.1 ± 1.2 μM) and 8PG (7.4 ± 3.7 μM). These experimental affinity values are in general good agreement with our computational data, where these three compounds are pointed as good inhibitors of the enzyme.

The kinetic assays are still in the process of optimization, but preliminary assays with both forms of the enzyme showed that *hl*GPa has a high activity while *hl*GPb did not show activity. An assay with *hl*GPa and caffeine (inhibitor) showed reduced enzyme activity. These results have shown to be quite

promising and it is intended in the future to determine the parameters V0, Vm and Km of the reaction, and to determine the inhibitory constants to the different inhibitors.

Crystallization assays were also attempted with *hl*GPb and with the inhibitors genistein and 8PN but still need to be optimized in order to determine the best crystallization conditions. In the future, it is intended to test different concentrations of protein and try to obtain crystals at a lower temperature of 4°C. It is also intended to perform crystallization assays with the protein in its more relaxed form and with the inhibitors for characterize structurally and clarify the inhibition reaction mechanism of enzyme.

Overall, the work described in this Thesis led to important conclusions suggesting natural flavonoids as potential inhibitors for human liver GP and that these can be explored as dietary supplements. These would be potentially useful in the diabetes treatments as also in decreasing the impact of diabetes in the patient's well-being.

The pioneer study, combining computational studies to identify potential inhibitory compounds, with interaction affinity studies by MST for their experimental validation, led to the publication of an article (Appendix 7.5) in the Bioorganic Chemistry journal with the title "Combined in Silico and in Vitro Studies to Identify Novel Antidiabetic Flavonoids Targeting Glycogen Phosphorylase".

