Computer-aided design & synthesis of 5,7-dihydroxyflavone derivatives as glycogen phosphorylase inhibitors

Ben A. Chetter. BSc(Hons)

In partial fulfilment for the requirements for the degree of Masters by research at the University of Central Lancashire

September 2016
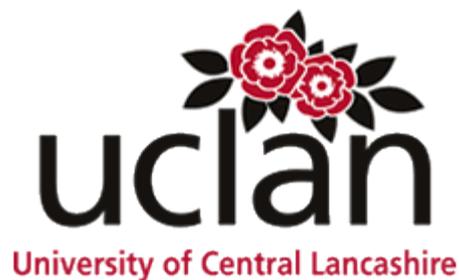
Concurrent registration for two or more academic awards

Either

 *I declare that while registered as a candidate for the research degree, I have not been a registered candidate or enrolled student for another award of the University or other academic or professional institution

Or

~~*I declare that while registered for the research degree, I was with the University's specific permission, a *registered candidate/*enrolled student for the following award:~~

_____

Material submitted for another award

Either

*I declare that no material contained in the thesis has been used in any other submission for an academic award and is solely my own work

Or

~~*I declare that the following material contained in the thesis formed part of a submission for the award of~~

_____

delete as appropriate

Collaboration

Where a candidate's research programme is part of a collaborative project, the thesis must indicate in addition clearly the candidate's individual contribution and the extent of the collaboration. Please state below:

Signature of Candidate _____

Type of Award   **Masters By Reseach_____**

School   **School of Physical and Computational Sciences_____**

Abstract

Diabetes is a chronic disease that affects approximately 415 million people worldwide and this number is expected to rise to 642 million by the year 2040. 90% of cases of diabetes are type 2. Although some treatments are already available for type 2 diabetes, there are considerable side-effects associated with these drugs that include a risk of hypoglycaemia. Hence, there is an immediate need for new and more effective drugs. Naturally occurring flavonoids are known inhibitors of glycogen phosphorylase, which is a validated target for controlling hyperglycemia in type 2 diabetes. By exploiting computational methods such as molecular docking with Glide and post-docking binding free energy calculations using advanced QM/MM-PBSA calculations, we have screened a large library of 1239 5,7-dihydroxyflavone analogues which have the potential to bind at the inhibitor site of glycogen phosphorylase. We compared these results to our benchmark ligand chrysin, which can inhibit glycogen phosphorylase with a K_i of 19.01 µm and identified thirteen ligands which are predicted to have better binding affinities exploiting a rigorous consensus scoring approach to reduce the chance of false positives. Favourable substituents on the predicted flavonoids included B ring hydroxyl and halogenated substituents in the ortho position and small hydrophobic groups on the meta/para positions directed towards the hydrophobic cavity of the binding site. Towards synthesis of these analogues, after many adjustments to the synthetic procedure it was identified that using methyl protection of the hydroxyls at the 5 and 7 position in parallel with the Baker-Venkataraman re-arrangement managed to achieve high purity initial compounds with a relatively good yield. The procedure is now ready to be employed to synthesize all other predicted 5,7-dihydroxyflavone derivatives. Kinetics experiments will validate their potency against glycogen phosphorylase and structure activity relationship analysis can help further guide lead optimization alongside pharmacokinetics and *in vivo* studies, until we can achieve a more potent and drug-like inhibitor of glycogen phosphorylase for the treatment of type 2 diabetes.

Contents

List of Figures

Figures

oxidation *with* versus *without* flavonoid present (β-carotene and linoleic acid system). The scavenging of free radicals using DPPH assay (% AR activity) was also measured versus a reference *without* flavonoid as described in the original reference

List of Tables

List of Schemes

Acknowledgements

I would like to thank my family for their support throughout the year and encouraging me to undergo this project.

A massive thank you to Joe Hayes and Tim Snape for supervising, guiding and occasionally kicking me through this project, I know it hasn't been easy for all three of us, but we made it… Just about…

Thank you to my friends who kept me alive and sane, even when I was losing my mind. And a special thanks to Zoe who put up with me while I slowly sunk into madness, her encouragement, cups of tea and food nights is what kept me going through the hard times. Also to Emma and Adam, who sheltered me and gave me a bed during the last few months. Their food, cups of tea and Netflix nights are all greatly appreciated!

A big hand to my friends in the office, (and in no particular order) Reece, Diego, Dan, Pippa, Laura and Jamie have kept me going, and all been very polite as I shouted at the computer and cried into my mug. And to Chris who helped me tremendously in labs.

Abbreviations

AO	Anti-oxidant
AR	Anti-radical
DFT	density functional theory
DPPH	1,1-diphenyl-2-picrylhydrazyl
ECG	Epicatechin 3-gallete
EG	Epicatechin
EGC	Epigallocatechin
EGCG	Epigallocatechin 3-gallete
G-1-P	α-D-glucose-1-phosphate
G-6-P	Glucose-6-phosphate
GCA	Gradient-corrected approximation
GCG	Gallocatechin gallete
GP	Glycogen Phosphorylase
GS	GlideScore
LBDD	Ligand based drug design
LDA	Local density approximation
LDL	Low density lipoproteins
MM	Molecular Mechanics
MO	Molecular orbital
MS	Mass Spectrometry
NAG	N-Acetyl- β-D-glucopyranosylamine
NCI	National Cancer Institute
NMR	Nuclear magnetic resonance
PBSA	Poisson-Boltzmann Solvation Energy
PDB	Protein Data Bank
PLP	Pyridoxal 5'-phosphate
QM	Quantum Mechanics
QSAR	quantitative structure activity relationship
RMSD	Root-mean-square-deviation
RNS	Reactive nitrogen species
ROS	Reactive oxygen species
RYGB	Roux-en-Y gastric bypass
SBDD	Structure based drug design

SCF	Self-consistent field
SP	Standard Precision
SUR-1	Sulphonylurea receptor
SWE	Schrödinger Wave equation
T1D	Type 1 diabetes
T2D	Type 2 diabetes
USDA	United States Department of Agriculture
vdW	Van der Waals
XP	Extra precision

Chapter 1 – Introduction

1.1 – Introduction to Diabetes

1.1.1 – Prevalence of Diabetes

Diabetes has become a worldwide problem, which affects both high and low income countries; the International Diabetes Federation claims diabetes affected 382 million worldwide in 2013, and by 2035 for that number to rise to 592 million [1]. Diabetes is also a massive strain on financial resources, with the estimated global cost (direct and indirect) being 245 billion U.S. dollars in 2012 alone [2], which will rise due to more people developing the condition and inflation.
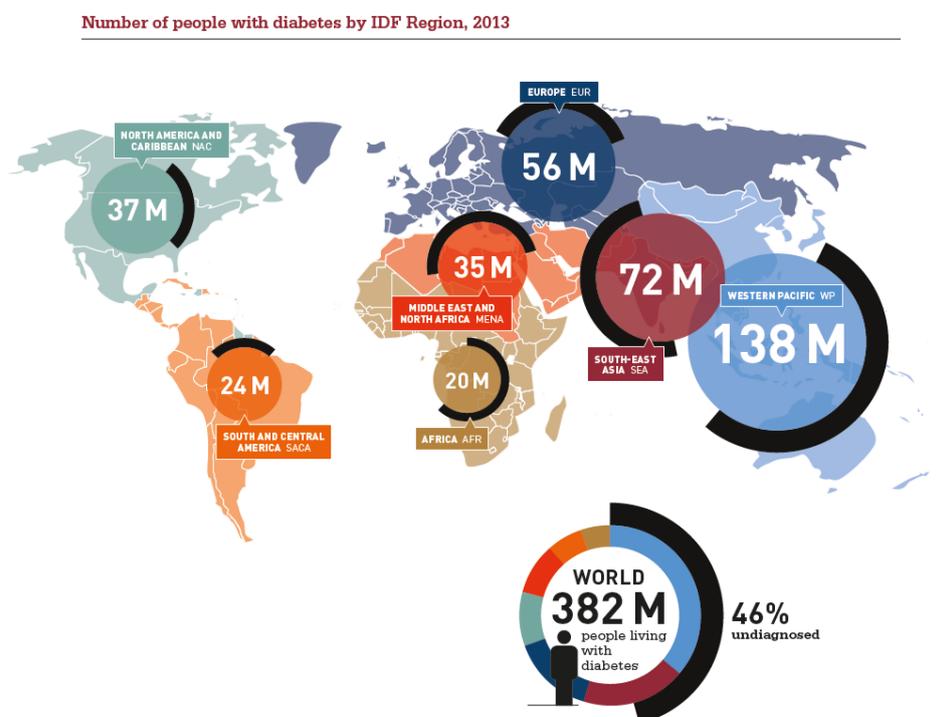
Figure 1.1 - Map showing the number of people worldwide with diabetes and how many people are affected in different regions (taken from [3]).

1.1.2 –Type 2 Diabetes

Type-2 diabetes (T2D) or non-insulin-dependent diabetes mellitus is a heterogeneous disease characterized by hyperglycaemia. It is estimated that 21 million people in the U.S. have been diagnosed with 8.1 million undiagnosed. 90-95% of people diagnosed with diabetes have T2D, and are generally middle aged or older.

T2D is characterised by three factors: 1) a resistance of insulin action on glucose uptake in tissues such as skeletal muscle, adipocytes and liver tissue, 2) relatively low production of insulin from the pancreas, and 3) impaired insulin action to inhibit the hepatic glucose production [4]. The insulin is a hormone which is created in the β-cells in the islets of Langerhans in the pancreas. Insulin has two major roles in the body: maintaining blood glucose levels and controlling glucose metabolism. It is important to regulate blood glucose levels in the blood, as too much glucose which is left to circulate in the blood (hyperglycaemia) can cause tissue damage over time [3], too little glucose (hypoglycaemia) can lead to tiredness, shaking and loss of consciousness [5]. Glucose can be regulated by insulin as insulin stimulates the translocation of the glucose transporter to the cell surface [6]. Glucose levels can also be regulated by decreasing glucose productivity, which can be achieved by decreasing glycogenolysis and gluconeogenesis. Both of these pathways produce glucose and are inhibited by insulin, however the key difference is that glycogenolysis produces glucose by breaking down glycogen, whereas gluconeogenesis forms glucose from pyruvate and other non-carbohydrate substances.

Chances of developing T2D can vary depending on several risk factors, including: obesity, poor diet, physical inactivity, age (>45 years), ethnicity, family history of diabetes and high blood glucose during pregnancy affecting the unborn child [3].

1.1.3 – Type 1 Diabetes

Type 1 diabetes (T1D) or diabetes mellitus, is an autoimmune condition which makes the body insulin-dependent. It's a metabolic disorder of multiple aetiology characterized by chronic hyperglycaemia with disturbances of carbohydrate, fat and protein metabolism resulting from defects in insulin secretion [7]. These defects occur as the immune system attacks the β cells in the pancreas, which stop them producing the insulin. T1D has many differences to Type 2 Diabetes, as shown in Table 1.1.

Table 1.1 – Differences between Type 1 and Type 2 diabetes [8]

Type 1 Diabetes	Type 2 Diabetes
Often diagnosed in childhood	Usually diagnosed in adults over 30
Not associated with obesity	Usually associated with high levels of obesity
Associated with above average ketone levels at diagnosis	Associated with high blood pressure and/or cholesterol levels at diagnosis
Treated with insulin injections or pumps	Usually initially treated without medication or with tablets
Cannot be controlled without taking insulin	Sometimes possible to come off diabetes medication

1.1.4 – Gestational Diabetes

Gestational diabetes is the third main type of diabetes which affects roughly 18% of pregnant women, though it can be diagnosed during the first trimester, this type of diabetes doesn't usually take effect until the second or third trimester [9]. The problem arises when the woman's body produces more insulin than normal, increasing glucose levels in the blood, which insulin cannot bring under control. Symptoms of gestational diabetes are similar to T2D, with increased thirst, urination, tiredness, blurred vision and recurrent infections. This form of diabetes can usually be controlled with a healthy diet and regular exercise, however, some women require medication to keep their blood glucose levels under control [9]. This form of diabetes usually goes away after pregnancy, however, studies have shown that women who suffer from gestational diabetes are more likely to develop T2D in the future [10].

1.2 – Signs & Symptoms

Many people are undiagnosed with Type 2 diabetes, and can remain undiagnosed for many years; this may be due to how mild early symptoms are, which include symptoms such as being thirsty, increased appetite, increased urination, balantis [11], vomiting or stomach pains, lack of concentration or interest and unexplained weight loss. These symptoms can be very gradual taking a few years in some cases. However, when T2D develops symptoms become more prominent, which include: increased urination, feeling thirsty, fatigue, weight loss, blurry vision, itching, slow healing cuts & bruises, leg & feet ulcers, recurrent infections and loss of feeling or tingling in the feet [12].

Figure 1.2 – Symptoms of diabetes [13].

1.3 – Health Conditions Exacerbated by Type 2 Diabetes

People who suffer from diabetes are also at an increased risk of other avoidable medical complications such as angina, cardiac failure, strokes, retinopathy, renal replacement therapy and amputations (>100 amputations a week in the UK) [14].

Between 2007 and 2012, avoidable complications increased significantly

67% ANGINA

130% CARDIAC FAILURE

106% STROKE

33% RETINOPATHY

95% RENAL REPLACEMENT THERAPY*

60% AMPUTATIONS
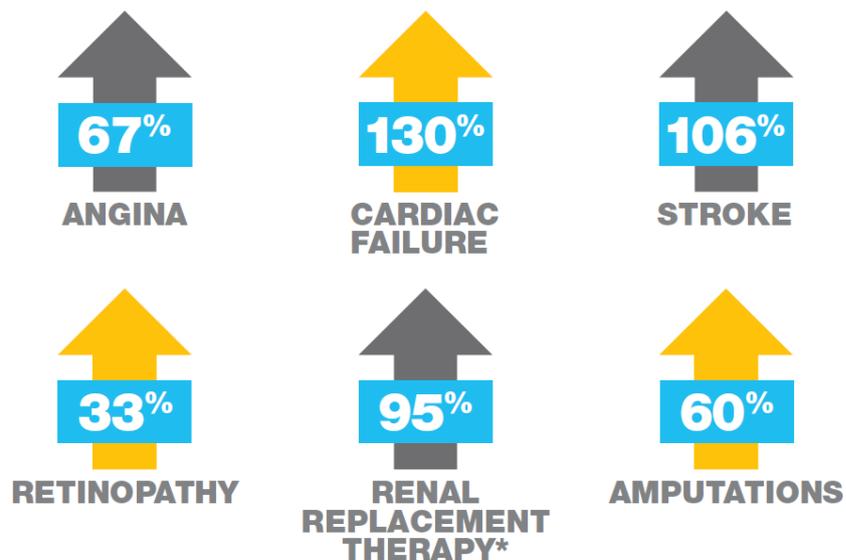
*Term used for life-supporting treatments required to treat end stage kidney disease

Figure 1.3 – Increase in avoidable complications between 2007 and 2012 [14].

These generally occur after several years of having T2D, and occurs due to tissue complications, such as in small blood vessels (microangiopathy). The frequency of arterial diseas such as atherosclerosis and macroangiopathy is increased in diabetics, leading to problems associated with coronary heart disease, strokes and peripheral vascular disease [15]. People with T2D are 32.0 per cent more likely to die earlier than the general population. It is estimated that there were around 22,060 in England and 1,926 additional deaths in Wales due to diabetes, probably due to the increased prevalence of the complications mentioned in **Fig 1.3 [16]**.

1.4 – Current Treatment

As it stands there is no known cure for type 2 diabetes. Also, given the nature of the condition, it is unlikely that a cure is possible any time soon, but instead effective lifelong treatment is currently needed. To avoid early symptoms getting worse and to reduce the possibility of developing extra health problems later in life, an early diagnoses is vital. For many people diagnosed with T2D, the first step to take is changes in the patient's lifestyle, as overweight or obese people account for 80-85% of the risk of developing T2D [14]. So altering people's lifestyle is the foundation of effective treatment. This includes a change of diet and exercise. For example with a 7% weight loss and at least 700 calories/week of physical activity, there is a 58% reduction in the incidence rate of diabetes [17].

If lifestyle changes are impractical for the patient or they are unable to achieve and maintain any significant weight loss, then surgical intervention is a current treatment option. This involves surgically helping the patient to lose weight via procedures such as gastric bands and Roux-en-Y gastric bypasses (RYGB). The RYGB method usually results in a higher sustained weight loss of roughly 40-50%, while gastric bands only work 20-30% of the time, this may be due to the patients' mentality and after care rather than the actual surgery. With RYGB surgery, patients with diabetes prior to surgery, resolution of the disease is reported in 65-90% of cases. Combined with the weight loss associated with post operation effects (decreased nutrient intake and changes in gut hormones), this can improve insulin resistance and β-cell function [18]. After RYGB surgery patients showed a significant decrease in blood glucose (P=0.005) and insulin levels (P=0.02) showing a clear effect on hormonal changes related to T2D [19]. This suggests that the control of diabetes may be a direct effect of the RYGB and not a secondary outcome [20], leading to potentially using this surgery as an early method to combat the onset of diabetes.

1.4.1 - Oral Hypoglycaemic Agents

As T2D is a progressive disease, eventually lifestyle changes alone will not be enough to control T2D. One source of treatment is oral hypoglycaemic agents such as metformin from the biguanide family, which acutely decreases hepatic glucose production. This is achieved by a mild and transient inhibition of the mitochondrial respiratory-chain complex 1 [21]. Metformin has also shown beneficial side effects by reducing LDL (Low Density Lipoprotein) cholesterol and triglyceride levels, these

reductions help cardiovascular problems associated with diabetes. There are some negative side effects, the most common is gastrointestinal upset, and more serious side effects include anorexia or diarrhoea. When taken in excess or by people with contraindications such as renal impairment, lactic acidosis (lactate build up in the blood) can be a problem [11]. This drug should also not be administered to people with hepatic, cardiac or respiratory failure or people with a history of alcohol abuse. Overall though, when prescribed properly, the overall risk of the drug is low and has been effectively used for decades, this is one reason it is generally one of the first drugs to be administered to the patient [18]. In principle metformin can adopt roughly 28 forms, depending on environment and protonation.
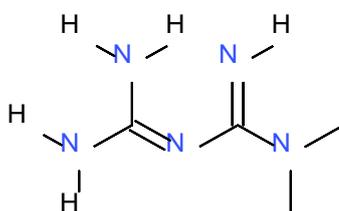
Figure 1.4 – Structure of Metformin [22].

1.4.2 - Sulphonylureas
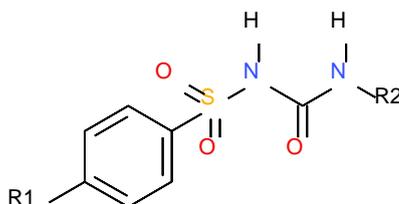
Figure 1.5 – General structure of sulphonylureas.

Sulphonylureas are generally second line treatments for patients after metformin alone [11]. The general scaffold is shown in Figure 1.5 and the different types currently marketed in Table 1.2. Sulphonylureas stimulate insulin secretion, which is achieved by binding to the sulphonylurea receptor (SUR-1) on the β-cell plasma membrane which closes the K_{ATP} channel. This causes the calcium channels, leading to an influx of calcium and exocytosis of insulingranules. A serious side effect to sulphonylureas is hypoglycaemia, particularly with glibenclamide. Side effects and effectiveness also dependant on a few other risk factors, such as alcohol intake, age, impaired renal

function and intake of drugs that displace sulphonylureas from their plasma protein binding sites. There is also the risk of weight gain [15].

Table 1.2 – Different sulphonylureas and their structures.

Name	Structure. [23] [24] [25]
Glibenclamide	
Gliclazide	
Glimepiride	
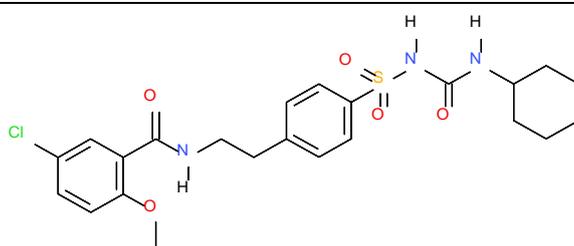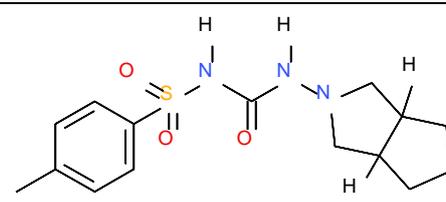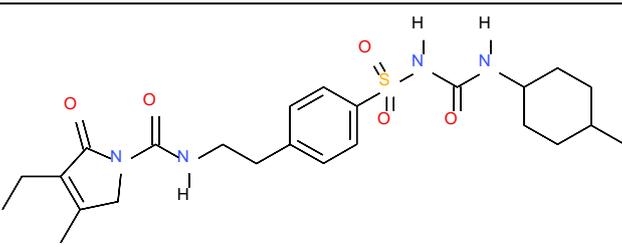
1.4.3 - Meglitinide Analogues
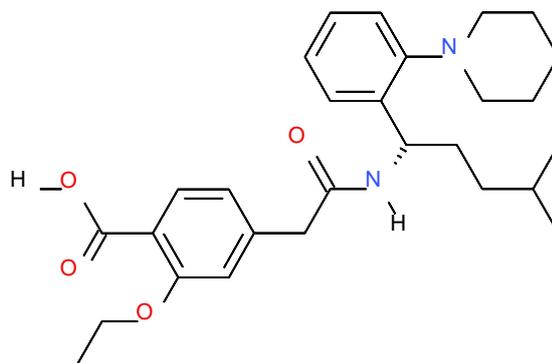
Figure 1.6 – Structure of Repaglinide [26]

Meglitinide analogues are very similar to sulphonylureas, as they also stimulate insulin secretion by inhibiting the same ATP-sensitive channels (SUR-1), however, the pharmacokinetic and pharmacodynamics properties are different to the sulphonylureas.

Meglitinide analogues such as repaglinide (Figure 1.6) are absorbed rapidly, stimulating the release of insulin in a matter of minutes, then being rapidly metabolised for excretion [27]. Side effects include greater weight loss (up to three kg in three months) and hypoglycaemia than metformin, but less diarrhoea [28].

1.4.4 - Insulin injections

In the most severe cases of T2D, most patients will be advised to take a drug such as metformin along with the insulin injections. However, for some people this is not a viable option and they must take insulin alone [29]. This should always be combined with the lifestyle changes (diet & exercise) as much as possible, to increase effectiveness. There are multiple ways of administering insulin: Insulin administered with a syringe, insulin pens, insulin pumps and inhaled insulin. If administered properly in the correct areas, then insulin injections rarely cause problems, aside bruises [11]. In some cases lipoatrophy can occur, when injecting into the same area for long periods of time. This is less of a risk recently, as modern insulin is highly purified and causes a much smaller immune response.[15]

1.5 - Synopsis

Type 2 diabetes mellitus is a worldwide problem, accounting for 90% of all diabetes cases. There is a demand for more effective and safer drugs for combating such a widespread and problematic illness. This thesis will investigate glycogen phosphorylase (GP), as a target for T2D, and flavonoid analogue inhibitors of the caffeine binding site in particular. Flavanoids have already shown promise as GP inhibitors. Liver GP is an important target in T2D, as this is the enzyme that catalyses the breakdown of glycogen to glucose, so inhibition will reduce the amount of glucose production from the liver released into the blood.

This project aims to design and synthesise analogues of 5,7-dihydroxyflavone derivatives which can inhibit GP more effectively than chrysin. This can be achieved by exploiting docking and post-docking computational methods, then, using these predictions to select candidates for synthesis. A synthetic protocol was optimised for these syntheses which ultimately led to the Baker-Venkataraman re-arrangement. The final products are to undergo kinetics experiments at our collaborators (University of

Thessaly, Greece) to determine their true potencies and X-ray crystallography to confirm their binding interactions with GP. This data will direct for further lead optimisations efforts using structure activity relationship methods, with the ultimate aim of designing a novel, safe and effective compound which can be used in the treatment of T2D.

-Chapter 1 – A general introduction to diabetes, with explanations of its global effect, types of diabetes, the different ways it can affect people and some of the current methods used in treatment.

-Chapter 2 – Glycogen Phosphorylase as a target for treatment of type two diabetes, as well as focus on glycogenolysis and how inhibiting this pathway can be an effective way to prevent hyperglycaemia.

-Chapter 3- Flavanoids as potential agents for the inhibition of GP. Flavanoids are naturally occurring, low molecular-weight polyphenolic compounds [30]. These compounds have already shown potential to inhibit GP.

-Chapter 4 – Computational background/theory presented, including a description of docking and in particular the programme Glide, from Schrödinger software, as well as quantum mechanics methods necessary to describe π-stacking effects of flavonoids binding at the GP caffeine binding site.

-Chapter 5 – Synthetic methods and synthetic routes for compounds made, with a particular focus on the Baker-Venkataram rearrangement.

-Chapter 6 – Computational results and discussions.

-Chapter 7- Synthetic results and discussions.

-Chapter 8 – Conclusion and future work.

Chapter 2 - Glycogen Phosphorylase

2.1 – Glycogenolysis and Gluconeogenesis

Glycogenolysis and gluconeogenesis are both ways in which the body produces glucose, which are both inhibited by insulin. However, gluconeogenesis is the production of glucose from the substrate pyruvate and other non-carbohydrate substances (Figure 2.1). Whereas glycogenesis is the production of glucose from the breakdown of glycogen (Figure 2.2), which is a complex carbohydrate. Gluconeogenesis also occurs in the kidneys, not just the liver and muscles [31, 32].
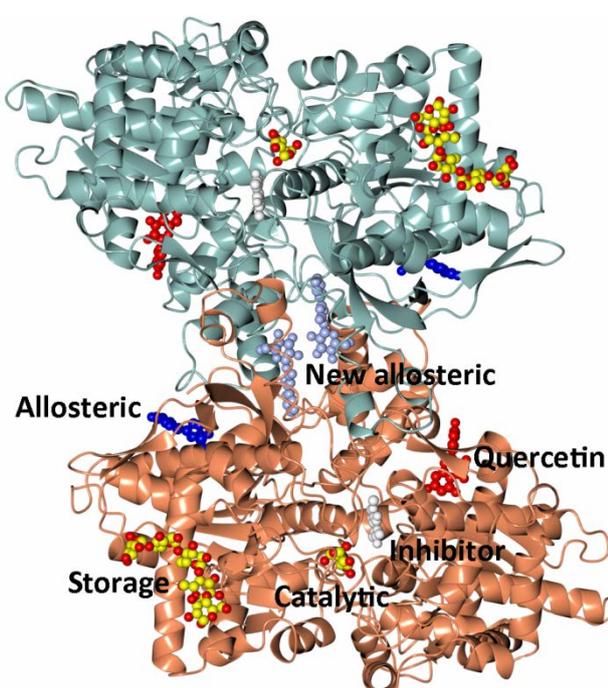
Figure 2.1. The glycogen phosphorylase dimer with the various ligand binding sites

Gluconeogenesis is triggered when a sufficient amount of carbohydrate substrates are not available, for example, when starving or during intense physical exertion. Gluconeogenesis is the reverse process of glycolysis (breakdown of glucose for energy), using similar enzymes. This is made possible by three extra steps in the gluconeogenesis pathway, which are regulated by other enzymes.[32]

2 x Pyruvate

Pyruvate carboxylase — 2 ATP + 2 HCO$_3^-$ / 2 ADP

2 x Oxaloacetate

PEP carboxykinase — 2 GTP / 2 GDP + CO$_2$

2 x Phosphoenolpyruvate

2 x 3-Phosphoglycerate

Phosphoglycerate kinase — 2 ATP / 2 ADP

2 x 1,3-Bisphosphoglycerate

2 NADH + H$^+$ / 2 NAD

Fructose-1,6-bisphosphate

Fructose-1,6-bisphosphatase — Pi

Fructose-6-phosphate

Glucose-6-phosphate

Glucose-6-phosphatase — Pi

Glucose

GLYCOGEN

Pi

Debranching enzyme — Glycogen phosphorylase

+ cAMP

GLUCOSE 1-PHOSPHATE

Phosphoglucomutase

Mg^{2+}

GLUCOSE 6-PHOSPHATE

Glucose 6-phosphatase

Pi

GLUCOSE

Figure 2.2 – The process of gluconeogenesis.[33]
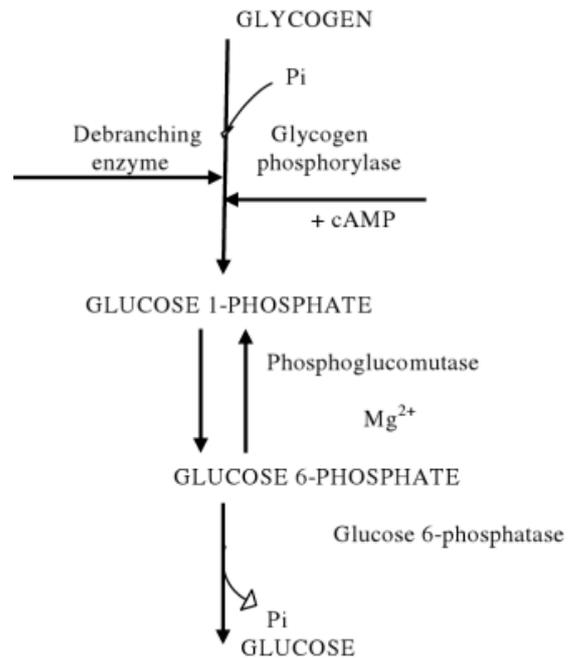
Figure 2.3–The process of glycogenolysis.[33]

2.2 – Glycogen Phosphorylase

Glycogen Phosphorylase is an enzyme involved in glycogen metabolism (glycogenolysis), which catalyses the first step in the intracellular degradation of glycogen to give α-D-glucose-1-phosphate (G-1-P). Because of this, it has been a prominent target for T2D as many antihyperglycaemic agents focus on the inhibition of the GP's inhibitor sites.

2.3 - Structure, Function and Regulation of Glycogen Phosphorylase

GP is the main regulatory enzyme for human carbohydrate metabolism, as it catalyses the breakdown of glycogen to G-1-P, which then undergoes isomerisation and interconversion to glucose-6-phosphate (G-6-P) by the enzyme phosphoglucomutase.[34] This G-6-P can then enter the glycolytic pathway or it can be

dephosphorylated by Glucose-6-phosphorylase to glucose and transported out of the cell. GP exists as a dimer of two interconvertible forms, there is a T-state and an R-state which exist in equilibrium, with a molecular weight of roughly 97500Da (Figure 2.3). The predominantly inactive T-state is the dephosphorylated form, GPb, which has low activity and substrate affinity. Then by either phosphorylation or AMP [35], the GPb is converted to the active R-state 'ser14-phosphorylated' form (GPa) which has a much higher activity and substrate affinity [36]. This reaction is reversed with the protein phosphate 1, which is released in response to insulin, and dephosphorylates the GPa back to GPb. There are also three isoforms of GP termed as "muscle", "brain" or "liver" GP, depending on which tissue the isoform is preferentially expressed, each has a distinct physiological role.

Muscle-GP is usually in it's inactive GPb form. As muscles need a lot of energy, sometimes without a lot of warning, it is important to supply enough glucose to power the muscle when needed. As muscles do not have the glucose-6-phosphorylase enzyme, glucose is formed slightly differently, glycogen breaks down into the G-1-P and undergoes interconversion to G-6-P, as normal, but then the molecule enters glycolysis. GP itself starts in the inactive form, but as the muscle is worked ATP is used so AMP levels increase, this causes the GP to convert from the inactive T state to the active R state. This process generates the metabolic energy needed for muscle contraction.

Brain GP exists at much lower concentrations as muscle and liver, however, it is still important, as small amounts of glycogen is stored in the brain (averaging 3-12 µmol/g), these low concentrations suggest that the GP and glycogen is not a primary energy source, [37] but as an emergency backup supply of glucose during times of anoxia, ischaemia, aglycemia or severe hypoglycaemia.
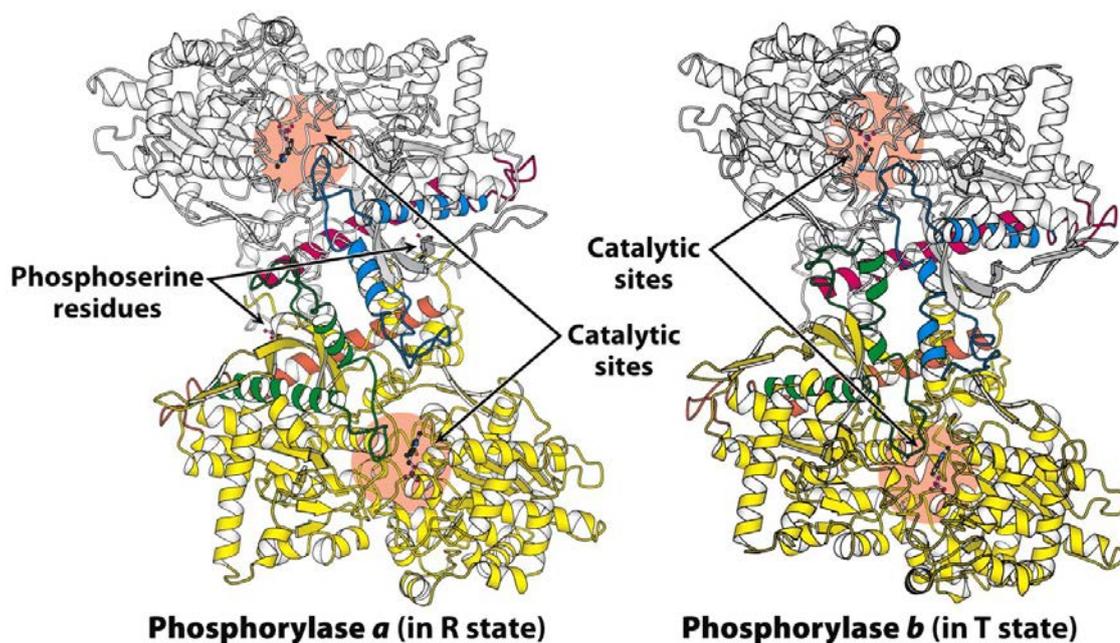
Figure 2.4 - Binding sites and conformational changes of glycogen phosphorylase.

There are currently six sites on the GP enzyme; the catalytic site, caffeine binding site, glycogen storage site, AMP binding site, the indole binding site and the quercetin binding site (Figure 2.1). The structure of GP has been confirmed through multiple extensive X-ray crystallographic studies and information about the protein-inhibitor interactions in each binding site has been well researched [36, 38], perhaps with the exception of quercetin binding site, as this has only been recently discovered. These studies have allowed research into using these sites as targets for compounds, which may be able to manipulate the GP to prevent unwanted glycogenolysis under high glucose concentrations, which gives rise to numerous opportunities for GP manipulation. While the inhibitor site and quercetin binding site will be described in Chapter 3 due to their propensity to bind flavonoids, a brief description of other key binding sites on GP is given here.

2.4 - Catalytic site

The catalytic site is the active site of GP, where the glycogen is phosphorylased to G-1-P. The site contains the essential pyridoxal 5'-phosphate (PLP) as a cofactor, located at the centre of the GP monomer subunit, which is accessible to bulk solvent through a 15 Å long channel. [36] This is where domains 1 (residues 1-484) and 2 (residues 485-842) join. The catalytic site is inhibited by many compounds (150 have been studied in the

last decade) [39] and has been noted that many are glucose based analogues. When the inhibitor binds to the catalytic site, it stabilises the conformation at the 280s loop (residues 282-287) located between helices α-7 (residues 261-274) and α-8 (residues 289-314), which blocks the substrate (glycogen) reaching the site, this promotes the inactive T-state conformation. This also prevents the conformational changes needed to form the functional catalytic site, such as the movement of Arg569 to the catalytic cleft to create the phosphate recognition site [40].

When the enzyme is activated via the conversion from the T-state to the R-state, the 280s loop becomes disordered and displaced, which creates the recognition site for the substrate phosphate, along with the opening of the 15 Å channel, this in turn allows the residue Arg569 to enter the catalytic cleft in place of Asp283, creating the phosphate recognition site mentioned earlier. The creation of this site allows the glycogen substrate to enter the catalytic site and promote a favourable electrostatic environment for PLP, due to the substrate phosphate site being within hydrogen bonding distance of the 5'-phosphate group of PLP [41].

Whilst finding treatments for human liver GP would be ideal, most studies are carried out using rabbit muscle GPb (RMGPb), this is due to its availability and its ability to form good quality crystals. HLGPb and RMGPb share the same primary sequence as both the human muscle and liver phosphorylase (80%). AMP binds to both HLGPb and RMGPb and the residues that contact AMP are the same. There are no changes in the HLGPb sequence compared to RMGPb and out of the 171 amino acid difference between the two enzyme, 85 of them (29%) are conservative as well as the fact that most of the amino acid changes have no impact on the enzymes function. The active sites of HLGPb and RMGPb are also identical in both amino acid sequence and structural architecture, which means conclusions for the active site, drawn by binding studies with RMGPb, are applicable to HLGPb [42].

Drug design for inhibitors of the catalytic site have been generally focused around glucose based analogues (Figure 2.4) with α- and β- substitutions on the anomeric C1 atom.
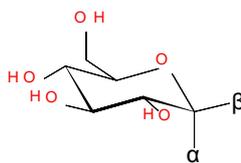
Figure 2.5 – A glucose analogue, displaying
α- positioning and β- positioning.

Figure 2.5 shows the binding of glucose based analogue bound at the GPb catalytic site. The α- and β- substitutions have an effect on binding in the catalytic site, as this orientation can make binding stronger or weaker depending on sterics and intermolecular binding, for example, the α-D-glucose (**1** in **Fig 2.7**) inhibits the GP at the catalytic site with a K_I value of 1.7 µm for GPb inhibtion in rabbit muscle (RM). Whereas the β-D-glucose inhibitor (**2** in **Fig 2.7**) binds to RMGPb with a K_i value of 7.4 µm. The use of β- substituted glucose exploits the catalytic sub-site called the β-cavity, which is an empty space at the β-1-C configuration lined by both polar and non-polar groups. Due to the location of the catalytic site, the GP monomer structure is viable for modelling inhibitor binding at this catalytic site. [43]

An important interaction used in inhibitor design, for the GPI catalytic site, is hydrogen bonding; hydrogen from the NH of ligands forms the hydrogen bond to the His377 via the main chain oxygen. N-Acetyl- β-D-glucopyranosylamine (NAG) (**5** in **Fig 2.7**) also binds to the His377 oxygen, this inhibits RMGPb with a K_i of 32µm, which is roughly 20 times better than α-D-glucose. NAG also binds to the catalytic site with very little structural conformational changes within the catalytic site and residues of the 280s loop are stabilised in the same conformation observed in the RMGPb - α-D-glucose complex.
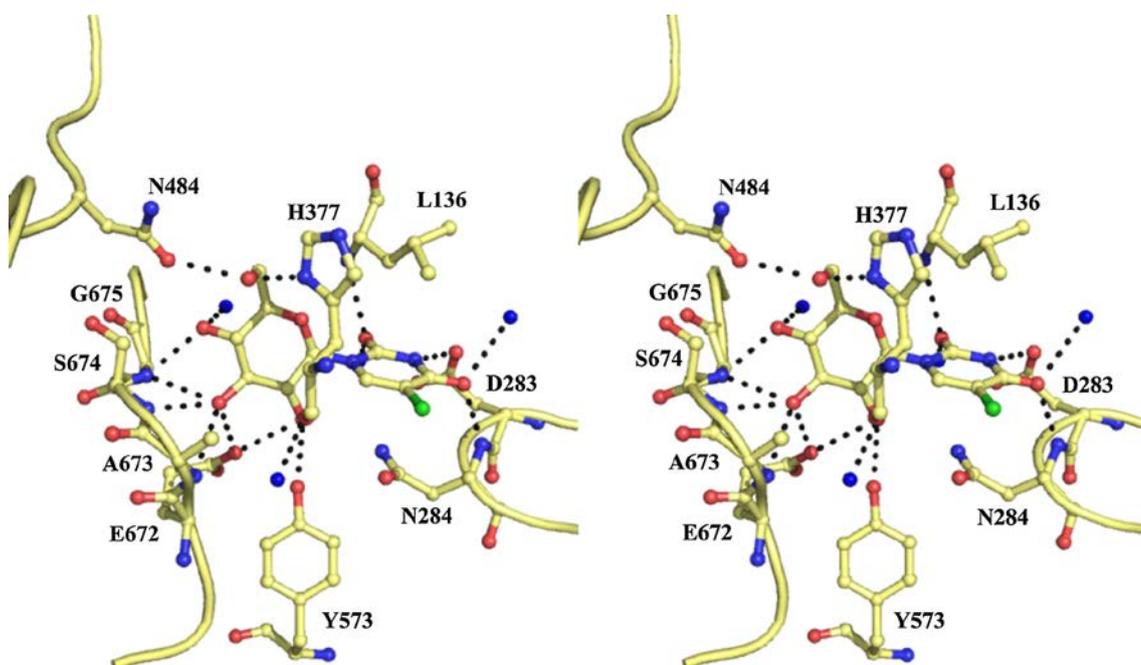
Figure 2.6 – 1-(β-D-glucopyranosyl)-5-chlorouracil bound to the catalytic site or RMGPb. Hydrogen bonds represented in black.
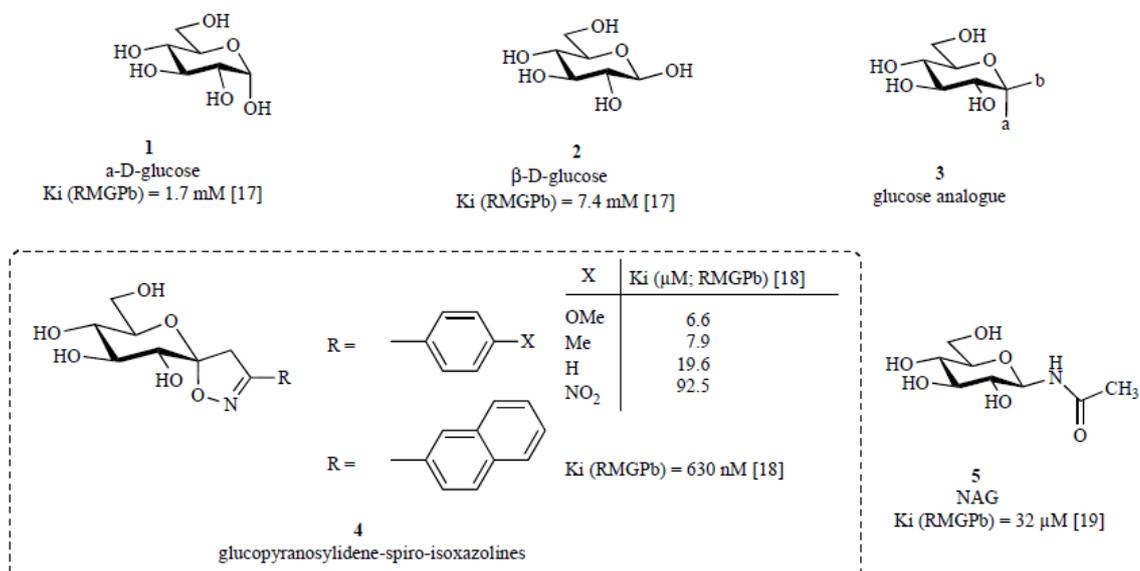
Figure 2.7 – Examples of glucose-based analogue inhibitors together, with their K_i values.