6. References

[1] I. D. F. D. Atlas, *463 PEOPLE LIVING WITH DIABETES million*. 2019.
[2] "WHO - Diabetes." [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/diabetes. [Accessed: 12-Jun-2020].
[3] M. C. Petersen and G. I. Shulman, "Mechanisms of insulin action and insulin resistance," *Physiol. Rev.*, vol. 98, no. 4, pp. 2133–2223, 2018.
[4] M. Stumvoll, B. J. Goldstein, and T. W. Van Haeften, "Type 2 diabetes: Principles of pathogenesis and therapy," *Lancet*, vol. 365, no. 9467, pp. 1333–1346, 2005.
[5] R. A. Haeusler, S. Camastra, B. Astiarraga, M. Nannipieri, M. Anselmino, and E. Ferrannini, "Decreased expression of hepatic glucokinase in type 2 diabetes," *Mol. Metab.*, vol. 4, no. 3, pp. 222–226, 2015.
[6] S. R. Naik and G. R. Kokil, *Development and discovery avenues in bioactive natural products for glycemic novel therapeutics*, 1st ed., vol. 39. Copyright © 2013 Elsevier B.V. All rights reserved., 2013.
[7] M. Krssak *et al.*, "Alterations in postprandial hepatic glycogen metabolism in type 2 diabetes," *Diabetes*, vol. 53, no. 12, pp. 3048–3056, 2004.
[8] L. Somsak, V. Nagy, Z. Hadady, N. Felfoldi, T. Docsa, and P. Gergely, "Recent Developments in the Synthesis and Evaluation of Glucose Analog Inhibitors of Glycogen Phosphorylases as Potential Antidiabetic Agents," *Front. Med. Chem. - Online*, vol. 2, no. 1, pp. 253–272, 2005.
[9] A. A. Spasov, N. I. Chepljaeva, and E. S. Vorob'ev, "Glycogen phosphorylase inhibitors in the regulation of carbohydrate metabolism in type 2 diabetes," *Russ. J. Bioorganic Chem.*, vol. 42, no. 2, p. 133–142A, 2016.
[10] R. Perfetti, P. S. Barnett, R. Mathur, and J. M. Egan, "Novel therapeutic strategies for the treatment of Type 2 diabetes," *Diabetes. Metab. Rev.*, vol. 14, no. 3, pp. 207–225, 1998.
[11] N. F. Brás, P. A. Fernandes, and M. J. Ramos, "Understanding the Rate-Limiting Step of Glycogenolysis by Using QM/MM Calculations on Human Glycogen Phosphorylase," *ChemMedChem*, vol. 13, no. 15, pp. 1608–1616, 2018.
[12] S. Y. Tan *et al.*, "Type 1 and 2 diabetes mellitus: A review on current treatment approach and gene therapy as potential intervention," *Diabetes Metab. Syndr. Clin. Res. Rev.*, vol. 13, no. 1, pp. 364–372, 2019.
[13] L. Somsak *et al.*, "New Inhibitors of Glycogen Phosphorylase as Potential Antidiabetic Agents," *Curr. Med. Chem.*, vol. 15, no. 28, pp. 2933–2983, 2008.
[14] E. G. Krebs and E. H. Fischer, "The phosphorylase b to a converting enzyme of rabbit skeletal muscle," *BBA - Gen. Subj.*, vol. 20, no. C, pp. 150–157, 1956.
[15] D. Barford, S. H. Hu, and L. N. Johnson, "Structural mechanism for glycogen phosphorylase control by phosphorylation and AMP," *J. Mol. Biol.*, vol. 218, no. 1, pp. 233–260, 1991.
[16] J. L. Treadway, P. Mendys, and D. J. Hoover, "Expert Opinion on Investigational Drugs Glycogen phosphorylase inhibitors for treatment of type 2 diabetes mellitus," *Expert Opin. Investig. Drugs*, vol. 10, no. 3, pp. 439–454, 2001.
[17] N. B. Livanova, N. A. Chebotareva, T. B. Eronina, and B. I. Kurganov, "Review: Pyridoxal 5'-phosphate as a catalytic and conformational cofactor of muscle glycogen phosphorylase b," *Biokhimiya*, vol. 67, no. 10, pp. 1317–1327, 2002.
[18] D. Palm, H. W. Klein, R. Schinzel, M. Buehner, and E. J. M. Helmreich, "The Role of Pyridoxal 5'-Phosphate in Glycogen Phosphorylase Catalysis," *Biochemistry*, vol. 29, no. 5, pp. 1099–1107, 1990.
[19] J. L. Buchbinder and R. J. Fletterickr, "Role of the active site gate of glycogen phosphorylase in allosteric inhibition and substrate binding," *J. Biol. Chem.*, vol. 271, no. 37, pp. 22305–22309, 1996.
[20] L. N. Johnson, "Glycogen phosphorylase: control by phosphorylation and allosteric effectors," *FASEB J.*, vol. 6, no. 6, pp. 2274–2282, 1992.
[21] L. Agius, "Role of glycogen phosphorylase in liver glycogen metabolism," *Mol. Aspects Med.*, vol. 46, pp. 34–45, 2015.
[22] J. W. Hudson, K. L. Hefferon, and M. M. Crerar, "Comparative analysis of species-independent,