2.5 - Allosteric Site

The allosteric site (AMP binding) site has been one of the most promising GP inhibitor binding site with respect to identification of potent, drug like molecules [18]. The site is located at the subunit, which is approximately 30 Å from the catalytic site and is partially exposed to solvent. The allosteric site can recognise a variety of phosphorylated compounds such as ATP, NADH, IMP, G-6-P, UDP-glucose, 2-deoxy-glucose-6-P, β-glycerophosphate and inorganic phosphate. Inhibition of this site works by direct inhibition of AMP binding and/or indirect inhibition of substrate binding via stabilisation of the T or T' state conformation (superscript prime refers to the symmetry related monomer subunit). Compounds that activate the allosteric site, such as ATP and G-6-P, can change the equilibrium between the T and R-state [43]. The first potent and powerful AMP site inhibitors were a series of dihydropyridine dicarboxylic acids, shown in **Table 2.1** along with the potency of the different R groups. [18]

Table 2.1 – Potencies of selected dihydropyridine AMP site inhibitors.
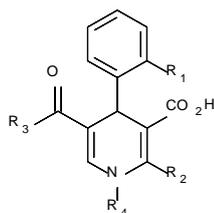
Entry	R_2	R_2	R_3	R_4	$hLGP$(a) K_i (μM)	Cell EC_{50} (μM)
1	Cl	CO_2H	(*i*-Pr)-O	Et	0.039	2.2
2	H	CO_2H	(*i*-Pr)-O	Et	0.395	6.5
3	Cl	H	(*i*-Pr)-O	Et	15.6	NA
4	Cl	CO_2H	(*i*-Pr)-O	H	0.692	11.0
5	Cl	CO_2H	(*i*-Pr)-NH	Et	9.1	NA
6	Cl	CO_2H	(*i*-Pr)-O	CH_2Ph	0.011	1.13
7	Cl	CO_2H	(*i*-Pr)-O	(4-Cl-benzyl)	0.002	0.48
8	Cl	CO_2H	(*i*-Pr)-O	(3,4-diOMe-benzyl)	0.004	0.27

2.6 - Indole Site

The indole binding site (also known as the new allosteric site) is located in the central cavity region of the subunit dimerization interface and is found 33 Å away from the catalytic site and is 50 Å from the glycogen storage site. [18] It has been found that when the indole site is inhibited the enzyme undergoes conformational changes induced by the binding of indole site inhibitors, which prevents activity. [18]

Inhibitors of this site have been proven to be very powerful at lowering blood glucose in animal models of diabetes.

Chapter 3 – Flavonoids as Glycogen Phosphorylase Inhibitors

3.1 - Introduction

Nutraceuticals are products derived from natural sources which provide health benefits, along with their basic nutritional value found in foods. An area that can benefit tremendously from the use of nutraceuticals is T2D. T2D, which has been labelled as the epidemic of the 21st century, is one of the fastest growing chronic diseases globally and the fourth leading cause of mortality [44]. At present, preventative and therapeutic drug strategies for T2D do not achieve adequate control of blood glucose because they have limited efficacy, major mechanism-based side effects, lose their effectiveness over time, are not well-tolerated in some patients or may have poor risk-benefit ratios [45]. To prevent T2D, it is currently recommended to have a generous intake of plant products and reduced intake of fat from beef, pork and dairy sources, while 57% of diabetics pursue complementary and alternative therapies that assist in managing T2D. While more than 1200 plants and herbal extracts appear to affect blood glucose regulation, scientific evidence to support their efficacy against T2D is limited [45]. Research on novel antidiabetics to date has been dominated by studies on newly synthesised compounds and only recently have phytogenic compounds started to attract interest in T2D research. Phytogenic compounds have displayed a very promising capability to inhibit glycogen metabolism enzymes which make them promising targets for the treatment of T2D [42]. Despite the fact that plants have been the single most productive source of bioactive material for various nutraceutical applications, it is estimated that only 5 to 15 % of the 250,000 to 750,000 existing species of higher plants have been tested for bioactive compounds. Thus, the vast majority of the world's structurally diverse natural compounds have not yet been tested for biological activity and therefore there is considerable potential for the discovery of novel bioactive compounds against T2D. In comparison to current pharmaceutical approaches, the study of plant-based remedies modulating physiological effects can provide natural effective oral antihyperglycaemic agents that potentially have little or no side effects for T2D patients.

T2D is characterized by an excess of glucose production from the liver. One molecular approach aims to reduce this production, and involves inhibition of glycogenolysis as this process accounts for more than 70% of the glucose production, whilst most of the glucose formed from gluconeogenesis (*de novo* synthesis of glucose) is also cycled

through the glycogen pool [46, 47]. As described in Chapter 2, GP is an important enzyme in blood glucose homeostasis, catalysing the breakdown of glycogen to glucose-1-phosphate (Glc-1-P) which leads to the release of monomeric glucose from glycogen deposits [48]. GP is a validated target for the discovery of antihyperglycaemic agents [49, 50] since GP inhibitors can affect blood glucose levels and hepatic glycogen balance, as demonstrated by *in vitro* cell biology experiments [51] and animal studies in rodent models of diabetes [52-54].

3.2. - Flavonoids

Flavonoids are polyphenols and ubiquitous components in food of plant origin. Studies in diabetic and non-diabetic rats have shown that they do have anti-hyperglycemic properties due to their effect on glycogen metabolism [55-57]. They are consumed daily as part of a healthy diet of fruit and vegetables and are recognised for many of their health benefits including antioxidant, antibacterial, anticancer, anti-inflammatory and antidiabetic activities, along with protection against cardiovascular disease [58-61]. Data such as the United States Department of Agriculture (USDA) database for flavonoids facilitates the assessment of daily flavonoid intake [62-64]. However, the amount of flavonoids present in food is influenced by a number of factors and different populations have different diets. As an example, based on an analysis of the population's commonly consumed foods, such as vegetables, fruits and beverages [65], the consumption of flavonoids in the USA has been reported as 20 mg/day, which is significantly different compared to Holland, which is > 70 mg/day. A study of 100 edible fruits and vegetables consumed in Harbin city, China, revealed the predominant flavonoids found in all of the foods that were analysed was kaempferol, followed by luteolin and quercetin [66]. Red wine and teas are also rich in flavonoids, and extracts from these sources have been explored for their GP inhibitory potential (*vide infra*) [67-69].

Flavone   Flavonol   Flavanone

Flavanol   Isoflavone   Anthocyanidin
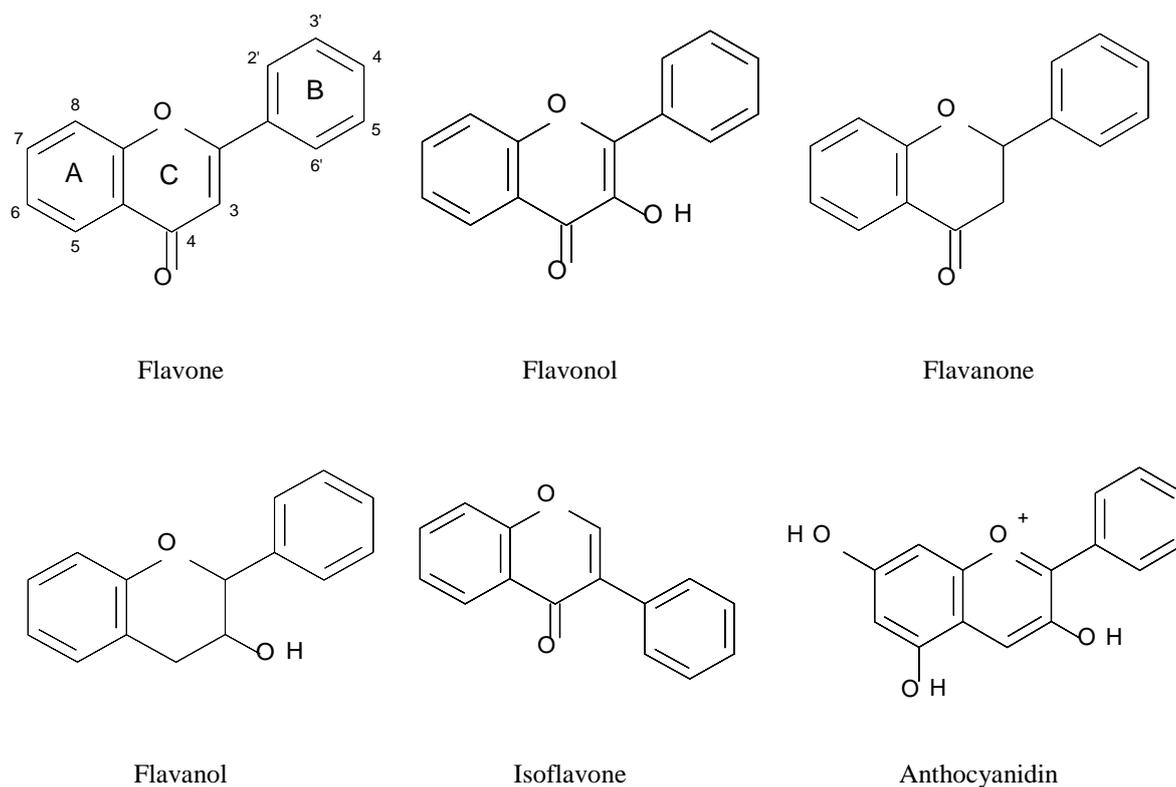
Figure 3.1 - Structural classification of flavonoids

Structurally, flavonoids have a generic scaffold which consist of two aromatic rings (A and B), joined by three carbons which are usually in a heterocyclic ring C containing oxygen. It is the differences on the heterocyclic ring C that classifies flavonoids as either flavones, flavanols, flavonols, flavanones, isoflavonoids and anthocyanidins (Fig. 3.1). The substitution patterns for different flavonoids of each type relevant to this discussion are shown in Table 3.1. Within the subgroups of flavonols and flavones, the flavonol quercetin is the most commonly found in foods. Isorhamnetin, kaempferol and the flavones apigenin and luteolin are also common. Thousands of structurally unique flavonoids have been found and databases containing their structural, biological and physicochemical properties are available from sources such as the National Cancer Institute (NCI) database (https://cactus.nci.nih.gov/ncidb2.2/) and the ZINC docking database (www.zinc.docking.org) [70], while the flavonoid content in selected foods are available from the United States Department of Agriculture (USDA) flavonoid database.

3.3 – Glycogen Phosphorylase Inhibition by Flavonoids

In recent years, natural flavonoid derivatives from the structural classes presented in Figure 3.1 have been investigated as GP inhibitors [68, 69, 71-75]. The results of these mainly IC_{50} binding assay experiments against RMGPa and/or RMGPb along with some K_i values from kinetics experiments are shown in Table 3.2.

The leading flavonoid inhibitors of GP are on the low µM range. Jakobs et al. found in their study that almost all flavonoids tested inhibited phosphorylated active RMGPa and unphosphorylated AMP activated GPb, however inhibition of GPa was two to four times stronger than inhibition for RMGPb [71]. The active form of GP is therefore more strongly inhibited which is favourable for the regulation of blood glucose levels in T2D patients. However, inhibition is very sensitive to the substitution patterns of the flavonoids; a number of flavonoids in each structural class demonstrate little to no inhibition, which is believed to be due to the substitutions. The best inhibitors from different studies are the flavonols quercetin (IC_{50}(GPa) = 4.8 µM; IC_{50}(GPb) = 20.9 µM) [71] and quercetagetin (K_i(GPb) = 3.5 µM) [73], flavones baicalein (IC_{50}(GPa) = 11.2 µM; IC_{50}(GPb) = 10.2 µM) [71], hypolaetin (IC_{50}(GPb) = 15.7 µM) and 6-hydroxyluteolin (IC_{50}(GPb) = 11.6 µM) (Kato et al., 2008) [73]. The flavone chrysin seems to be more potent for GPb (IC_{50} = 15.3 µM) than GPa (IC_{50} > 27.5 µM) [71]. Unfortunately, no flavanone or isoflavone showed inhibition based on the few compounds tested. Of the anthocyanidins, cyanidin (IC_{50}(GPa) = 3.0 µM; IC_{50}(GPb) = 9.0 µM) and delphinidin (IC_{50}(GPa) = 3.1 µM; IC_{50}(GPb) = 10.7 µM) were the most potent [71], while for the catechins the best inhibitors of GPa were EGCG (IC_{50}(GPa) = 7.7 µM; IC_{50}(GPb) = 34 µM) and ECG (IC_{50}(GPa) = 12.5 µM; IC_{50}(GPb) $\geq$ 27 µM) [68, 71], while GCG has also exhibited low µM inhibition for GPb (IC_{50} = 34 µM) [68]. Recently, binding of gallic acid as well as its dimer ellagic acid to GPb were studied by crystallography (*vide infra*) and kinetics. Ellagic acid is a potent inhibitor of GPb (IC_{50} = 12.1; K_i =13.4 µM) as determined by two studies [71, 76] and approximately 2-4 times more potent for GPa (IC_{50} = 3.2 µM; K_i =3900 µM) [71]. On the other hand gallic acid demonstrated poor inhibitory potential in comparison [71, 76].

Table 3.1. Substitution patterns of relevant flavonoids. Numbering of substituents according to the basic structure given below
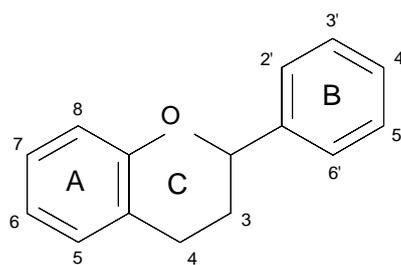
Basic Structure[a]

	3	5	7	3'	4'	+OH	+OCH$_3$
Flavones							
DHF				OH	OH		
Chrysin		OH	OH				
Apigenin		OH	OH		OH		
Luteolin		OH	OH	OH	OH		
Tricin		OH	OH	OCH$_3$	OH		5'
Baicalein		OH	OH			6	
Hypolaetin		OH	OH	OH	OH	8	
6-hydroxyluteolin		OH	OH	OH	OH	6	
Isoscullarein		OH	OH		OH	8	
Flavonols							
Quercetin	OH	OH	OH	OH	OH		
Quercetagetin	OH	OH	OH	OH	OH	6	
(-)-Epicatechin (EC)	OH	OH	OH	OH	OH		
Isorhamnetin	OH	OH	OH	OCH$_3$	OH		
Kaempferol	OH	OH	OH		OH		
Flavanones							
Naringenin		OH	OH		OH		
Hesperitin		OH	OH	OH	OCH$_3$		
Eriodictyol		OH	OH	OH	OH		
Isoflavones							
Daidzein			OH		OH		
Genistein		OH	OH		OH		
Catechins							
(+)-Catechin	OH	OH	OH	OH	OH		
(-)-Epicatechin	OH	OH	OH	OH	OH		
(EGC)[b]	OH	OH	OH	OH	OH	5'	
(ECG) (2S,3S)[c]	Galloyl	OH	OH	OH	OH		
(EGCG) (2S,3S)[d]	Galloyl	OH	OH	OH	OH	5'	
(GCG) (2R,3S)[e]	Galloyl	OH	OH	OH	OH	5'	
(CG) (2R,3S)[f]	Galloyl	OH	OH	OH	OH		
Anthocyanidins							
Pelargonidin	OH	OH	OH		OH		
Cyanidin	OH	OH	OH	OH	OH		
Delphinidin	OH	OH	OH	OH	OH	5'	
Peonidin	OH	OH	OH	OCH$_3$	OH		
Malvidin	OH	OH	OH	OCH$_3$	OH		5'

a) Refer to Fig 3.1. for the different structural classes. b) Epigallocatechin. c) Epicatechin 3-gallate. d) Epigallocatechin 3-gallate. e) Gallocatechin gallate. f) Catechin 3-gallate.

Table 3.2. IC$_{50}$ values (μM) for inhibition of GPa and GPb of flavonoids tested in concentrations up to 50 μM.

Compound	IC$_{50}$ GPa (μM)	IC$_{50}$ GPb (μM)	Compound	IC$_{50}$ GPa (μM)	IC$_{50}$ GPb (μM)
Flavones			**Anthocyanidins**		
DHF	>50 [71]	>50 [71]	Pelargonidin	43.6 [71]	6.2 [71]
Chrysin	>27.5 [71]	15.3 [71]	Cyanidin	3.0 [71]	9.0 [71]
		19.0 [76]	Delphinidin	3.1 [71]	10.7 [71]
		19.01 (K_i) [75]	Peonidin	25.1 [71]	17.6 [71]
Apigenin	>30 [71]	No effect up to 30 [71]	Malvidin	>50 [71]	>50 [71]
Luteolin	15.6 [71]	28.8 [71]	**Catechins**		
		29.7 [73]	Catechin	No effect [71]	No effect [71]
		31.7 [73]	Epicatechin	No effect [71]	No effect [71]
Tricin	>50 [71]	No effect [71]			No effect [73]
Baicalein	11.2 [71]	10.2 [71]	Epigallocatechin	No effect [71]	No effect [71]
Hypolaetin		15.7 [73]			No effect [71]
6-hydroxyluteolin		11.6 [73]	ECG	12.5 [71]	50 [71]
Isoscullarein		46.1 [73]			27 [68]
Flavonols			EGCG	7.7 [71]	33.9 [71]
Quercetin	4.8 [71]	20.9 [71]			34 [68]
		33.5 [73]	EGC		>400 [68]
Quercetagetin		9.7 [73]	GCG		6.3 [68]
		3.5 (K_i) [73]	CG		35 [68]
(-)-Epicatechin (EC)		290 [68]	**Other**		
Flavanones			Ellagic acid	3.2 [71]	12.1 [71]
Naringenin	>50 [71]	No effect [71]		7.5 (K_i) [76]	13.4(K_i) [76]
Hesperitin	No effect [71]	No effect [71]	Gallic acid	No effect [71]	No effect [71]
Eriodictyol	No effect [71]	No effect [71]		3900(K_i)[76]	1730(K_i) [76]
Isoflavones			Caffeine		130 [77]
Daidzein	No effect [71]	No effect [71]			
Genistein	>50 [71]	No effect [71]			

The most favourable flavonoid inhibitors of GP discussed are shown in Figure 3 where the substitution patterns exhibit diversity. While initially it was believed that flavonoids may exert their inhibitory potential through binding at the same GP binding site, it is now clear due to recent work [68, 73, 75] that their mode of action/binding site is extremely sensitive to the structural class and/or substitution patterns. The binding site of most flavonoids is unknown (*c.f.*, for example in Fig. 3.3).

quercetin
$IC_{50}(GPa) = 4.8$ [71]
$IC_{50}GPb = 20.9$ [71], (33.5) [73], (34.8)[f]

Binding site: quercetin

quercetagetin
$IC_{50}(GPb) = 9.7$ [73], (3.5(K_i)) [73]

Binding site: allosteric

hypolaetin
$IC_{50}(GPb) = 15.7$ [73]

Binding site: unknown

6-hydroxyluteolin
$IC_{50}(GPb) = 11.6$ [71]
Binding site: unknown

cyanidin
$IC_{50}(GPa) = 3.0$ [71]
$IC_{50}(GPb) = 9.0$ [71]

Binding site: unknown

delphinidin
$IC_{50}(GPa) = 3.1$ [71]
$IC_{50}(GPb) = 10.7$ [71]

Binding site: unknown

EGCG
$IC_{50}(GPa) = 7.7$ [71]
$IC_{50}(GPb) = 33.9$ [71], (34.0) [68]

Binding site: unknown

GCG
$IC_{50}(GPb) = 6.3$ [68]

Binding site: new allosteric site

ECG
$IC_{50}(GPa) = 12.5$ [71]
$IC_{50}(GPb) = 27.0$ [68], (50) [71]

Binding site: unknown

Ellagic acid
$IC_{50}(GPa) = 3.2$ [71]
$IC_{50}(GPb) = 12.1$ [71], (13.4) [76]
Binding site: inhibitor

Chrysin
$IC_{50}(GPa) = 27.5$ [71]
$IC_{50}(GPb) = 15.3$ [71], (19.01(K_i)) [75]
Binding site: inhibitor

Figure 3.2 - Most potent flavonoid inhibitors of GP known to date.

3.4 - Flavonoid Binding Sites

3.4.1 - Inhibitor Site

The GP inhibitor site (Figure 3.1), which is also known as the caffeine binding site ($K_i \sim$ 0.1 mM) [78, 79], is located on the surface of GP at the entrance to the catalytic site (some 12 Å from it). It is a hydrophobic binding pocket and when in the T state conformation, Phe285 from the 280s loop is stacked close to Tyr613 with these two aromatic residues forming the core of the inhibitor site. When inhibitors bind at this site, they stabilize the GP T-state conformation and block access to the catalytic site, thereby inhibiting the enzyme. Inhibition at this site is synergistic with glucose, suggesting that inhibition could be regulated by blood glucose levels and would decrease as normoglycaemia is reached and diminish the risk of hypoglycemia. Purines, nucleosides, nucleotides [49], some indirubins [80-82] and flavonoids [75] together with flavopiridol [74, 75] are known to bind at this site, by exploiting π-π stacking interactions of their aromatic rings between the side chains of Phe285 and Tyr613.

(a)   (b)
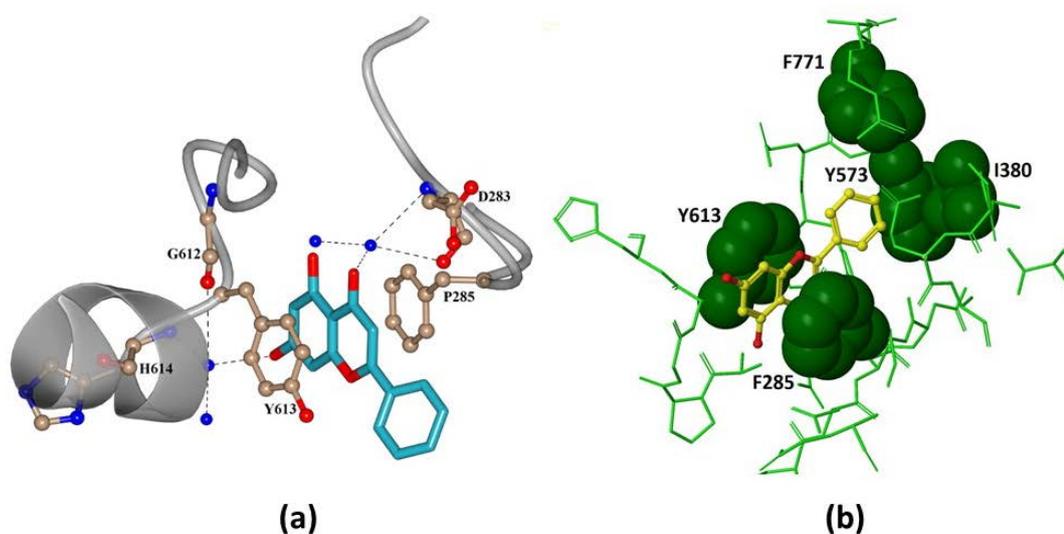
Figure 3.3 - Binding of the flavone chrysin at the GP inhibitor site as determined using X-ray crystallography (PDB code: 3EBO) [75]. The key interactions are shown in **(a)** with the compound sandwiched between the Phe285 and Tyr613 residues exploiting π-π stacking interactions and a number of water bridging (blue) protein-ligand interactions. In **(b)**, the hydrophobic pocket lined by Phe771, Tyr573 and Ile380 in the vicinity of the flavonoid ring B is highlighted.

The binding of chrysin at the inhibitor site (PDB code: 3EBO) [75] is shown in Figure 3.3. As first thought, the main characteristic feature for chrysin binding to GPb is the

stacking interactions between the two aromatic residues, Phe285 (from the 280s loop) and Tyr613. When chrysin binds to GPb, it forms a few polar/polar interactions and exploits numerous van der Waals contacts dominated by the substantial contacts made to Phe285 and Tyr613 by the aromatic ring [83]. Of particular significance to binding of flavonoids at this site is the hydrophobic pocket formed by Phe771, Tyr573 and Ile380, as highlighted in Figure 3.3(b). It has been postulated that flavonoids with polar OH substituents on ring B and in particular at position 4' are unfavorable for inhibitor site binding on the basis of quantum mechanics – molecular mechanics / Poisson Boltzmann surface area (QM/MM-PBSA) binding free energy (ΔG_{bind}) calculations [75]. In support of this, calculated ΔG_{bind} values for quercetagetin with 3' and 4' OH substituents suggested poor binding at the site, in agreement with multiple inhibition studies using Dixon plots which suggest that quercetagetin binds at the allosteric site [73]. The hypothesis was further corroborated by the recent discovery of the quercetin binding site discussed below [67].

The most potent natural compound currently known to bind at the inhibitor site is ellagic acid. Kyriakis et al. recently published a comprehensive kinetics and crystallographic study on the binding properties of gallic acid and its dimer ellagic acid to GP [76]. Ellagic acid was found to bind with K_i's of 7.5 and 13.4 μM for GPa and GPb, respectively [76]. In comparison, the binding of gallic acid displays weaker K_i's of 3.9 and 1.7 mM, respectively. Ellagic acid (but not gallic acid) binds in a strongly synergistic mode with glucose. The inhibition behavior of both compounds was additionally investigated with respect to caffeine and allosteric activator AMP. Multiple inhibition studies using Dixon plots revealed that the binding of ellagic or gallic acid and caffeine is mutually exclusive, that is, they compete for the same binding site. Similar analysis using AMP which binds at the allosteric site revealed that gallic or ellagic acids and AMP are not mutually exclusive. The experiments, therefore, indicated that gallic and ellagic acid bind to the inhibitor binding site. The reported crystal structures of the RMGPb-gallic acid (PDB code: 4Z5X) and RMGPb-ellagic acid (PDB code: 4YUA) confirmed binding at the site. The binding modes of both ligands are similar. As expected, the characteristic binding feature is stacking interactions between Phe285 and Tyr613, as shown in Figure 3.4 for the RMGPb - ellagic acid complex. Ellagic acid exploits nearly twice as many van der Waals interactions with GPb compared to gallic acid; it forms one direct hydrogen bond to the Asn282 sidechain (ND atom) but has a number of protein-ligand water bridged interactions with Asp283, Ile570, Ala610, Gly612, His614 and Met615. On binding at the inhibitor site, solvent

accessible surface area calculations indicate that ~ 74% of both ligands become buried, almost all from non-polar groups. Kyriakis et al. found an approximate correlation between ligand binding affinities acting at the inhibitor site and total buried surface areas (protein + ligand).
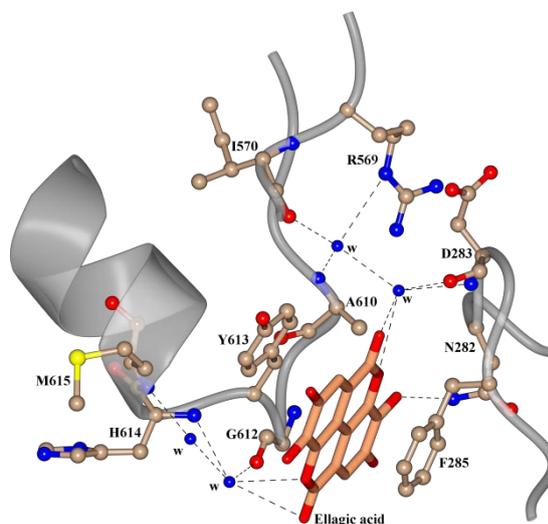
Figure 3.4 - Binding of the ellagic acid at the GP inhibitor site as determined using X-ray crystallography (PDB code: 4YUA) [76].

The most potent compounds binding at the inhibitor site discovered to date are flavopiridol (a synthetic flavonoid) [74, 75] and flavopiridol analogues [84] (Table 3.3). While flavopiridol has a K_i of ~1.2 µM, analogues of flavopiridol with a flattened tetrahydro-pyridine ring and halogen substituted phenyl ring B have revealed potencies ranging from 0.83–2.75 µM. The most favorable substitution in this series is the 3-Cl analogue. Halogen substitution and binding effects through the σ-hole phenomenon are now widely recognized and exploited in rational drug design [85], a successful example of which can also be found for the design of GP catalytic site inhibitors [39]. Replacement of the B phenyl ring of the flavopiridol derivatives by a methyl ($Ki = 241$ µM) or a *tert*-butyl (no inhibition) group was found to be detrimental to potency [84].

Table 3.3. The chemical structures of the most potent inhibitors known to bind at the GP inhibitor site together with their K_i's for RMGPb inhibition

Flavopiridol	K_i (µM)	Flavopiridol derivatives	R	K_i (µM)
	1.16 [74] 1.24 [75]		H	1.28 [84]
			2-Cl	0.99 [84]
			3-Cl	0.83 [84]
			4-Cl	1.89 [84]
			2-F	2.75 [84]
			2-Br	1.50 [84]
			3,5-diCl	2.50 [84]

The crystal structure of flavopiridol in complex with RMGPb (PDB codes: 1C8K, 3EBP) has been solved [74, 75] and the structural features of inhibition compared to those of chrysin (PDB ID: 3EBO) [75]. The common cores of flavopiridol and chrysin superimpose quite well [75] stacked between Phe285 and Tyr613. Flavopiridol, the binding of which is shown in Figure 3.6 makes a few polar/polar interactions with water molecules and has numerous van der Waals contacts. The chlorophenyl ring B extends towards the hydrophobic pocket lined by Phe771, Tyr573 and Ile380 (as highlighted in Figure 3.3(b)) but the chlorine atom does not form a strong halogen bond with GPb and is mainly involved in van der Waals interactions.
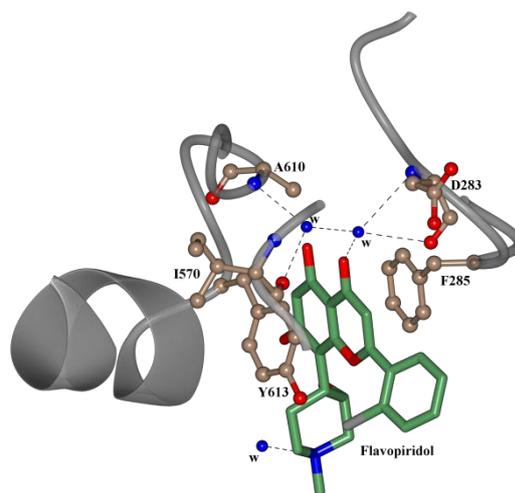
Figure 3.5 - Binding of flavopiridol at the GP inhibitor site as determined using X-ray crystallography (PDB code: 3EBP) [75].

3.4.2 - Quercetin Binding Site

Recent screening of thirteen polyphenolic extracts obtained from the vinification byproducts of *Vitis vinifera* that displayed significant inhibitory potency for RMGPb *in vitro* (IC_{50} values in the range of low μg/mL) revealed that the most active ingredient of these extracts is the flavonoid quercetin [67]. Quercetin was found to bind to a novel binding site, which was unknown at the time. This new site is 15 Å away from the active site, 43 Å from the allosteric site and 32 Å from the inhibitor site [67]. The site is a shallow groove, created by Lys544, Arg551, Lys655, and Tyr548 on one side and Glu120 and Glu123 on the other side. Quercetin, upon binding to GP, forms hydrogen bond and van der Waals interactions with Arg551, Lys544. It is also involved in water-bridging interactions with Glu121, Lys655, Leu494, Cys495, and Glu654 (Fig. 3.7).
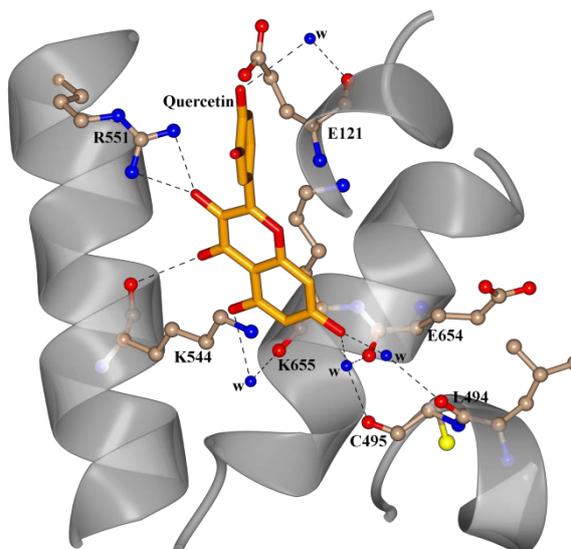
Figure 3.6 - Binding of quercetin at the GP inhibitor site as determined by X-ray crystallography (PDB code: 4MRA) [67].

3.4.3 - Allosteric and New Allosteric Sites

As stated earlier, quercetagetin binds at the allosteric site on the basis of multiple inhibition studies using Dixon plots [73]. With respect to the new allosteric site, binding of gallocatechin gallate (GCG) (K_i = 6.3 µM for RMGPb) at this site has been proposed also on the basis of multiple inhibition studies using Dixon plots [68]. The new allosteric site is located inside the central cavity formed by the association of the two GP subunits, 15 Å from the allosteric site, 33 Å from the catalytic site and 37 Å from the inhibitor site.

3.5 – Flavonoid Anti-Oxidant Effects

Multiple studies have been conducted investigating flavonoid antioxidant and antiradical activities. Reactive oxygen species (ROS) and reactive nitrogen species (RNS) are both beneficial and harmful to cells [86]. However, when there is an imbalance between free radical production and antioxidant capacity, a state of oxidative stress arises. Oxidative stress plays an important role in the pathogenesis of a number of disorders such as cardiovascular disease, inflammation, and neurodegenerative diseases [87, 88]. Elevated ROS levels have also been connected to T2D [89, 90] and the

resulting oxidative stress is a major cause of hyperglycemia-induced diabetic complications [89-93]. Hyperglycemia stimulates the formation of ROS/RNS from sources such as the mitochondrial complex II, oxidative phosphorylation, glucose autooxidation, NAD(P)H oxidase, lipooxygenase, cytochrome P450 monooxygenases and nitric oxide synthases (NOSs). NADPH oxidases produce superoxide and/or hydrogen peroxide as intermediates of redox reactions [94]. Multiple independent strategies that can alleviate mitochondrial ROS production were shown to prevent some of the typical secondary complications that arise from type 2 diabetes, including the activation of protein kinase C or NF-kB and the formation of advanced glycation end products. Superoxide is also generated by glucose auto-oxidation which is related to the formation of glycated proteins in the plasma of diabetic patients [95, 96]. The interaction of these advanced glycated proteins with their corresponding cell surface receptors further stimulates ROS production and reduces intracellular glutathione levels [97]. Furthermore, hyperglycemia boosts cell-mediated low-density lipoprotein (LDL) peroxidation in endothelial cells [98].
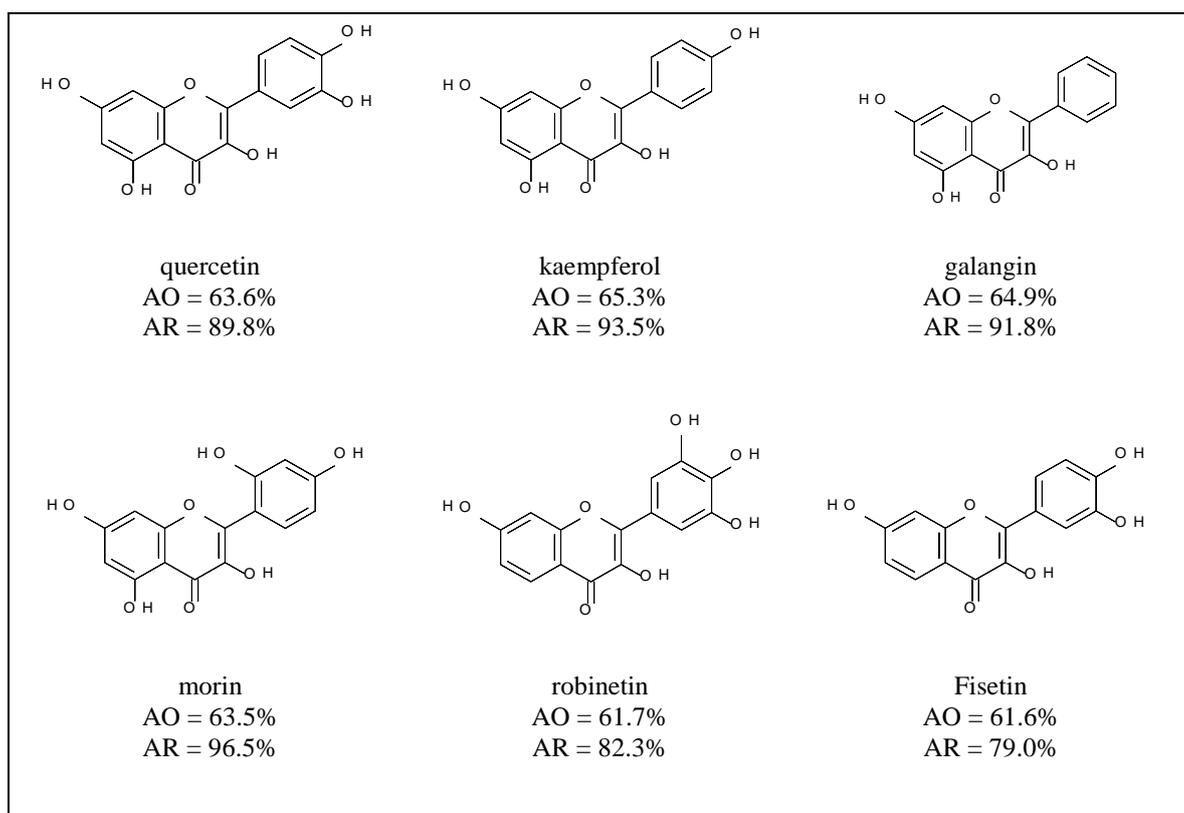
Figure 3.7 - Antioxidant (AO) and antiradical (AR) activities of flavonoids [99] as discussed in the text. % AO activity was calculated as % inhibition of oxidation *with* versus *without* flavonoid present (β-carotene and linoleic acid system). The scavenging of free radicals

using DPPH assay (% AR activity) was also measured versus a reference *without* flavonoid as described in the original reference [99].

Treatment with antioxidants has been shown to alleviate many diabetic complications [100-103], although the benefits of antioxidant supplements are controversial [104]. Flavonoids have been reported as scavengers of a wide range of ROS and RNS, as well as inhibitors of lipid peroxidation [105, 106]. The number of hydroxyl groups, a C2-C3 double bond (C ring) and a 3-OH group have been referenced as conditions for favourable antioxidant and antiradical activities [99, 107-109]. Burda et al. additionally highlighted the importance of a flavonol structure or 4'-OH substituent to antiradical activity [99]. In a study of 42 structurally diverse flavonoids for their antioxidant and antiradical activities [99] measured by heat-induced oxidation in a β-carotene and linoleic acid system and by 1,1-diphenyl-2-picrylhydrazyl (DPPH) decoloration test, respectively, among the most active that also exhibits potent GP inhibition was quercetin (63.6% antioxidant (AO) activity; 89.8% antiradical (AR) activity). Other examples of flavonoids which revealed both good antioxidant and antiradical activities (Fig. 3.8) were kaempferol (65.3% AO; 93.5% AR), galangin (64.9% AO; 91.8% AR), morin (63.5% AO; 96.5% AR), robinetin (61.7% AO; 82.3% AR) and fisetin (61.6% AO; 79% AR). All of these compounds are flavonols and have a 3-OH substituent. Flavonols with a free 3-OH were very effective antioxidants and also had high activities towards scavenging DPPH radicals. Blockage of the 3-OH group by glycosylation or methylation resulted in a complete loss of antioxidant activity, consistent with other studies [110]. Flavonoids without hydroxyl groups or with hydroxyls only at C-5 and/or C-7 had no effect on scavenging of free radicals; the effect of a 4'-OH substituent on the hand, as mentioned above, was significant. It would be interesting to test robinetin and fisetin for their GP activity to establish the importance of the 5-OH group to potency, and morin because of the 3'-OH and 5'-OH disubstitution on ring B. Galangin, with its unsubstituted B ring, is potentially a potent GP inhibitor acting at the inhibitor site.

3.6 - Conclusion

The development of antidiabetic medication from natural products has recently begun to receive considerable attention. A number of studies have highlighted the benefits of polyphenolic natural compounds with hypoglycemic effect and their importance in the management of diabetic complications. These compounds can have nutraceutical potential and/or can provide scaffolds for the development of more effective

pharmaceuticals for T2D treatment. This section highlights the potential antihyperglycaemic effects of natural flavonoids acting through inhibition of GP, a validated target for T2D treatment. A number of these compounds are low µM inhibitors and the structural basis of the potency as currently known has been analyzed, providing information for further structure based inhibitor design efforts. These substances can also serve as lead compounds with sufficient structural diversity to be of considerable importance in the development of new antihyperglycaemic agents. Some of the compounds also exhibit anti-oxidant effects relevant to prevention of diabetic complications [111].