isozyme-specific amino-acid substitutions in mammalian muscle, brain and liver glycogen phosphorylases," *Biochim. Biophys. Acta (BBA)/Protein Struct. Mol.*, vol. 1164, no. 2, pp. 197–208, 1993.
[23] V. L. Rath *et al.*, "Activation of human liver glycogen phosphorylase by alteration of the secondary structure and packing of the catalytic core," *Mol. Cell*, vol. 6, no. 1, pp. 139–148, 2000.
[24] C. B. Newgard, P. K. Hwang, and R. J. Fletterick, "The family of glycogen phosphorylases: Structure and functio," *Crit. Rev. Biochem. Mol. Biol.*, vol. 24, no. 1, pp. 69–99, 1989.
[25] J. L. Treadway, P. Mendys, and D. J. Hoover, "Glycogen phosphorylase inhibitors for treatment of type 2 diabetes mellitus," *Expert Opin. Investig. Drugs*, vol. 10, no. 3, pp. 439–454, 2001.
[26] F. Llavero *et al.*, "Mcardle disease: New insights into its underlying molecular mechanisms," *Int. J. Mol. Sci.*, vol. 20, no. 23, pp. 1–15, 2019.
[27] V. L. Rath *et al.*, "Human liver glycogen phosphorylase inhibitors bind at a new allosteric site," *Chem. Biol.*, vol. 7, no. 9, pp. 677–682, 2000.
[28] J. M. Hayes, *Computer-Aided Discovery of Glycogen Phosphorylase Inhibitors Exploiting Natural Products*. Elsevier Inc., 2017.
[29] N. Oikonomakos, "Glycogen Phosphorylase as a Molecular Target for Type 2 Diabetes Therapy," *Curr. Protein Pept. Sci.*, vol. 3, no. 6, pp. 561–586, 2005.
[30] J. M. Hayes, A. L. Kantsadi, and D. D. Leonidas, "Natural products and their derivatives as inhibitors of glycogen phosphorylase: Potential treatment for type 2 diabetes," *Phytochem. Rev.*, vol. 13, no. 2, pp. 471–498, 2014.
[31] M. Donnier-Maréchal and S. Vidal, "Glycogen phosphorylase inhibitors: A patent review (2013-2015)," *Expert Opin. Ther. Pat.*, vol. 26, no. 2, pp. 199–212, 2016.
[32] A. L. Kantsadi *et al.*, "Biochemical and biological assessment of the inhibitory potency of extracts from vinification byproducts of Vitis vinifera extracts against glycogen phosphorylase," *Food Chem. Toxicol.*, vol. 67, pp. 35–43, 2014.
[33] S. L. Berg JM, Tymoczko JL, "Epinephrine and Glucagon Signal the Need for Glycogen Breakdown.," in *Biochemistry*, 5th ed., W. H. Freeman, Ed. New York, 2002.
[34] S. Jakobs, D. Fridrich, S. Hofem, G. Pahlke, and G. Eisenbrand, "Natural flavonoids are potent inhibitors of glycogen phosphorylase," *Mol. Nutr. Food Res.*, vol. 50, no. 1, pp. 52–57, 2006.
[35] A. Froger and J. E. Hall, "Transformation of Plasmid DNA into E. Coli using the heat shock method," *J. Vis. Exp.*, no. 6, p. 2007, 2007.
[36] M. Jerabek-Willemsen *et al.*, "MicroScale Thermophoresis: Interaction analysis and beyond," *J. Mol. Struct.*, vol. 1077, pp. 101–113, 2014.
[37] T. P. Geladopoulos, T. G. Sotiroudis, and A. E. Evangelopoulos, "A malachite green colorimetric assay for protein phosphatase activity," *Anal. Biochem.*, vol. 192, no. 1, pp. 112–116, 1991.
[38] P. A. Lanzetta, L. J. Alvarez, P. S. Reinach, and O. A. Candia, "An improved assay for nanomole amounts of inorganic phosphate," *Anal. Biochem.*, vol. 100, no. 1, pp. 95–97, 1979.
[39] J. C. Brooks-Bartlett and E. F. Garman, "The nobel science: One hundred years of crystallography," *Interdiscip. Sci. Rev.*, vol. 40, no. 3, pp. 244–264, 2015.
[40] A. L. Carvalho, J. Trincão, and M. J. Romão, "X-ray crystallography in drug discovery.," *Methods Mol. Biol.*, vol. 572, pp. 31–56, 2009.
[41] G. Rhodes and J. Cooper, *Crystallography Made Crystal Clear 3e*. 1994.
[42] C. Growth, "Hanging Drop Vapor Diffusion Crystallization Hanging Drop Vapor Diffusion Crystallization," *Hapt. Res.*, vol. 1, pp. 1–2, 2017.
[43] Hampton research, "Sitting Drop Vapor Diffusion Crystallization," *Handbook*, vol. i, pp. 1–3, 2001.
[44] NOVAGEN, *pET Systems Manual*, 11th ed. 2006.
[45] GE Healthcare Life Sciences, *HiTrap™ TALON® crude, 1 ml and 5 ml TALON Superflow™*. 2011.
[46] U. B. Ericsson, B. M. Hallberg, G. T. DeTitta, N. Dekker, and P. Nordlund, "Thermofluor-based high-throughput stability optimization of proteins for structural studies," *Anal. Biochem.*, vol. 357, no. 2, pp. 289–298, 2006.
[47] M. Jerabek-Willemsen, C. J. Wienken, D. Braun, P. Baaske, and S. Duhr, "Molecular interaction studies using microscale thermophoresis," *Assay Drug Dev. Technol.*, vol. 9, no. 4, pp. 342–353,