Chapter 4 - Computational Approaches to Drug Design

4.1 - Introduction

Medicines are often derived from natural sources such as plants as described in detail for flavonoids targetting GP in the previous chapter. Even though this is a continued area of study, there is now the potential to synthesise a large array of compounds which aren't currently found in nature. Combine this with the endless possibilities of derivatives and sub groups creates a vast library of potential drugs. However, this is not practical, as manually synthesising these compounds, then taking them through to human-trails is incredibly expensive and time consuming. This enormous library can be reduced, through target based screening during the early stages of drug discovery, which is far more cost-effective than physically testing each compound. Key to this process in modern day drug design efforts is the application of molecular modelling methods with the benefit significantly savings in terms of both time and money. These methods have become an integral part of drug discovery projects both in academia and industry. There are two main types of drug design: structure-based drug design (SBDD) and ligand-based drug design (LBDD). In SBDD, the structure of the target is known and can be exploited to guide ligand design efforts. For LBDD, the structure of the target is unknown and relies on binding data being available for a set of ligands which can then be compared and modifications suggested exploiting methods such as quantitative structure activity relationship (QSAR) analysis. When a set of actives (compounds which have activity against the target site) are known for a given target, pharmacophore modelling can also be used to identify common structural features (and distances between them) such as hydrogen bond donors, hydrogen bond acceptors, halogen groups, aromatic groups and hydrophobic group positions that are crucial to activity. The pharmacophores can then be used to search many of the available on-line databases such as ZINC (www.zinc.docking.org) [112], the database we have employed in the current study. For SBDD, docking is the most commonly applied method, whereby ligand binding to the receptor is explicity predicted exploiting the many known biological macromolecular targets solved and deposited in the RCSB Protein Database (www.pdb.org). This can be used to screen and identify potential hit compounds from different sources and the on-line databases previously mentioned. Although primarily used for hit identification, docking accuracy has increased dramatically in recent years

such that docking may now also help in lead optimization efforts, suggesting modifications that have the potential to improve activity [113].

In the current work, targeting the inhibitor site of GP, the core interactions in the binding site are π-stacking interactions between the flavonoid(s) and Phe285 and Tyr613. Accurate modelling of these interactions are currently beyond the scope of docking methods due to their electronic nature, so that we employ docking with constraints only to generate initial potential binding poses for all our predicted flavone analogues but then more accurately describe the protein-ligand binding using quantum mechanics (QM). As it would be impossible to describe the full system using QM, only the active site is modelled using QM, with the rest of the protein modelled using molecular mechanics (MM) in a method known as QM/MM. The current chapter will describe the foundations of the key computational approaches used in this work, specifically docking (with an emphasis on Glide) and the quantum mechanics methods we used to describe the active site binding process (based on density functional theory - DFT).

4.2 - Molecular Docking

Ligand-protein docking is an *in silico* approach and an extremely important step in drug design. The aim is to structurally predict the most likely binding modes and energetically the binding affinities between a small molecule and its target (Figure 4.1). Docking mimics the way in which a ligand can bind to the target site of a protein and will then generate multiple binding geometries, referred to as poses.
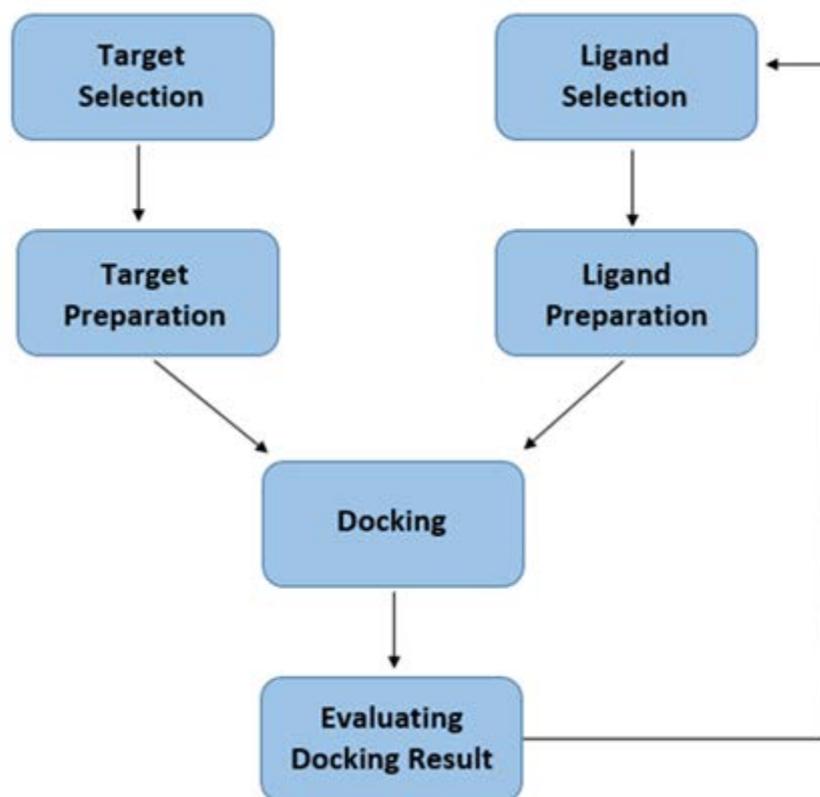
Figure 4.1 - A basic workflow showing the generalised proccess of docking

A scoring function then predicts the binding energies between the ligand and receptor for each of these generated poses. These are then evaluated and ranked based on their binding energies; the top ranked pose, in theory, should be the most favourable conformation of the ligand in the binding site.

There are two types of docking, rigid and flexible. In rigid docking, the protein and ligand are treated as rigid, restricting the search space. The ligand is pre built before-hand. An image of the target binding site is built up from multiple overlapping spheres of different sizes, which have been mapped from the molecular surface of the target protein conformation and structure. With this map, the ligand is then overlaid with the centre of generated spheres (Fig 4.2). If the conformation fits the ligand is then placed into the target site, minimised (the conformation is adjusted so that the structure is at the lowest energy), and the position is scored.
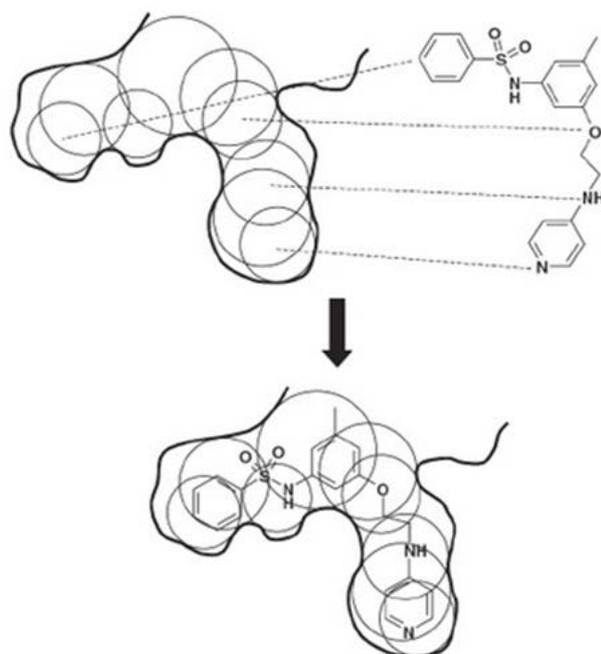
Figure 4.2 – A diagram showing how the overlapping spheres build the active site, then how important atoms on the ligand are orientated to fit within the newly generated site.

Flexible docking however, generates different conformations of each ligand 'on-the-fly' using a search algorithm. This style is much more effective, and is much more commonly used. This is because when conformational change becomes a bigger factor, the accuracy of rigid docking decreases; for example, hydrogen bonds and large scale features, such as hydrocarbon chains, might be missed due to a lack of flexibility. Also, many flexible ligands may be incorrectly dismissed if they are docked as rigid bodies in the wrong conformation.[114] The docking process starts by breaking down the molecule into its conformationally flexible and its rigid fragments. An example of a flexible docking protocol is shown in Figure 4.3. This is done to save time, as the rigid fragment can be left in place, rather than running multiple calculations on the 'same' conformation. Then the orientated fragments are determined producing n ligand conformations, and are joined onto the different rigid placements, so, for each rigid orientation, n different conformations will be created.
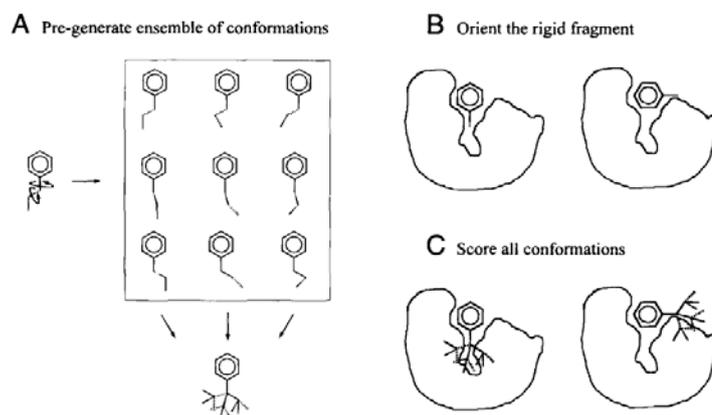
Figure 4.3 - Overview of rigid docking ligand ensemble method. A: rigid atoms are fixed in position and the conformational space of the rest of the molecule is systematically sampled at 60° or 120° increments. B: The rigid fragment is orientated in the target site. C: All *n* flexible fragments of the molecule are scored in the orientation of the rigid fragment.

4.3 - Scoring Functions

Scoring functions are in place to distinguish whether or not poses are correct or incorrect, or whether the ligand is active or inactive. This is achieved by calculating the binding affinity for the particular pose, if a ligand is deemed active it means that it is predicted to have biological activity.

Scoring functions take into account the various contributions of free energy binding as an additive equation. For example

$$\Delta G_{bind} = \Delta G_{solvent} + \Delta G_{conf} + \Delta G_{int} + \Delta G_{rot} + \Delta G_{t/r} + \Delta G_{vib}$$

where $\Delta G_{solvent}$ is how the solvent effect arises from the interactions between the solvent and the ligand, protein and intermolecular complex. ΔG_{conf} arises from conformational changes within the protein and in the ligand. In the majority of cases, the protein does not change significantly on binding and most docking methods assume a rigid receptor. However, the ligand changes from a large variety of conformations in solution to a single dominant conformation in the bound state. Where ΔG_{int} is the free energy due to specific protein-ligand interactions. ΔG_{rot} is the free energy loss associated with holding internal rotations of the protein and the ligand in place. Where $\Delta G_{t/r}$ is the loss in translational and rotational free energy due to two bodies (the ligand and the receptor forming a single body) known as the intermolecular complex. This is

usually assumed to be the same for all ligands and so is ignored if you are only interested in the relative binding strengths of different ligands to a certain protein. ΔG_{vib} is the free energy due to changes in vibrational modes. This contribution is hard to calculate and so is generally ignored [115].

There are three classes of scoring functions: Classical force-field-based, Empirical and knowledge-based potentials.

Classical force-field based: These functions are developed based on physical atomic interactions. These interactions include van der Waals interactions, electrostatic interactions and bond stretching, bending, torsional forces, between the ligand and the ligand-receptor complex. The docking is scored based on terms of molecular mechanics forcefields.

Empirical: Empirical scoring functions are derived to reproduce data obtained from experimentally determined complex structures [116]. This approach is useful, but is hindered by the requirement for experimentally obtained analytical data.

Knowledge-based potentials: (also referred to as statistical-potential based scoring functions) employ energy potentials which are derived from the structural information embedded in experimentally determined atomic structures [117].

4.4 - Glide Docking and Scoring Functions

Glide, the program we employ in this work, uses a funnel like system of hierarchical filters to search for possible conformations of the ligand in the active-site of the target receptor. (Figure 4.4)
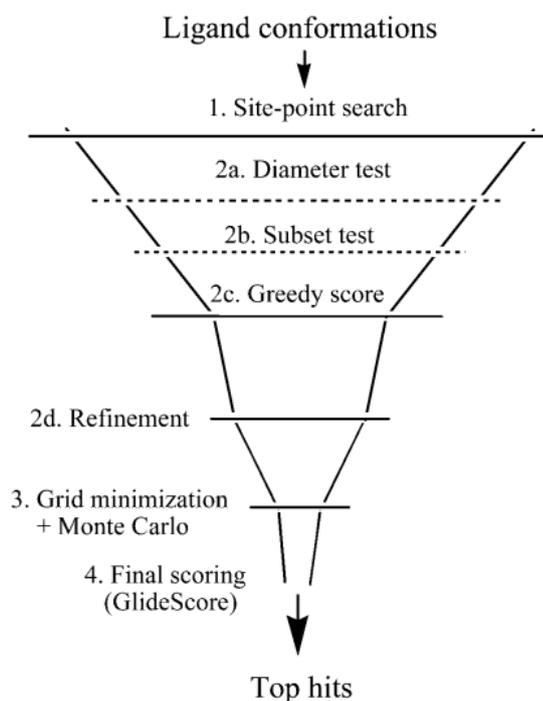
Figure 4.4 - Glide docking "funnel" stages, showing the glide docking hierarchy.

The shape and properties of the receptor are mapped on a grid, which is prepared in the pre-processing steps, by different sets of fields that provide progressively more accurate scoring of the ligand pose. The next step produces a set of initial ligand conformations. These are then subjected to initial screens into the entire phase space to find promising ligand poses, skipping poses that sterically clash with the receptor. These poses are then taken forward and are minimised using the OPLS-AA minimisation forcefield, with a small amount of the lowest energy poses being subjected to a Monte Carlo procedure which randomly adjust some of the torsional angles, and then minimises them, to find potentially better poses. Finally the minimised poses are re-scored with Schrödinger's GlideScore scoring function, which ranks the successful poses. GlideScore (GS) (2.5) is derived from modifying the ChemScore function, as follows:

GS = (vdW) + (Coul) + (Lipo) + (H-bonds) + (Metal) + (Rewards) + (RotB) + (Site)

Where (vdW) is the van der Waals evergy taking into account non-bonding interactions. Where (Coul) is the Coulumbic term, which accounts for electron-electron repulsion. (Lipo) is the lipophilic term which is derived from experimental data. (Metal) is for any metal-ligand interactions that may be present. (Rewards) takes into account favourable bindings whilst penalising polar groups in regions of hydrophobicity. (RotB) stands for rotatable bonds, scoring lower for non-moving rotatable ligand binding bonds. Finally (Site) gives higher scores for non-hydrogen bonding areas of hydrophobicity.

4.5 - Quantum Mechanics Methods

4.5.1-Introduction

Electrons are very light particles, they display both wave and particle characteristics, and must be described in terms of wavefunction, Ψ. Quantum mechanics explicitly represents the electrons in the calculation, so it is possible to derive properties that depend upon the electronic distribution. The term *ab intio* is Latin for "from the beginning", so *ab initio* methods relate to methods derived from first principles, purely based on theory and with no input from experiment. Most *ab initio* methods try to solve the time-independant Schrödinger Wave equation (SWE), which is given as:

$$\left\{ -\frac{\hbar^2}{2m} \nabla^2 + V \right\} \Psi(r) = E\Psi(r) \qquad \textbf{Eq. (4.1)}$$

Where E is the energy of the particle. The equation can be written in shorthand form with (equation $\hat{H}\,\Psi = E\,\Psi$). Where $\hat{H}$ is the Hamiltonian operator:

$$\hat{H} = -\frac{\hbar^2}{2m} \nabla^2 + V \qquad \textbf{Eq. (4.2)}$$

V in this equation represents the potential energy. Whereas the first term in the equation is the kinetic energy operator with ∇ having the following form:

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \qquad \textbf{Eq. (4.3)}$$

The Hamiltonian operator $\hat{H}$ acts on the wave function Ψ, which characterizes the particle's motion. It contains all the measurable information about the particle. The wavefunction represents the probability amplitude for finding a particle at a given point in space at a given time, where $|\Psi|^2$ is the probability distribution of the particles, the integration of the probability of finding the particle over space must be 1:

$$\int \Psi^* \Psi \, d\tau = 1 \qquad \textbf{Eq. (4.4)}$$

The wavefunction of a complete system is commonly composed of a set of single-particle orbitals [118]

The Hamiltonian can be written in full as:

$$\hat{H} = \sum_i^{electrons} \frac{-\hbar^2}{2m_e} \nabla_i^2 + \sum_A^{nuclei} \frac{-\hbar^2}{2m_A} \nabla_A^2 + \sum_i^{electrons} \sum_A^{nuclei} \frac{-e^2 Z_A}{r_{iA}} +$$

$$\sum_{i>j}^{electrons} \frac{e^2}{r_{ij}} + \sum_{A>B}^{nuclei} \frac{e^2 Z_A Z_B}{r_{AB}} \qquad \textbf{Eq. (4.5)}$$

Where the consecutive terms in the equation represents:

- Kinetic energy of the electrons (T_e)
- Kinetic energy of the nuclei (T_n)
- Electrostatic interaction between the electrons and the nuclei (V_{e-n})
- Electrostatic interaction between the electrons (V_{e-e})
- Electrostatic interactions between the nuclei (V_{n-n})

4.5.2 - Variational Theorem

According to the Variational Theorem, the expectation value of the Hamiltonian is the variational energy according to Eq 4.6:

$$\frac{\int \Psi^* \hat{H} \Psi d\tau}{\int \Psi^* \Psi d\tau} = E_{var} \geq E_{exact} \qquad \text{Eq. (4.6)}$$

where an approximate wavefunction for a molecular system, when substituted into the Schrödinger equation, will always yield a higher energy than the actual energy of the system. The more precise the wavefunction that is chosen, the closer the calculated energy will be to the true energy. The computational method using this principle to obtain approximations to correct wavefunctions is called the Variational Method. [119]

4.5.3 - The Born-Oppenheimer Approximation

Since exact solutions to the SWE are not available for systems consisting of more than two particles, only approximate solutions may be obtained. One of the first approximations to be introduced, the Born-Oppenheimer Approximation, concerns the mobility of the nuclei in a system. Since the masses of the nuclei are much greater than those of the electrons, we can consider the nuclei stationary and the Schrödinger equation can be solved for the electrons moving in the stationary potential generated by the fixed nuclei approximation. This means that the Hamiltonian operator may be approximately written in terms of two separate operators, the nuclear operator $\hat{H}_{nuc}$ and the electronic operator $\hat{H}_{el}$ may be expressed as:

$$\hat{H}_{el} = T_e + V_{n-n} + V_{e-e} + V_{e-n} \qquad \text{Eq. (4.7)}$$

4.5.4 – Hartree-Fock Approximation

Another key approximation is the Hartree Approximation, this assumes that a many electron wavefunction can as written as a product of one electron functions, in other words the all electron wave function Ψ is described as a simple product of one electron wavefunctions $\psi(\vec{r}_i)$ (just under 2.11), i.e, the orbital functions.

$$\Psi = \Psi(\vec{r}_1, \vec{r}_2, \vec{r}_3, \ldots \ldots \vec{r}_N) \approx \psi_1(\vec{r}_1)\psi_2(\vec{r}_2)\psi_3(\vec{r}_3) \ldots \ldots \psi_N(\vec{r}_N) \quad \textbf{Eq. (4.8)}$$

If we use the variational energy, solving the many electron Schrödinger equation is reduced to solving a series of one electron Schrödinger equations, each electron interacts with the average distribution of other electrons. An effective Hamiltonian operator $h(i)$ can be defined to act only on the electron i and the problem can be rewritten as a system of one electron wave functions:

$$h_i\psi_i(\vec{r}_i) = \varepsilon_i\psi_i(\vec{r}_i) \quad \textbf{Eq. (4.9)}$$

However, the Hartree Approximation makes no distinctions with respect to spin, and the one electron wavefunctions have to be described not only in terms of the position of the electron, $\vec{r}_i$ but also in terms of the spin, $\vec{s}_i$. $\vec{s}_i$ can be one of two orthogonal spin states $\alpha(\omega)$ or $\beta(\omega)$, corresponding to "spin up" $\omega = +\frac{1}{2}$ and "spin down" $\omega = -\frac{1}{2}$ respectively. The single electron wavefunction is therefore written more correctly in the form of a so called one electron spin orbital (Molecular Orbital, MO):

$$\chi(\vec{x}) = \chi(\vec{r}, s) = \begin{cases} \psi(\vec{r})\alpha(\omega) \\ or \\ \psi(\vec{r})\beta(\omega) \end{cases} \quad \textbf{Eq. (4.10)}$$

However, an N electron wavefunction expressed as a function of N spin orbitals:

$$\Psi = \Psi(\vec{r}_1, \vec{r}_2, \vec{r}_3, \ldots \ldots \vec{r}_N) \approx \chi_1(\vec{r}_1)\chi_2(\vec{r}_2)\chi_3(\vec{r}_3) \ldots \ldots \chi_N(\vec{r}_N) \quad \textbf{Eq. (4.11)}$$

is still unacceptable, as it does not have the property of antisymmetry. The Pauli Exclusion principle states that a many electron wavefunction must be antisymmetric with respect to the interchange of the coordinates $\vec{x}$ (both space $(\vec{r})$ and spin$(\vec{s})$) of any two electrons. What is needed therefore, is an approximation which is as easy to handle as the Hartree one and also respects Pauli's Principle. One solution to this lies in the use of Slater Determinants suggested by Fock. The simplest antisymmetric wavefunction which can be used to describe the ground state of an N electron system is a single Slater determinant.

$$\Psi(\vec{x}_1, \vec{x}_2, \ldots \ldots \vec{x}_N) = \Phi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_i(\vec{x}_1) & \chi_j(\vec{x}_1) & \cdots & \chi_k(\vec{x}_1) \\ \chi_i(\vec{x}_2) & \chi_j(\vec{x}_2) & \cdots & \chi_k(\vec{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_i(\vec{x}_N) & \chi_j(\vec{x}_N) & \cdots & \chi_k(\vec{x}_N) \end{vmatrix} \qquad \textbf{Eq.}$$

(4.12)

Where Φ represents the Slater determinant. The mathematical properties of the Slater determinant guarantee the above mentioned property that is necessary for any multielectron wavefunction: antisymmetry. Antisymmetry, the exchange of the positions and spins of any two electrons, is equivalent to the exchange of two rows in the determinant and results in a change of sign in Ψ. [120]

Hartree-Fock theory is based on the variation principle in quantum mechanics. If we take Hartree-Fock wave function, equation (4.13)

$$\Psi = |\chi_1 \chi_2 \cdots \chi_3| \qquad \textbf{Eq. (4.13)}$$

and put it into the variational energy expression (4.6), and minimise the energy with respect to changes in the orbitals

$$\partial E / \partial \phi_i = 0 \qquad \textbf{Eq. (4.14)}$$

this yields the Hartree-Fock equation

$$f(i)\chi_i(\vec{x}_i) = \varepsilon_i \chi_i(\vec{x}_i) \qquad \textbf{Eq. (4.15)}$$

$f(i)$ represents the one electron Fock operator for electron i. The eigenvalues ε_i are interpreted as molecular orbital energies. These orbital energies have a simple physical interpretation: they give the amount of energy necessary to take the electron out of the molecular orbital, which corresponds to the negative of the experimentally observable ionization potential (Koopmans' theorem).