2011.
[48] M. K. Battersby and G. K. Radda, “The stereospecificity of the glucose-6-phosphate binding site of glycogen phosphorylase b,” *FEBS Lett.*, vol. 72, no. 2, pp. 319–322, 1976.
[49] J. L. Ekstrom *et al.*, “Structure-activity analysis of the purine binding site of human liver glycogen phosphorylase,” *Chem. Biol.*, vol. 9, no. 8, pp. 915–924, 2002.
[50] A. M. Hassell *et al.*, “Crystallization of protein-ligand complexes,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 63, no. 1, pp. 72–79, 2006.
[51] C. B. H. Luong, M. F. Browner, R. J. Fletterick, and B. L. Haymore, “Purification of glycogen phosphorylase isozymes by metal-affinity chromatography,” *J. Chromatogr. B Biomed. Sci. Appl.*, vol. 584, no. 1, pp. 77–84, 1992.
[52] B. Lederer and W. Stalmans, “Human liver glycogen phosphorylase. Kinetic properties and assay in biopsy specimens,” *Biochem. J.*, vol. 159, no. 3, pp. 689–695, 1976.

7. Appendix

7.1. Appendix 1 –*HL*GP express vector and recombinant protein sequence.

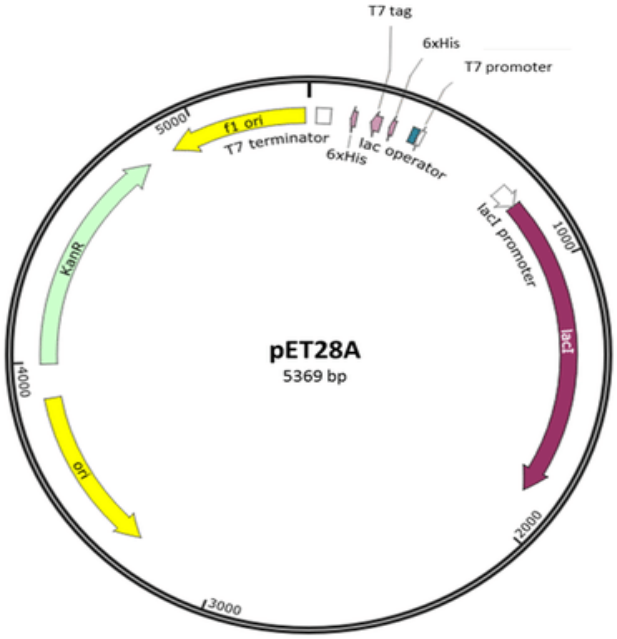


Figure 7.1 – Plasmid map of pET28A. The plasmid harbours an origin of replication (ori), a repressor for the *lac* promoter (lacI), a transcriptional promoter from the T7 phage (T7 promoter), a lactose operator (lac operator), an affinity purification tag (6 × His), a T7 transcriptional terminator (T7 terminator), and a kanamycin resistance gene (KanR) [44].

*hl*GP – Recombinant Protein Sequence (nzytech)

MGSSHHHHHHSSGLVPRGSHMASMAKPLTDQEKRRQISIRGIVGVENVAELKKSFNRHLHFTLVKDRNVATTRDYYFA
LAHTVRDHLVGRWIRTQQHYYDKCPKRVYYLSLEFYMGRTLQNTMINLGLQNACDEAIYQLGLDIEELEEIEEDAGLG
NGGLGRLAACFLDSMATLGLAAYGYGIRYEYGIFNQKIRDGWQVEEADDWLRYGNPWEKSRPEFMLPVHFYGKVEHTN
TGTKWIDTQVVLALPYDTPVPGYMNNTVNTMRLWSARAPNDFNLRDFNVGDYIQAVLDRNLAENISRVLYPNDNFFEG
KELRLKQEYFVVAATLQDIIRRFKASKFGSTRGAGTVFDAFPDQVAIQLNDTHPALAIPELMRIFVDIEKLPWSKAWE
LTQKTFAYTNHTVLPEALERWPVDLVEKLLPRHLEIIYEINQKHLDRIVALFPKDVDRLRRMSLIEEEGSKRINMAHL
CIVGSHAVNGVAKIHSDIVKTKVFKDFSELEPDKFQNKTNGITPRRWLLLCNPGLAELIAEKIGEDYVKDLSQLTKLH
SFLGDDVFLRELAKVKQENKLKFSQFLETEYKVKINPSSMFDVQVKRIHEYKRQLLNCLHVITMYNRIKKDPKKLFVP
RTVIIGGKAAPGYHMAKMIIKLITSVADVVNNDPMVGSKLKVIFLENYRVSLAEKVIPATDLSEQISTAGTEASGTGN
MKFMLNGALTIGTMDGANVEMAEEAGEENLFIFGMRIDDVAALDKKGYEAKEYYEALPELKLVIDQIDNGFFSPKQPD
LFKDIINMLFYHDRFKVFADYEAYVKCQDKVSQLYMNPKAWNTMVLKNIAASGKFSSDRTIKEYAQNIWNVEPSDLKI
SLSNESNKVNG N*

7.2. Appendix 2 – Reagents and solutions used throughout the experimental procedure

Table 7.2.1 – Composition of LB medium for expression. The medium was prepared with distilled water and then autoclaved for 20 minutes at 120 ° C.

Reagent	Concentration	Brand
Tryptone	10 g/L	BD Bacto™
NaCl	10 g/L	PanReac AppliChem
Yeast extract	5 g/L	NZYTech

Table 7.2.2 – Composition of TFBI solution used in competent cells preparation. The pH was adjusted to 5.8 with acetic acid and the solution was filtered.

Reagent	Concentration	Brand
KOAc	30 mM	LaborSpirit
RbCl	100 mM	Alfa Aesar
$CaCl_2$	10 mM	Fluka
$MnCl_2$	50 mM	Sigma
Glycerol	15%	LabChem

Table 7.2.3 – Composition of TFBII solution used in competent cells preparation. The pH was adjusted to 6.5 with KOH and the solution was filtered.

Reagent	Concentration	Brand
MOPS	10 mM	Sigma
RbCl	10 mM	Alfa Aesar
$CaCl_2$	75 mM	Fluka
Glycerol	15%	LabChem