The essence of the Hartree-Fock approximation is that it reduces the many electron problem into a single electron problem and explicit electron-electron interaction can be replaced with the Hartree-Fock potential, $v_{HF}(i)$, which can be understood as the average "field " or average potential experienced by the i-th electon due to all other electrons. This can be written mathematically as:

$$v_{HF}(i) = \sum_{j=1}^{N}(J_{ij} - K_{ij}) \qquad \textbf{Eq. (4.16)}$$

and the all electron Fock operator $\hat{F}$ expressed as:

$$\hat{F} = \hat{T} + \hat{V}_{NE} + \hat{J} - \hat{K} \qquad \textbf{Eq. (4.17)}$$

with $\hat{J}$ representing the Coulomb operator for electron-electron repulsion and $\hat{K}$ the exchange operator which is purely quantum mechanical rising from the fact that the wave function must change sign when exchanging two electrons

$$\hat{K}\chi_i = \left\{ \sum_j^{electrons} \int \chi_j \frac{e^2}{r_{ij}} \chi_i d\tau \right\} \chi_j \qquad \textbf{Eq. (4.18)}$$

The negative sign in eq 4.17 reminds us that spin correlation keeps the electrons apart and so reduces the classical Coulumbic repulsion.

Solving the Hartree-Fock equation requires an iterative approach known as the self consistent field (SCF) method. direct solution of the Hartree-Fock equations is not a practical proposition for molecules. The most popular strategy is to write each spin orbital as a linear combination of single electron orbitals (atomic orbital, LCAO):

$$\chi_i = \sum_\mu c_{\mu,i} \, \phi_\mu \qquad \textbf{Eq. (4.19)}$$

Where $c_{\mu,i}$ is the molecular orbital coefficient and ϕ_μ is the one-electron orbitals which are commonly called basis functions and often correspond to the atomic orbitals.

The objective is to determine the set of coefficients that gives the lowest energy of the system

$$\partial E / \partial c_{\mu,i} = 0 \qquad \textbf{Eq. (4.20)}$$

4.5.5 - Roothaan-Hall Equations

For a closed-shell system with N electrons in N/2 orbitals, the derivation of the Hartree-Fock equations was first proposed by Roothaan and independently by Hall (1951). They recast the Fock equations in matrix form, which can be solved using standard techniques and can be applied to systems of any geometry.

$$FC_i = \varepsilon_i S C_i \qquad \textbf{Eq. (4.21)}$$

Where F is the Fock matrix, C_i is the column vector of the molecular orbital coefficients, ε_i is the orbital energy, S is the overlap matrix. Hartree-Fock can predict

lots of properties with reasonable accuracy, e.g., equilibrium structures and relative energies. The one significant drawback to Hartree-Fock Theory is that it only treats electron correlation in an average, not instantaneous sense. In fact, the motions of electrons are correlated and they tend to avoid each other, giving rise to a lower energy. Their motion must be "correlated". For a given basis set, the correlation energy is:

$$E_{correlation} = E_{exact} - E_{HF} \qquad \textbf{Eq. (4.22)}$$

Neglect of instantaneous electron-electron interactions, gives rise to errors and is particularly relevant in accounting for dispersion interactions of significant importance to the current study. [118]

4.5.6 - Density Functional Theory

Density Functional Theory (DFT) is an approach to solve the electronic structure of atoms and molecules which has enjoyed an increasing interest since the late 1980s and 1990s. Its advantages include less demanding computational effort and less computer time and in many cases better agreement with experiment than is obtained from Hartree-Fock based methods. DFT replaces the wave function (which is the function of 3n co-ordinates, where n is the number of electrons) with the electron density (ρ) (which is a function of only 3 co-ordinates) as the fundamental unknown.

The 'functional' part of the name comes from the fact that the energy of the molecule is a function of the electron density, and the electron density is itself a function of the positions of the electrons $\rho(r_1)$.

Whereas in the Hartree-Fock method, the expression for energy is

$$E_{HF}(\Psi) = T_e + V_{ne} + \overbrace{J_{ee} - K_{ee}}^{V_{ee}} \qquad \textbf{Eq. (4.23)}$$

Where T_e is the electronic kinetic energy, V_{ne} is the potential energy due to the nuclear-electronic Coulumbic attraction. The last two terms on the right hand side correspond to the classical Coulomb repulsion of the electrons (J_{ee}) and non-classical (K_{ee}) due to the quantum nature of the electrons. The difference between these terms accounts for the electron-electron repulsion, V_{ee}. In 1964, P. Hohenberg and W. Kohn were able to prove

that the exact ground-state energy of a molecule is uniquely determined by its electron probability density. They showed it is possible to write

$$E[\rho] = E_{classical}[\rho] + E_{xc}[\rho] \qquad \textbf{Eq. (4.24)}$$

where $E_{classical}[\rho]$ is the sum of contributions of K.E., electron-nuclear interactions and the classical electron-electron potential energy, and $E_{xc}[\rho]$ is the exchange-correlation energy. This term takes into account all the non-classical electron-electron effects due to spin and applies small corrections to the K.E. part of $E_{classical}$ that arise from electron-electron interactions. The DFT expression for the energy is now:

$$E_{DFT}[\rho] = T_e[\rho] + V_{Ne} + J_{ee} + E_{XC}[\rho] \qquad \textbf{Eq. (4.25)}$$

where the exact Hartree-Fock exchange is replaced by the XC energy density functional that accounts for both the exchange and the correlation energy and can be written as:

$$E[\rho] = T[\rho] + \int \rho V_{nuclei} + \int \frac{\rho(r)\rho(r')}{|r-r'|} dr dr' + E_{XC}[\rho] \qquad \textbf{Eq. (4.26)}$$

where $T[\rho]$ is the Kinetic energy, $\int \rho V_{nuclei}$ is the electron-nuclei attraction, $\int \frac{\rho(r)\rho(r')}{|r-r'|} dr dr'$ is the electron-electron repulsion and $K_{XC}[\rho]$ is the exchange / correlation. While the Hohenberg-Kohn theorem just mentioned guarantees, the existence of $K_{XC}[\rho]$ but like so many existing theorems in mathematics, it gives no clue as to how it should be calculated. The relevant equations were deduced by Kohn and Sham in 1965 who showed that $\rho(r)$ can be expressed as a contribution of each electron present in the molecule and written as:

$$\rho(r) = \sum_i |\psi_i(r)|^2 \qquad \textbf{Eq. (4.27)}$$

where ψ_i is called a Kohn-Sham orbital and the density of the system is the sum of the square moduli of a set of one-electron orthonormal orbitals. T is also unknown but can be approximated if the density is associated with a wavefunction. Kohn and Sham introduced the orbital concept into DFT

$$T[\rho] = \sum_i \int \psi_i(r) \left(-\frac{\nabla^2}{2}\right) \psi_i(r) dr \qquad \textbf{Eq. (4.28)}$$

We need to define K_{XC} which can be approximated in various different ways exactly the same ansatz is used as Hartree-Fock, the only difference is in the Fockian operator

$$\hat{F}_{HF} = \sum_i \hat{h}(i) + \left(\sum_{j\epsilon occ} 2\hat{J}_j - \hat{K}_j\right)$$

$$\hat{F}_{KS} = \sum_i \hat{h}(i) + \left(\sum_{j\epsilon occ} 2\hat{J}_j\right) + \hat{E}_{xc} \qquad \textbf{Eq. (4.29)}$$

Where the first term of each of the above equations represents the one-electron operator for kinetic energy for each electron and electron-nuclear interaction. As with Hartree-Fock the Kohn-Sham orbitals are expressed as linear combination of atomic-centred basis functions, solved iteratively and self consistently (SCF method):

$$\psi_i(r) = \sum_v c_{vi}\phi_v \qquad \textbf{Eq. (4.30)}$$

The $E_{XC}[\rho]$ can be written as

$$E_{XC}[\rho] = E_X[\rho] + E_{corr}[\rho] \qquad \textbf{Eq. (4.31)}$$

where $E_X[\rho]$ represents the exchange energy density and $E_{corr}[\rho]$ the correlation energy density functional.

4.6 - Conclusion

Computational chemistry is an extremely powerful tool in the design of potential drugs and an extremely valuable tool in modern day drug discovery efforts throughout medicine. Its effectiveness to narrow down a library of ligands which have potential activity on the GP enzyme has already been demonstrated with a number of computationally driven success stories [43, 121]. Using these methods, we can find a few select ligands which are likely to bind at the inhibitor site, without the lengthy and expensive process of synthesising inactive compounds and then physically testing them. And with advances in computer technology, more powerful and lengthy computational methods can be used, for more accurate binding predictions reducing the chances of false positive predictions.

Chapter 5 –Synthetic Methods

5.1 - Introduction

Type-2 Diabetes (T2D) has become a considerable threat to human health on a global scale, and a massive social and economic burden [122]. Multiple problems arise from T2D, including an increase in the risk of patients developing other health concerns which would otherwise be avoidable, such as angina, cardiac failure, strokes, retinopathy, renal replacement therapy and amputations [14]. Diabetes mellitus is characterised by hyperglycemia (high levels of glucose in the bloodstream), peripheral insulin resistance and low insulin production in the pancreas [123]. Glycogen Phosphorylase (GP) has been identified as a key enzyme in T2D, as it catalyses the first step in the intracellular degradation of glycogen to give α-D-glucose-1-phosphate (G-1-P), via the glycogenolysis pathway, as shown in Figure 5.1. This function makes it an extremely promising target for T2D treatment, as inhibiting GP would reduce glucose production, which would reduce the level of hyperglycemia. It is important to note however, that inhibition of GP needs to be temporary, to avoid causing hypoglycaemia in the patient (too little glucose in the blood).
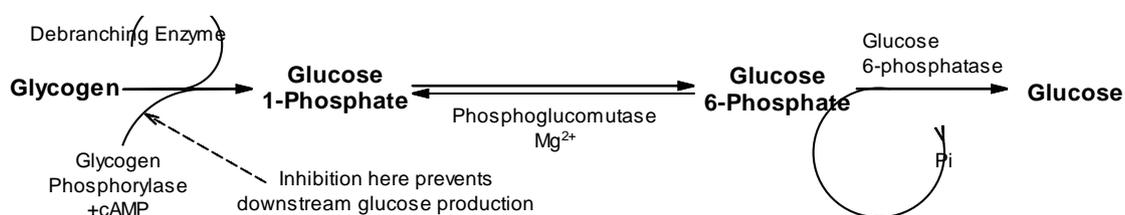
Figure 5.1 – The glycogenolysis pathway.[33]
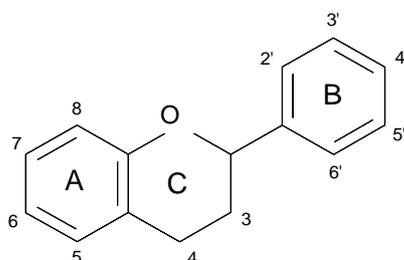
Figure 5.2 – Basic structure of flavonoids showing ring and atom labelling.

Flavonoids are polyphenols which are found in nature, which leads to them being consumed daily as part of a healthy diet of fruit and vegetables. They are also known for their other health benefits, such as antioxidant, antibacterial, anticancer, anti-inflammatory and antidiabetic properties, as well as protection against cardiovascular disease.[59-61, 124] They have a general scaffold of two aromatic rings (A and B) which are joined by a benzo-fused dihydropyran (ring C). The flavone structure, which is the focus of this project, has two hydroxy groups on the A ring in the 5 and 7 position, a ketone at the 4 position and a double bond between the 2 and 3 carbons (as seen in Fig 5.3)
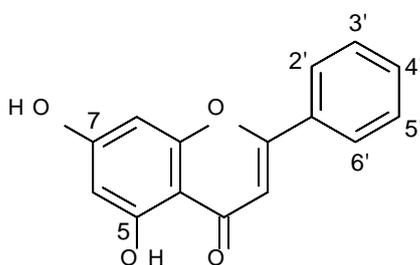
Figure 5.3 – Structure of chrysin

Flavonoids, including chrysin, have shown promise in T2D, by inhibiting GP via binding at the caffeine binding site, which is found on the surface of GP, 12 Å from the catalytic site. This is due to the two aromatic residues found within the inhibitor site: Phe285 and Tyr613, which can interact with the A & C rings of the flavone skeleton via π-stacking interactions. Binding at this site causes inhibition as the T-state conformation is stabilised which blocks access to the catalytic site (*c.f.* Chapter 2)

This phenolic structure can be synthesised in a number of ways, whilst keeping the desired hydroxyl groups on the 5 and 7 position intact, most of which, involve starting with the A and B rings, joined by a carbon chain, which is then cyclised into the C ring. Examples are given in the following section.

5.2 – Synthetic Methods.

5.2.1 – Baker-Venkataraman Rearrangement
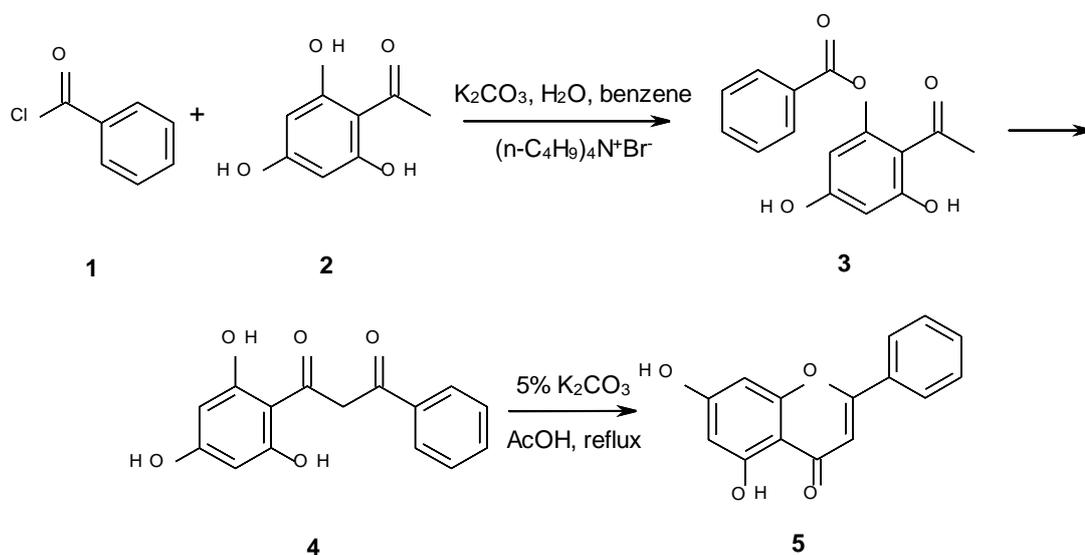
Scheme 5.1 - Synthesis of chrysin via the Baker-Venkataraman rearrangement (step 3-4).

Flavone derivatives can be synthesised from commercially available benzoyl chlorides and 2-acetylphenols, by exploiting the Baker-Venkataraman rearrangement [125]. Chrysin can be made via this route by using 2,4,6-hydroxyacetophenone (**2**) and benzoyl chloride (**1**) as starting materials. These have been experimentally observed to react under multiple conditions, where Zhiwei Chen *et al* [125] at Zhejiang University have observed a reaction when combining the two reactants with K_2CO_3, H_2O and benzene with $(n-C_4H_9)_4N^+Br^-$ as a phase-transfer catalyst stirring at reflux. The use of benzene is potentially dangerous, however, but creates a biphase layer, which requires the $(n-C_4H_9)_4N^+Br^-$ to act as a phase-transfer catalyst which needs removal afterwards. This can be avoided by the use of dry pyridine as a solvent and base, as observed by Bhawna Yvas *et al* [126] at Punjabi University, which is stirred at reflux then poured onto ice with hydrochloric acid to yield the solid product. The process can also be done in acetone, with the K_2CO_3 as base under reflux, as proven by Xing Zheng *et al* [127] at the University of South China. All three of these methods require basic conditions for the condensation to be achievable, as a base is needed for the formation of the enolate intermediate.

An intermediate product from this reaction is an ester, which can undergo the Baker-Venkataraman rearrangement to form the diketone as seen in Scheme 5.2.
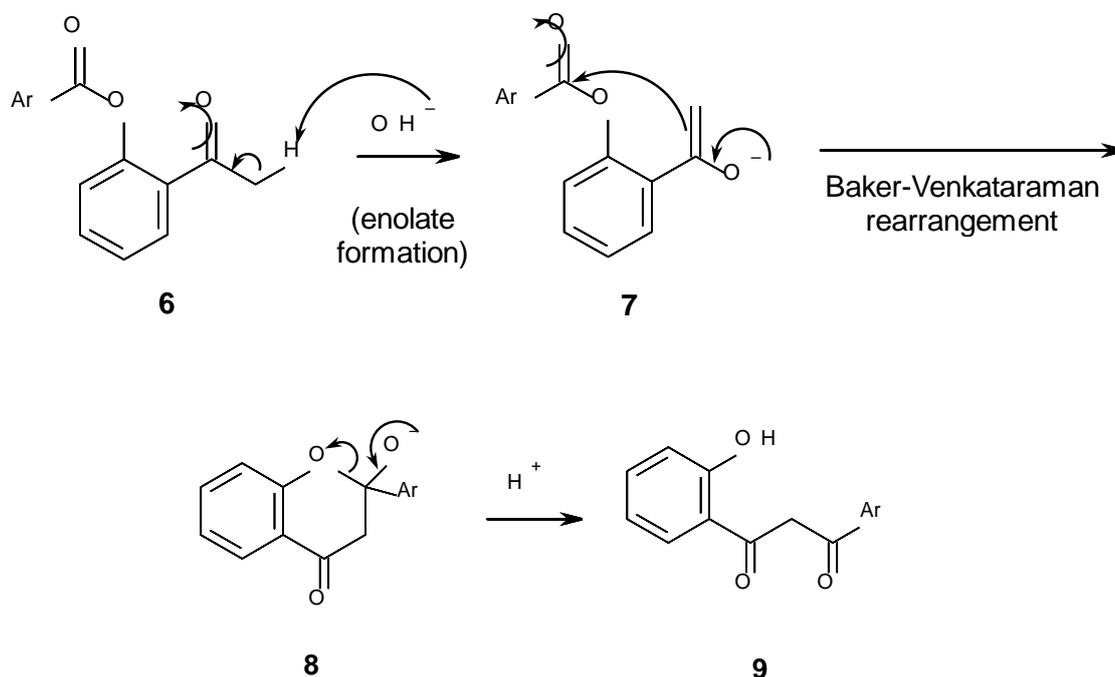
Scheme 5.2 – Baker-Venkataraman rearrangement of (**7**) to form the diketone intermediate, (**9**).

Cyclisation of the diketone to yield the flavone structure can be achieved by a variety of methods, for example, Zhiwei Chen *et al* [125] found that heating the diketone at reflux in 5% aqueous potassium carbonate for 6 hours then treating with acetic acid closed the C ring to form the flavone. Whereas Bhawna Yvas *et al* [126] did not use basic conditions but used strong acidic conditions for ring closure. Bhawna Yvas *et al* [126] used glacial acetic acid and concentrated sulphuric acid at reflux for 1 hour, which was then poured onto crushed ice to precipitate the flavone product. However, this process may be problematic due to the dangerous nature of concentrated sulphuric acid, so, Xing Zheng *et al* [127] overcame this by completing this reaction with the same glacial acetic acid, but instead mixed with anhydrous sodium acetate.