7.3. Appendix 3 – Thermofluor shift assay screens

Table 7.3.1 – Thermofluor shift assay buffer screen

	1	2	3	4	5	6	7	8	9	10	11	12
A	Citric acid pH 4.0	Sodium acetate pH 4.5	Sodium citrate pH 4.7	Sodium acetate pH 5.0	Potassium phosphate pH 5.0	Sodium phosphate pH 5.5	Sodium citrate pH 5.5	MES pH 5.8	Potassium phosphate pH 6.0	MES pH 6.2	Sodium phosphate pH 6.5	Sodium cacodylate pH 6.5
B	MES pH 6.5	PIPES pH 6.7	MOPS pH 7.0	HEPES pH 7.0	Ammonium acetate pH 7.3	Sodium phosphate pH 7.5	Tris pH 7.5	Imidazol pH 8.0	HEPES pH 8.0	Tris pH 8.0	Bicine pH 8.5	Tris pH 8.5
C	Bicine pH 9.0	CAPS pH 9.0	Glycine-HCl pH 9.5	Glycine-HCl pH 10.0	CAPS pH 10.5	CAPS pH 11.0	Citric acid pH 4.0 + 300 mM NaCl	Sodium acetate pH 4.5 + 300 mM NaCl	Sodium citrate pH 4.7 + 300 mM NaCl	Sodium acetate pH 5.0 + 300 mM NaCl	Potassium phosphate pH 5.0 + 300 mM NaCl	Sodium phosphate pH 5.5
D	Sodium citrate pH 5.5 + 300 mM NaCl	MES pH 5.8 + 300 mM NaCl	Potassium phosphate pH 6.0 + 300 mM NaCl	MES pH 6.2 + 300 mM NaCl	Sodium phosphate pH 6.5 + 300 mM NaCl	Sodium cacodylate pH 6.5 + 300 mM NaCl	MES pH 6.5 + 300 mM NaCl	PIPES pH 6.7 + 300 mM NaCl	MOPS pH 7.0 + 300 mM NaCl	HEPES pH 7.0 + 300 mM NaCl	Ammonium acetate pH 7.3 + 300 mM NaCl	Sodium phosphate pH 7.5 + 300 mM NaCl
E	Tris pH 7.5 + 300 mM NaCl	Imidazol pH 8.0 + 300 mM NaCl	HEPES pH 8.0 + 300 mM NaCl	Tris pH 8.0 + 300 mM NaCl	Bicine pH 8.5 + 300 mM NaCl	Tris pH 8.5 + 300 mM NaCl	Bicine pH 9.0 + 300 mM NaCl	CAPS pH 9.0 + 300 mM NaCl	Glycine-HCl pH 9.5 + 300 mM NaCl	Glycine-HCl pH 10.0 + 300 mM NaCl	CAPS pH 10.5 + 300 mM NaCl	CAPS pH 11.0 + 300 mM NaCl
F	Citric acid pH 4.0 + 1 M NaCl	Sodium acetate pH 4.5 + 1 M NaCl	Sodium citrate pH 4.7 + 1 M NaCl	Sodium acetate pH 5.0 + 1 M NaCl	Potassium phosphate pH 5.0 + 1 M NaCl	Sodium phosphate pH 5.5 + 1 M NaCl	Sodium citrate pH 5.5 + 1 M NaCl	MES pH 5.8 + 1 M NaCl	Potassium phosphate pH 6.0 + 1 M NaCl	MES pH 6.2 + 1 M NaCl	Sodium phosphate pH 6.5 + 1 M NaCl	Sodium cacodylate pH 6.5 + 1 M NaCl
G	MES pH 6.5 + 1 M NaCl	PIPES pH 6.7 + 1 M NaCl	MOPS pH 7.0 + 1 M NaCl	HEPES pH 7.0 + 1 M NaCl	Ammonium acetate pH 7.3 + 1 M NaCl	Sodium phosphate pH 7.5 + 1 M NaCl	Tris pH 7.5 + 1 M NaCl	Imidazol pH 8.0 + 1 M NaCl	HEPES pH 8.0 + 1 M NaCl	Tris pH 8.0 + 1 M NaCl	Bicine pH 8.5 + 1 M NaCl	Tris pH 8.5 + 1 M NaCl
H	Bicine pH 9.0 + 1 M NaCl	CAPS pH 9.0 + 1 M NaCl	Glycine-HCl pH 9.5 + 1 M NaCl	Glycine-HCl pH 10.0 + 1 M NaCl	CAPS pH 10.5 + 1 M NaCl	CAPS pH 11.0 + 1 M NaCl	Control	Control	Control			

Table 7.3.2 – Thermofluor shift assay additive screen

	1	2	3	4	5	6	7	8	9	10	11	12
A	DMSO 5%	DMSO 10%	DMSO 15%	DMSO 20%	DMSO 25%	DMSO 30%	DMSO 5%	DMSO 10%	DMSO 15%	DMSO 20%	DMSO 25%	DMSO 30%
B	DMSO 5%	DMSO 10%	DMSO 15%	DMSO 20%	DMSO 25%	DMSO 30%	PEG 400 5%	PEG 400 10%	PEG 400 15%	PEG 400 20%	PEG 400 25%	PEG 400 30%
C	PEG 400 5%	PEG 400 10%	PEG 400 15%	PEG 400 20%	PEG 400 25%	PEG 400 30%	PEG 400 5%	PEG 400 10%	PEG 400 15%	PEG 400 20%	PEG 400 25%	PEG 400 30%
D	Ureia 0.3 M	Ureia 0.6 M	Ureia 1 M	Ureia 2 M	Ureia 0.3 M	Ureia 0.6 M	Ureia 1 M	Ureia 2 M	Ureia 0.3 M	Ureia 0.6 M	Ureia 1 M	Ureia 2 M
E	Imidazol 100 mM	Imidazol 500 mM	Imidazol 100 mM	Imidazol 500 mM	Imidazol 100 mM	Imidazol 500 mM	Glucose 5 mM	Glucose 10 mM	Glucose 20 mM	Glucose 30 mM	Glucose 40 mM	Glucose 5 mM
F	Glucose 10 mM	Glucose 20 mM	Glucose 30 mM	Glucose 40 mM	Glucose 5 mM	Glucose 10 mM	Glucose 20 mM	Glucose 30 mM	Glucose 40 mM	Glucose 5 mM + 1 mM AMP + 1 mM $MgCl_2$	Glucose 10 mM + 1 mM AMP + 1 mM $MgCl_2$	Glucose 20 mM + 1 mM AMP + 1 mM $MgCl_2$
G	Glucose 30 Mm + 1 mM AMP + 1 mM $MgCl_2$	Glucose 40 mM + 1 mM AMP + 1 mM $MgCl_2$	Glucose 5 mM + 1 mM AMP + 1 mM $MgCl_2$	Glucose 10 mM + 1 mM AMP + 1 mM $MgCl_2$	Glucose 20 mM + 1 mM AMP + 1 mM $MgCl_2$	Glucose 30 Mm + 1 mM AMP + 1 mM $MgCl_2$	Glucose 40 mM + 1 mM AMP + 1 mM $MgCl_2$	Glucose 5mM + 1 mM AMP + 1 mM $MgCl_2$	Glucose 10 mM + 1 mM AMP + 1 mM $MgCl_2$	Glucose 20 mM + 1 mM AMP + 1 mM $MgCl_2$	Glucose 30 Mm + 1 mM AMP + 1 mM $MgCl_2$	Glucose 40 mM + 1 mM AMP + 1 mM $MgCl_2$
H	Proteína batch 1 + tampão + Dye	Proteína batch 1 + tampão + Dye	Proteína batch 1 + Tris + Dye	Proteína batch 1 + Tris + Dye	Proteína batch 2 + tampão + Dye	Proteína batch 2 + tampão + Dye	Proteína batch 2 + Tris + Dye	Proteína batch 2 + Tris + Dye	Tampão + Dye	Tampão + Dye	Tampão + proteína	Tampão + proteína

7.4. Appendix 4 – Activity assay

Solutions: All reusable plastic and glassware material were rinsed with 1 N HCl to remove any phosphate residue

A - Malachite green solution (0.045% [w/v])

B - Ammonium molybdate (4.2% [w/v]) in 4 N HCl

C - Sodium citrate (34% [w/v])

D - 2% (w/v) Nonidet P-40

E - Color reagent (MG + AM) – Mix 1 volume of ammonium molybdate solution (B) with 3 volumes of malachite green solution (A). Stir for 20 min at room temperature until clear and add 0.1 mL of "D"/5 ml of mix, stir, and filter.