5.2.2 – Microwave Synthesis

George W. Kabalka *et al* [128] took the diketone structure (**9**) and synthesised a variety of flavones by using microwave assisted methods. This was achieved by adding 1.0 mmol of the diketone compound and 0.1 mmol of $CuCl_2$ in 3 ml of ethanol, then

subjecting the mixture to microwave at 80°C at 100W for 5 minutes. This mixture was extracted and purified by flash column chromatography to yield flavones in up to 98% yield. Julio A. Seijas *et al* [129] also created the flavone structure via the diketone pathway by using microwave techniques. However, this was achieved without the use of solvent. This varies slightly to the techniques discussed earlier as this method uses ethyl benzoylacetate (Fig 5.4) with 1,3,5-trihydroxybenzene, rather than the benzoyl chloride and 2-acetylphenol method in Scheme 5.1. However, without solvent, this meant that the microwave was set to 240°C at 800W, which is substantially more powerful than that used by George W. Kabalka *et al.*
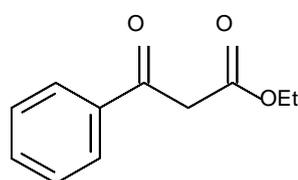
10

Figure 5.4 – Structure of ethyl benzoylacetate.

10   **11**   **5**
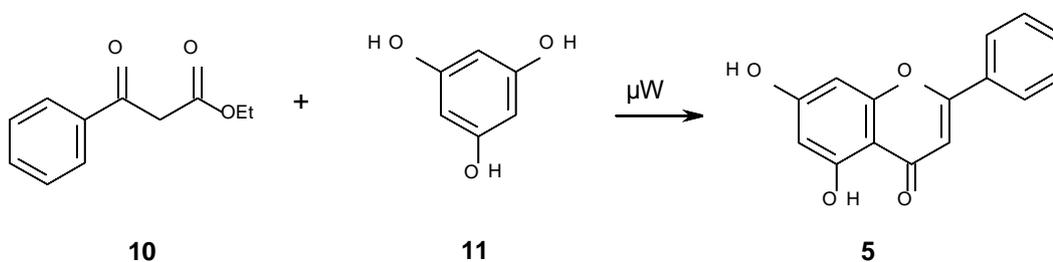
Scheme 5.3 – Microwave assisted synthesis of chrysin in solvent free conditions.

Though a mechanism was not provided, it is assumed that the mechanism follows a path similar to Scheme 5.4.
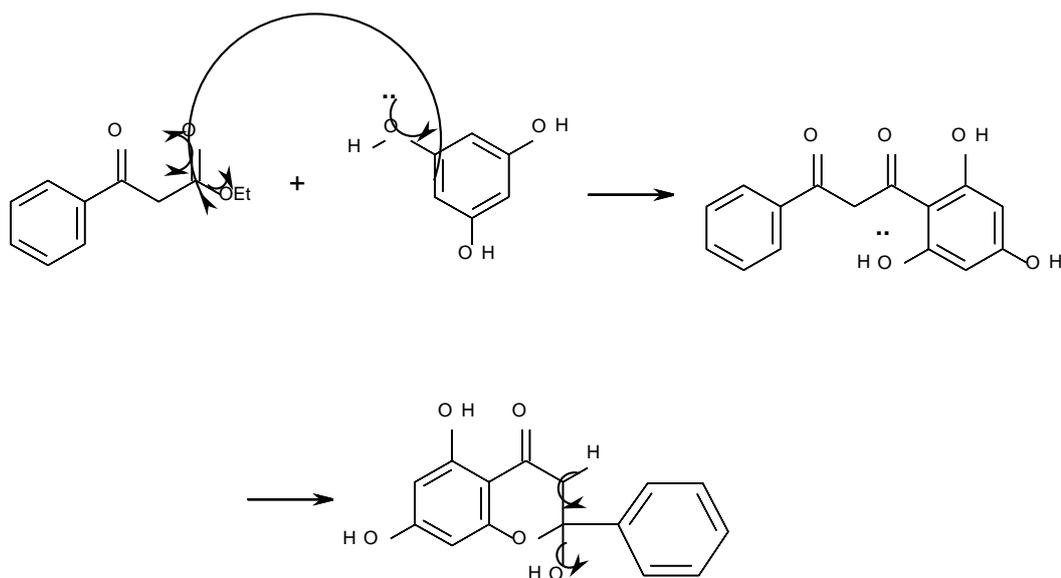
Scheme 5.4 – Proposed mechanism for the microwave assisted synthesis of chrysin.

This has an advantage over traditional 'wet' methods, as there are no harmful solvents, acids, or bases used during the synthesis. Julio A. Seijas *et al* [129] also recorded extremely high yields using this method, of 96% when synthesising chrysin. Microwave synthesis also has the added advantage of taking much less time in comparison to traditional methods, the reaction mentioned above (Scheme 5.3) requires 3 minutes of irradiation time, whereas wet methods can take longer than 24 hours. This in turn, means less energy is needed to power apparatus for long periods of time.

5.2.3 – Single Ketone Synthesis

Many reaction pathways in the synthesis of flavones form a single ketone as an intermediate, which are derivatives of 1-(2-hydroxyphenyl)-3-phenyl-2-propenone (Fig 5.5) or 2′-hydroxychalcone (Fig 5.6). These structures can undergo the cyclocondensation required to form the C ring (*c.f.* Fig 5.2) to create the desired flavone structure.
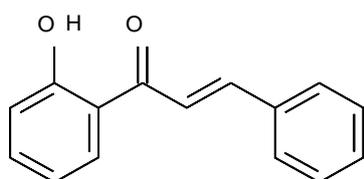
Figure 5.5 – Structure of 1-(2-hydroxyphenyl)-3-phenyl-2-propenone.
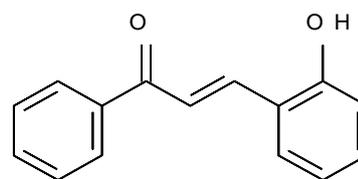
Figure 5.6 – Structure of 2′-hydroxychalcone.

Both of these can be synthesised under basic conditions, 1-(2-hydroxyphenyl)-3-phenyl-2-propenone was synthesised by Abhay S. Zambare *et al* [130] using sodium hydroxide in ethanol at 15-20°C to undergo an aldol condensation, as seen in Scheme 5.5.
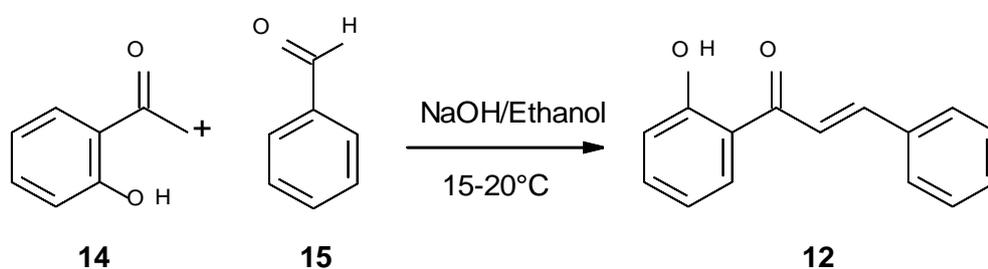
Scheme 5.5 – Synthesis of 1-(2-hydroxyphenyl)-3-phenyl-2-propenone.

Koneni V. Sashidhara *et al* [131] synthesised 2'-hydroxychalcone by reacting salicylaldehyde (**16**) and acetophenone (**17**) together with 10% aq. KOH and ethanol at reflux (see Scheme 5.5) via a variant of the aldol reaction, the Claisen-Schmidt condensation.
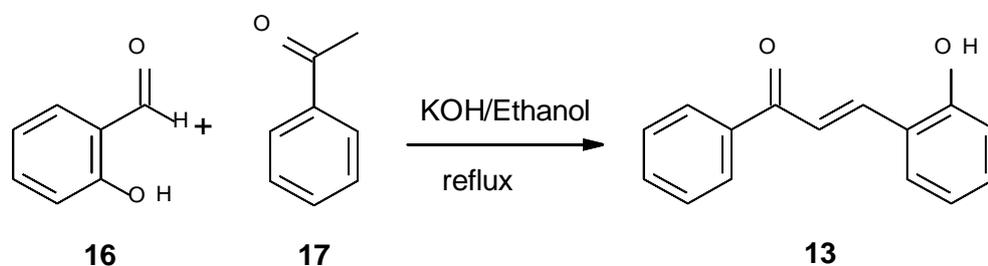
Scheme 5.6 – Synthesis of 2'-hydroxychalcone.

The use of 1-(2-hydroxyphenyl)-3-phenyl-2-propenone has been investigated in the synthesis of flavones, with positive results in many different systems. Surender Kumar *et al* [132] cyclised 1-(2-hydroxyphenyl)-3-phenyl-2-propenone to form flavone in DMSO and sodium tellurite (Na_2TeO_3) at 130-140°C with yields of 80% (Scheme 5.6)
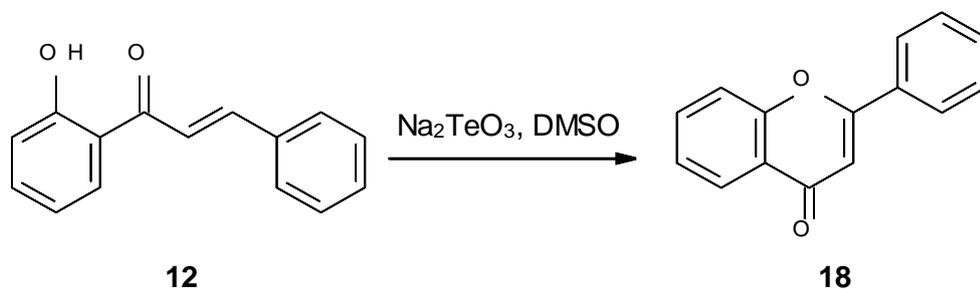
Scheme 5.7 – Synthesis of flavone from the cyclocondensation of 1-(2-hydroxyphenyl)-3-phenyl-2-propenone using sodium tellurite in DMSO.

Kalvi Hemanth Kumar *et al* [133] found that using iron(III) chloride hexahydrate ($FeCl_3.6H_2O$) in methanol at reflux was also a viable method of cyclocondensing 1-(2-hydroxyphenyl)-3-phenyl-2-propenone to form flavone (Fig 5.7), this method was also found to be effective in ethanol, but unsuccessful in non-alcoholic solvents (DMF,DMSO, ethyl acetate, THF and toluene). This method did give lower yields in comparison to previously mentioned methods; $FeCl_3.6H_2O$ in methanol gave only 55% yields, and also took 10 hours for completion, which is longer than the previously discussed microwave methods.
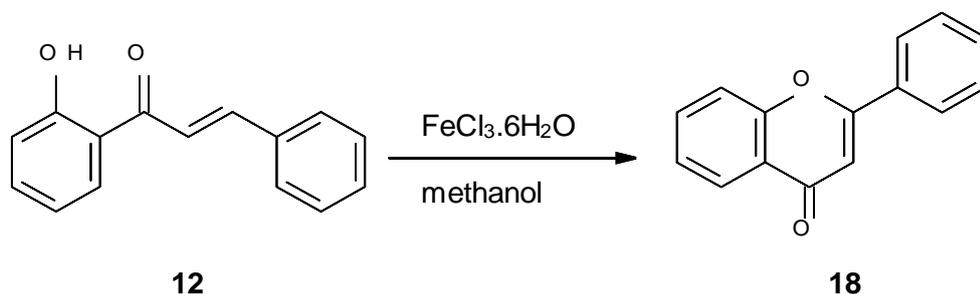
Scheme 5.8 - Synthesis of flavone from the cyclocondensation of 1-(2-hydroxyphenyl)-3-phenyl-2-propenone using iron(III) chloride hexahydrate.

Ethanol has been used repeatedly for the cyclocondensation step for the formation of the C ring; Abhay S. Zambare *et al* [130] used 10 mol% oxalic acid in ethanol to give the cyclised flavone with excellent yields of 95% over 6 hours. Alcoholic solvents were

also tested and methanol gave yields of 92% over 6.5 hours, which backs up the previously mentioned solvent choice [133]. However, unlike the iron(III) chloride hexahydrate route, using oxalic acid as a catalyst allowed a wider range of solvents to be used, with acetonitrile over 12 hours giving 85% yield and THF over 10 hours giving 88% yield. This makes the oxalic method far more favourable than the iron(III) chloride hexahydrate method. However, this is still unfavourable in comparison to the microwave assisted methods mentioned earlier.

Abu T. Khan *et al* [134] also managed to achieve the cyclised product by exploiting ammonium bromide and vanadium pentoxide alongside hydrogen peroxide (Scheme 5.9)
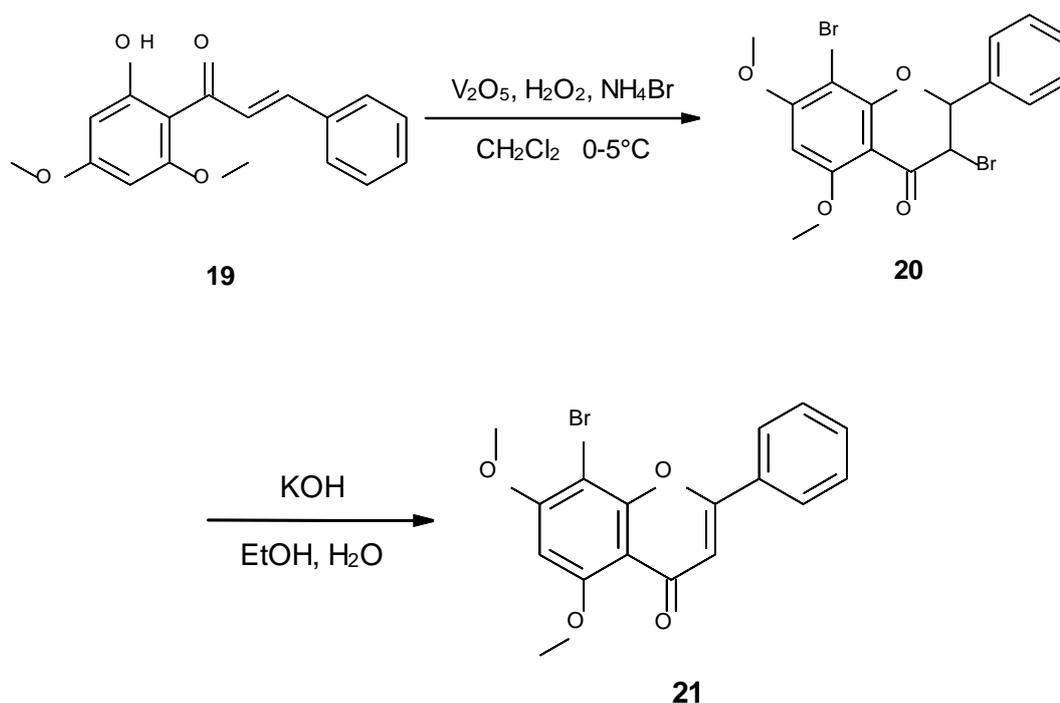
Scheme 5.9 – Synthesis of 8-bromochrysin derivative from 2'-hydroxy-4',6'-dimethoxychalcone.

Unfortunately, this method only gave an overall yield of 62%; which means this method is not optimal, as the yields are poor in comparison to previously mentioned methods, and also introduces an undesired bromine at the 8-position, which would need further

work to remove, such as exploiting a Pt(III) surface for hydrodebromination methods.[135] The methyl groups on the hydroxyl groups would also need to be removed via de-methylation techniques, such as HBr with AcOH at reflux for 24 hours as reported by Chavi Yenjai *et al* [136] in which they achieved a 95% yield.

Koneni V. Sashidhara *et al* [131] investigated the cyclocondensation of 2′-hydroxychalcone with the use of different catalysts (Table 5.1) and solvents (Table 5.2).
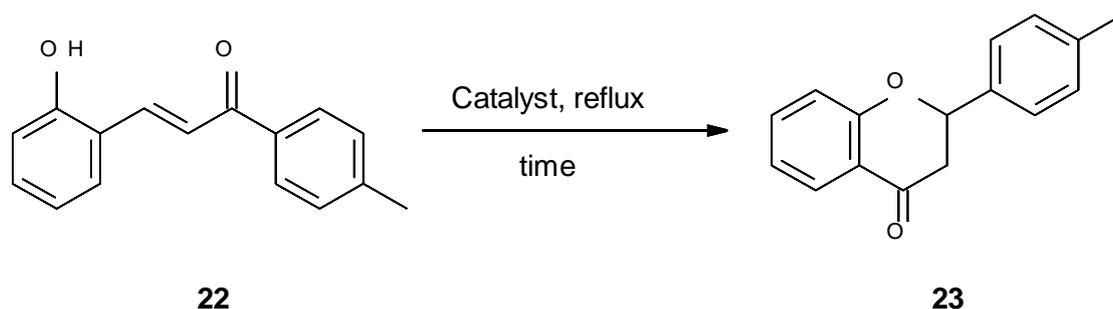
Scheme 5.10 – Synthesis of 2-(4-methylphenyl)-4H-chromen-4-one.

Table 5.1 – Optimisation of reaction conditions for the synthesis of 2-(4-methylphenyl)-4H-chromen-4-one.

Entry	Catalyst	Time (h)	Yield (%)
1	$AlCl_3$	2.5	28
2	$ZnCl_2$	3.0	10
3	$BF_3.OEt_2$	3.0	0
4	$SnCl_2.2H_2O$	2.0	54
5	$HgCl_2$	3.0	5
6	$FeCl_3$	3.0	5
7	$SnCl_4$	3.0	0
8	I_2	1.5	74

It was reported that iodine gave the best results under these conditions, they also reported that the use of 10 mol % of iodine was optimal. More surprisingly however, is

the use of solvent; as no solvent achieved yields close to performing the synthesis neat, as seen in Table 5.2.

Table 5.2 – Optimisation of reaction solvents for the synthesis of 2-(4-methylphenyl)-4H-chromen-4-one.

Entry	Solvent	Yield
1	1,4-Dioxane	56
2	Acetonitrile	0
3	Ethanol	33
4	Methanol	12
5	THF	15
6	DMSO	20
7	Neat	74

Using iodine as a catalyst with no solvent at 110-130°C for 1.5h gives a 74% yield for the 2-(4-methylphenyl)-4H-chromen-4-one, and also gives a 72% yield for flavone. Sashidhara [131] suggested a plausible mechanism, as shown in Figure 5.10:
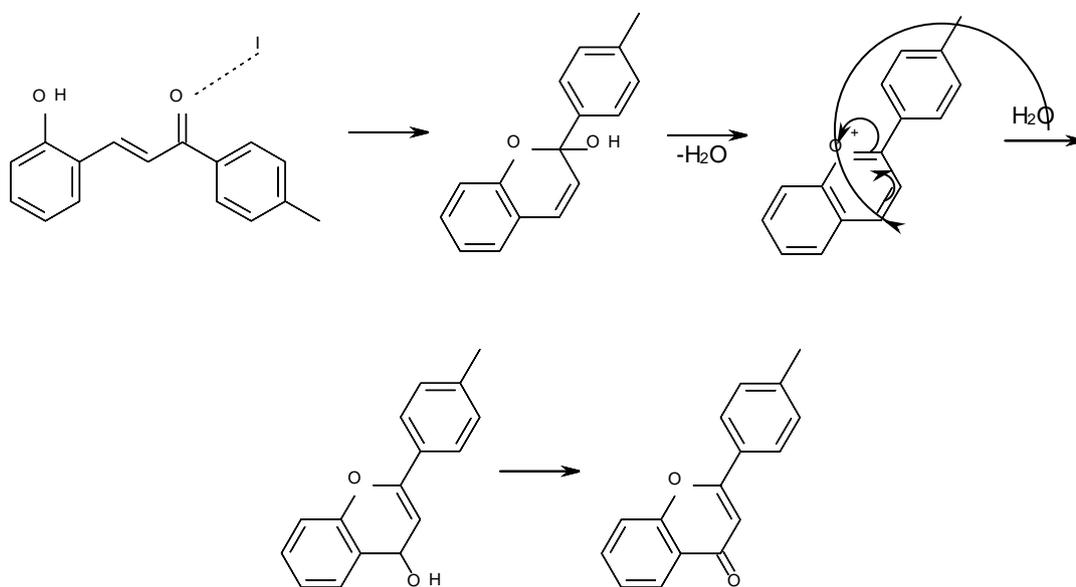
Figure 5.7 – Proposed mechanism of reaction of flavones using iodine as a catalyst.