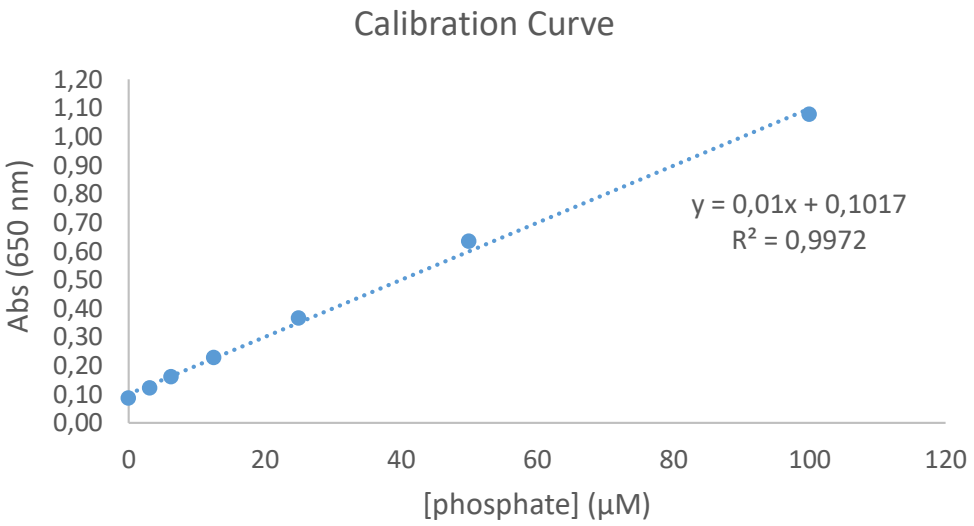


Figure 7.4.1 – Phosphate calibration curve in buffer containing 50 mM HEPES pH 7.5, 100 mM KCl, 2.5 mM EDTA, 2.5 mM $MgCl_2$ and 0.5 mM AMP. The phosphate concentrations used were 100 µM, 50 µM, 25 µM, 12.5 µM, 6.25 µM, 3.12 µM and 0 µM.

7.5. Appendix 5 – Article Order Confirmation

ELSEVIER

Rights and Access

Combined In Silico and In Vitro Studies to Identify Novel Antidiabetic Flavonoids Targeting Glycogen Phosphorylase

Corresponding author	Dr Natercia F. Bras
E-mail address	nbras@fc.up.pt
Journal	Bioorganic Chemistry
Article number	104552
Our reference	YBIOO_104552
PII	S0045-2068(20)31850-2

Order Confirmation

Thank you for taking the time to complete the Rights and Access form.

Order number	OACSRYBIOO1045520
Order date	11 December 2020
Research Funders	Fundação para a Ciência e Tecnologia Grant numbers: CEECIND/02017/2018,UIDB/50006/2020,UIDB/04378/2020,NORTE-01-0145-FEDER-000024,PTDC/QUIQFI/28714/2017
Publishing Option	Subscription
Publishing Agreement	• I am one author signing on behalf of all co-authors of the manuscript I may post the accepted manuscript in my institutional repository and make this public after an embargo period. To ensure the sustainability of peer-reviewed research in journal publications, I may not share the final article publicly, for example on ResearchGate or Academia.edu. Further details on Elsevier Sharing Policy here. Based on information provided the embargo period/end date is 24 months.