5.2.4 – Synthesis Using a Metallic Catalyst

An alternative method involves the use of palladium-based catalyst to synthesise the flavone compound. Hua Miao *et al* [137] at Harvard University achieved high yields under basic conditions under a carbon dioxide atmosphere, as shown in Scheme 5.11.
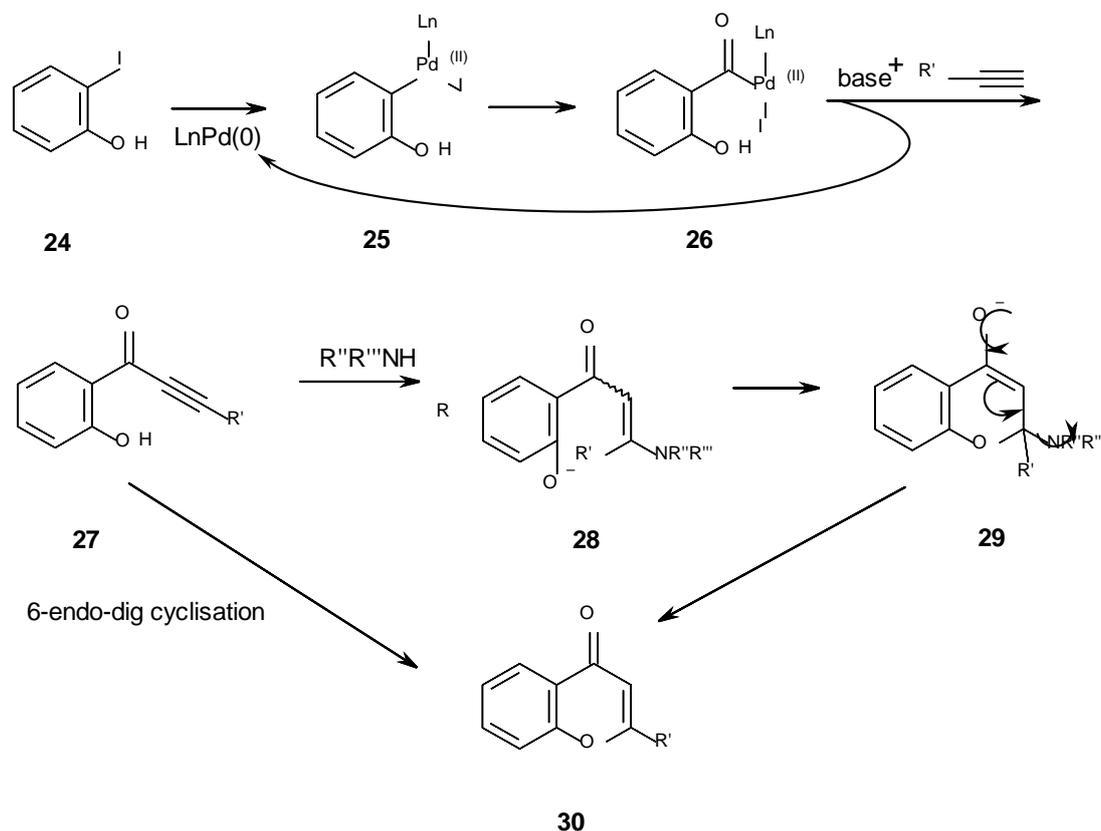
Scheme 5.11 - Mechanistic interpretation for the synthesis of flavones using a palladium based catalyst.

Unlike other methods mentioned, this approach starts with iodophenol (**24**), which undergoes oxidative addition to form the organo metallic intermediate (**25**), then the carbon monoxide is inserted to give the carbon monoxide insertion product (**26**). With base and the desired R group attached to an alkyne carbon-carbon bond (**27**). For the synthesis of flavones, the R group needs to be aromatic. Compound (**27**) can then undergo 6-endo-dig cyclisation to directly form the desired flavone (**30**). Alternatively, (**27**) can undergo a Michael addition, to attach the NR"R''' diethylamine group, forming the alkene compound (**28**) which can undergo 6-endo-trig cyclisation (**29**) which

spontaneously rearranges to form the flavone (**30**). There is a major drawback with this method, as compound (**27**) can rearrange to form a palladium compound. This regioselective problem has been resolved by Torii and Kalinin by using an excess of diethylamine to competitively form compound (**28**). However, the conditions required are 120°C and under pressure of 20kg/cm^2, which is impractical compared to other methods discussed previously.

5.3 - Conclusion

T2D is a chronic and widespread illness, affecting millions of patients. With the promising approach of inhibition of the GP enzyme using flavone derivatives, better treatment for T2D is achievable. The synthetic methods presented in this chapter are all viable options for the synthesis of flavones, which in turn means they can be considered for the synthesis of chrysin and it's analogues. Due to it's synergy with the computational methods and using benzoyl chlorides with the Zinc Docking Database, and also due to the previous expertise within the group, to follow the synthetic route via the Baker-Venkataraman rearrangement, via the 2,4,6-hydroxyacetophenone and the benzoyl chloride. The synthesised analogues can then be taken forward to kinetics experiments, to test whether or not they have activity against GP. These results will help influence the next step, as structure activity relationships will be performed, comparing the computational results to the kinetic experiments. These results will then be used to refine and suggest further more ligands, for synthesis aiming for greater activity.

Chapter 6 - Computational Screening for New Flavonoid Derivatives Targeting the Inhibitor Site

6.1 - Introduction

Type-2 Diabetes (T2D) is a chronic heterogeneous disease which is characterised by hyperglycaemia. T2D is caused by a disorder of insulin secretion, insulin resistance, and by the production of hepatic glucose in the liver, of which, 70% comes from the breakdown of glycogen by the GP enzyme in the glycogenolysis pathway.[138] This means that GP is an extremely promising target for inhibition in the treatment of T2D. GP has several different binding sites: the allosteric, the catalytic, the inhibitor, the new allosteric, the quercetin binding and the glycogen storage site. The inhibitor site is the main focus in this study, and how flavone derivatives can bind to the site by exploiting the π-stacking interactions. The inhibitor site of GP is found on the surface of the enzyme, roughly 12 Å from the catalytic site. When the enzyme is in the less active T state, Phe285, which is on the 280s loop (residues 282-287) is stacked close to Tyr613, which is on the α19 helix (residues 613-631). These two aromatic residues form the core of the inhibitor site. The ligands that bind to this site are heavily influenced by these residues, this is because of π-stacking interactions between the Phe285 and the Tyr613 sidechains, forming a sandwich-like complex. This promotes the T state conformation of the enzyme via stabilisation of the closed position of the 280s loop, which in turn blocks access to the catalytic site. When the enzyme is converted from the T state to the R state, the 280s loop changes its shape, which opens a channel that allows Arg569 to enter the catalytic site in place of Asp283, creating the recognition site for the substrate phosphate. It is this transition which enables the glycogen substrate to reach the catalytic site [40]. Inhibition of the binding site is usually synergistic with glucose, which suggests that inhibition could be regulated by blood glucose levels, and would decrease as normoglycaemia is achieved without causing hypoglycaemia [139]. Because of this property, design of inhibitors targeting the "inhibitor site" is particularly attractive from a drug design perspective. Additionally, as we have discussed previously (Chapter 3), natural products such as flavonoids provide a proven scaffold for design of potent GP inhibitors but to date the only flavonoid ligands confirmed to bind at the site are chrysin (by crystallography, PDB code 3EBO) and quercetagetin (multiple inhibition studies).
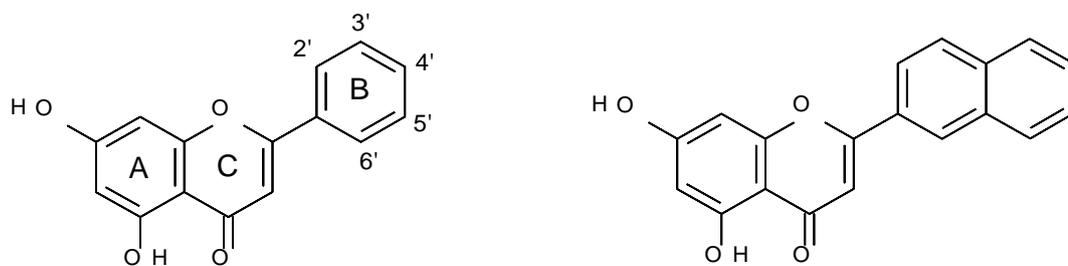
Figure 6.1 - Structure of chrysin and 5,7-dihydroxy-2-(naphthalen-2-yl)-4H-chromen-4-one

In this chapter, we present the virtual screening of a library of compounds exploiting the Zinc database (http://zinc.docking.org) [112] using 5,7-dihydroxy-4H-chromen-4-one as a core scaffold (Figure 6.1). Initial screening was performed using both Glide-SP and –XP but combining the results in a consensus scoring approach so as to potentially reduce false positive results. Two GP protein conformations were used for this purpose, one taking the solved GP protein conformation directly from its complex with chrysin (PDB code 3EBO) and the other from an induced fit docking of 5,7-dihydroxy-2-(naphthalen-2-yl)-4H-chromen-4-one (Figure 6.1) to 3EBO so as to expand the hydrophobic pocket (key residues Leu380, Tyr573 and Phe771; *c.f.* Figure 3.3 in Chapter 3) to allow binding of ligands with larger B ring substituents extending into the hydrophobic pocket. QM/MM-PBSA binding free energy calculations were then performed on the predicted best inhibitors that were also considered to be synthetically viable. Filters were employed so that the final most accurate QM/MM-PBSA predictions only considered the best possible ligands. The whole process is demonstrated in the flowchart shown in Figure 6.3 (on page 71).

6.2 - Computational Details

6.2.1 - Protein Preparation

The GPb receptor (PDB code: 3EBO) was prepared for calculations using Schrödinger's Protein Preparation Wizard [140]. This receptor was chosen as it is the solved structure with the lead compound chrysin bound at the inhibitor site. The waters within 5 Å of chrysin were retained, bond orders assigned and hydrogens added, with protonation states for basic and acidic residues based on pK_a's at a pH of 7. Finally, the system was

softly minimised using OPLS-AA (2005) forcefield with the root-mean-square deviation (RMSD) of heavy atoms kept within 0.3 Å of the crystallographic positions.

6.2.2 - Ligand Database Preparation

Using the Zinc docking database (http://zinc.docking.org/), benzoyl chloride derivatives were selected using criteria so that the final compounds would follow Lipinski's 'rule of five' [141] for oral bioavailability. To ensure this, the search for benzyl chlorides was given the following search criteria: $0 \leq MW \leq 386$ amu, ≤ 3 hydrogen bond donors and ≤ 7 hydrogen bond acceptors. There was also a limit on the number of rotatable bonds of the downloaded compounds to < 3 (5 rotatable bonds in total for final compounds limiting flexibility) and the formal charge to be 0 (to make the final ligand more suitable for synthesis). These restrictions gave 1169 benzoyl chloride derivatives. These were joined to the 5,7-dihydroxy-4H-chromen-4-one core using CombiGlide 3.9,[140] which exploits the program LigPrep, as shown in Figure 6.2. This process yielded 1239 minimised structures including tautomers (occurring only for the B ring), which were now ready to be employed in the docking studies.

5,7-dihydroxy-4H-chromen-4-one core

Benzoyl chloride substructure
(1169)

CombiGlide
↓

1239 Ligands
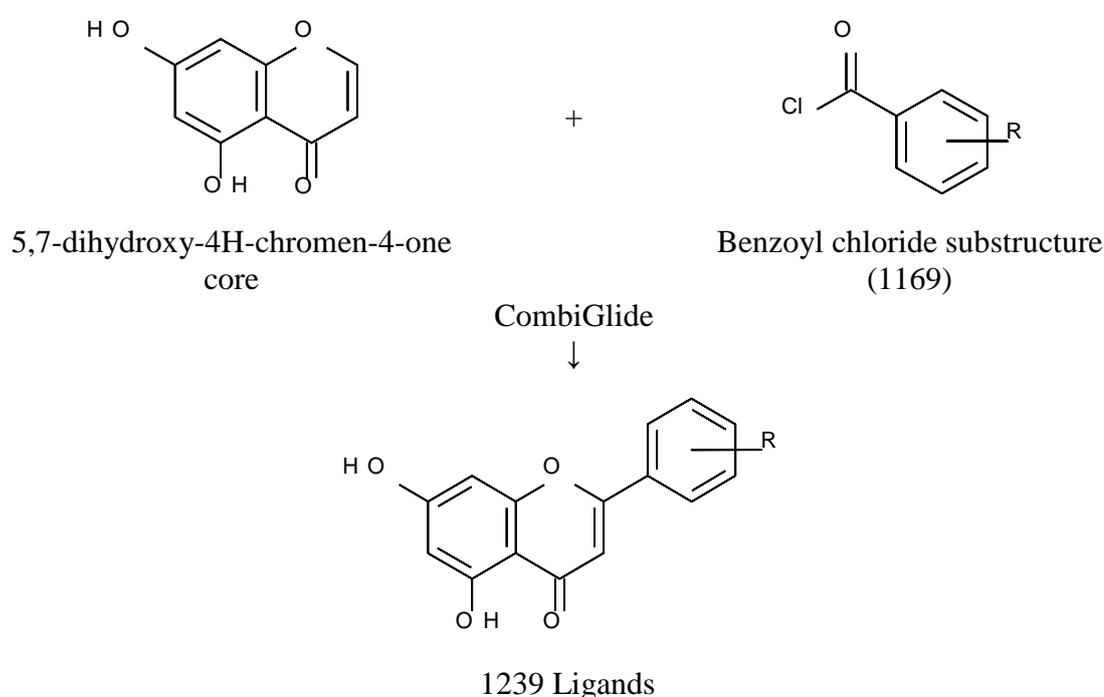
Figure 6.2 – The chrysin core and the benzyl chloride substructure used for searching the ZINC docking database and then being bound together with CombiGlide to form the final ligands for screening.

6.2.3. Docking

Docking of all ligand was performed to the "Original GPb Receptor" conformation from it complex with chrysin (*c.f.* 6.2.1) and also to an "Induced-fit GPb Receptor" GP conformation prepared so as to allow the binding of bulkier ligands.

Docking to Original GPb Receptor

In the docking calculations with Glide 6.9 [140, 142], the shape and properties of the inhibitor binding site were mapped onto grids with dimensions of 23.1 x 23.1 x 23.1 Å centred on the native chrysin ligand. Standard parameters were applied including van der Waals scaling for non-polar atoms (by 0.8) to include modest 'induced-fit' effects. Core constraints (maximum RMSD of 0.10 Å for core atoms positions) were applied on the A and C ring atoms (*c.f.* Figure 6.1) as well as 5 and 7 hydroxyl atoms. Docking calculations were performed using both Glide standard-precision (SP) and extra-precision (XP), both of which included post-docking minimization with strain correction. Intramolecular hydrogen bonds were also rewarded. 1 pose per ligand (each tautomer) was saved for both docking calculations. This Glide-SP and -XP docking produced 457 and 499 protein-ligand poses, respectively.

Docking to Induced-fit GPb Receptor

To expand the inhibitor site's hydrophobic pocket, induced fit docking (IFD) was initially performed. Using the 5,7-dihydroxy-2-(naphthalen-2-yl)-4H-chromen-4-one (Fig 6.1) as the ligand, a model for GP was obtained that would be able to accommodate larger substituents in the hydrophobic pocket in rigid-receptor docking calculations [143]. In Stage 1 of the IFD, 5,7-dihydroxy-2-(naphthalen-2-yl)-4H-chromen-4-one was docked to the inhibitor site using Glide-SP with 20 poses per ligand saved. The core of the ligand (rings A and C atoms) involved in the π-stacking was restrained to its original position using core constraints (tolerance 0.1 Å). The Leu380 sidechain was trimmed for this stage, as well as scaling of the van der Waals radii of non-polar atoms by 0.5. Stage 2 was a Prime refinement, with residues Phe771 and Leu380 sidechain positions in the hydrophobic pocket refined. For Stage III, up to 20 protein conformations within 30 kcal/mol of the lowest-energy structure were used for Glide-SP re-docking.

Using the top-ranked IFD pose based on IFDscore, the original 1169 ligands from the CombiGlide process, were redocked into the new induced-fit receptor using Glide-SP and Glide-XP using the exact same settings as for the "Original GP Receptor" and which gave 740 and 790 poses, respectively.

Consensus Scoring

Ligands were ranked using Consensus Scoring with a Simple Sum Rank of the Glide-SP and –XP ranks (based on GlideScore) for each receptor, "Original GP Receptor" and "Induced-fit GPb Receptor" (Eq. 6.1). In the case of ligands with tautomers only the best scoring tautomer was ranked in each case so that each ligand received only one rank. This gave more reliable ranks, by taking all calculations into account [144, 145].

$$simple \; sum \; rank = (Glide - SP \; Score) + (Glide - XP \; score) \qquad \textbf{Eq. (6.1)}$$

We then compared the RMSDs of all atoms for ligands that gave poses for both SP and XP. This was to see if SP and XP gave similar poses for the same receptor [145]. A RMSD cut-off less than 0.7 Angstroms was considered similar. Additionally, the ligands were manually investigated in detail and based on intuition, compounds that were not feasible with the Baker-Venterman re-arrangement, or if they had bulky or polar para substituents, were deleted, as these would be unfavourable. This process left 373 novel compounds. Based on this, the top 200 ligands for each of the two receptor conformations (with some ligands in top 200 for both receptors) proceeded to more accurate QM/MM-PBSA binding free energy calculations.

6.2.4 - QM/MM-PBSA Calculations

The best 200 SP poses for each receptor were used as input for QM/MM-PBSA binding free energy calculations. QM/MM interaction energies ($\Delta E_{QM/MM}$) were calculated directly for each of the poses, considering the bound and unbound states of the predicted protein-ligand complexes:

$$\Delta E_{QM/MM} = E_{complex} - E_{receptor} - E_{ligand} \qquad \textbf{Eq. (6.2)}$$

Ligand binding free energies were then calculated using QM/MM-PBSA according to the following equation:

$$\Delta G_{bind} = \Delta E_{QM/MM} + \Delta G_{solv} - T\Delta S_{MM}$$ **Eq. (6.3)**

All calculations employed DFT with the M06-2X functional [146] and the all electron LACVP*+ basis set [140] was used for the QM region (ligand, Phe 285 and Tyr 613 sidechains); this method was previously shown to be effective for halogen substituted ligands in studies of this type [39]. The GP enzyme (except for Phe285 and Tyr613 sidechains) was described using MM with the OPLS-AA(2005) forcefield [147]. No cut-off for non-bonded interactions was employed. Effectively, the MM region polarizes the QM region, with electrostatic interactions between QM wavefunction, and the van der Waals interactions between QM and MM atoms accounted for [147]. Bulk solvation effects were included using Poisson-Boltzmann Surface Area (PBSA), allowing ΔG_{solv} to be calculated. The default solute (internal) dielectric constant of 1.0 was used. An estimate for the loss of ligand entropy (ΔS_{MM}) on binding was calculated using MM with the OPLS-AA(2005) forcefield [147] and the Rigid Rotor Harmonic Oscillator (RRHO) approximation. Using this method, the change in vibrational, rotational and translational entropy of the ligands on binding was considered. All QM/MM and QM/MM-PBSA calculations were performed using Qsite 6.9 [140]. RRHO calculations were performed using MacroModel 11.0 [140]. We then compared the ΔG_{bind} values of ligands from both receptors results; the best ΔG_{bind} value was taken as the "true" value and the corresponding receptor-ligand pose considered (corresponding either to "Original GP receptor" or "Induced-fit receptor"). Ligands were ranked by ΔG_{bind} values and those with a value more negative than -11 kcal/mol taken forward (58 ligands) for even more accurate QM/MM-PBSA calculations following a Prime refinement of the corresponding receptor-ligand pose.

6.2.5 - Prime Minimisation

To prepare complexes for the more accurate QM/MM-PBSA calculations, Prime 4.2 [140] minimisation was used to refine the protein-ligand complexes using the preferred GPb conformation (either the "Original GPb receptor" or the "Induced-fit GPb Receptor" model) for each of the 58 ligand candidates. The chosen GPb model was based on the protein conformation which gave the best ΔG_{bind} value in the original QM/MM-PBSA calculations. For this Prime "protein refinement" minimization,

standard settings were employed with the OPLS3 forcefield [148] and VSGB solvation model.[149] Refined residues were the sidechains of amino acids: Asp283, Asn284, Glu382, Leu380, Hie571, Glu572, Tyr573, Arg770 and Phe771, which are all within the hydrophobic pocket and are also in close proximity to bound ligands. While the 5,7-dihydroxy-4H-chromen-4-one ligand core of each ligand was constrained using a force constant of 200 kcal/(mol.Å^2) (maintaining the π-stacking distances with Phe285 and Tyr613), the B ring groups were additionally refined (free) in the hydrophobic cavity with the aforementioned residues.

6.2.6 - Second stage QM/MM-PBSA Calculations

The more accurate QM/MM-PBSA binding free energies of these refined complex poses were determined using the Prime output files as input. Two sets of calculations with different QM regions were run for each refined protein-ligand complex:

1) **post-Prime QM/MM-PBSA(1)** – ligand, Phe285, Tyr613 as QM region
2) **post-Prime QM/MM-PBSA(2)** – ligand, Phe285, Tyr613 and additionally Asp283, Asn284, Glu382, Leu380, Hie571, Glu572, Tyr573, Arg770 and Phe771 as QM region

As before, DFT with the M06-2X functional and the all electron LACVP*+ basis set was used for the QM region, with the rest of the GP receptor was described using MM with the OPLS-AA (2005) forcefield, and ΔG_{bind} values calculated according to Eq. 6.3.
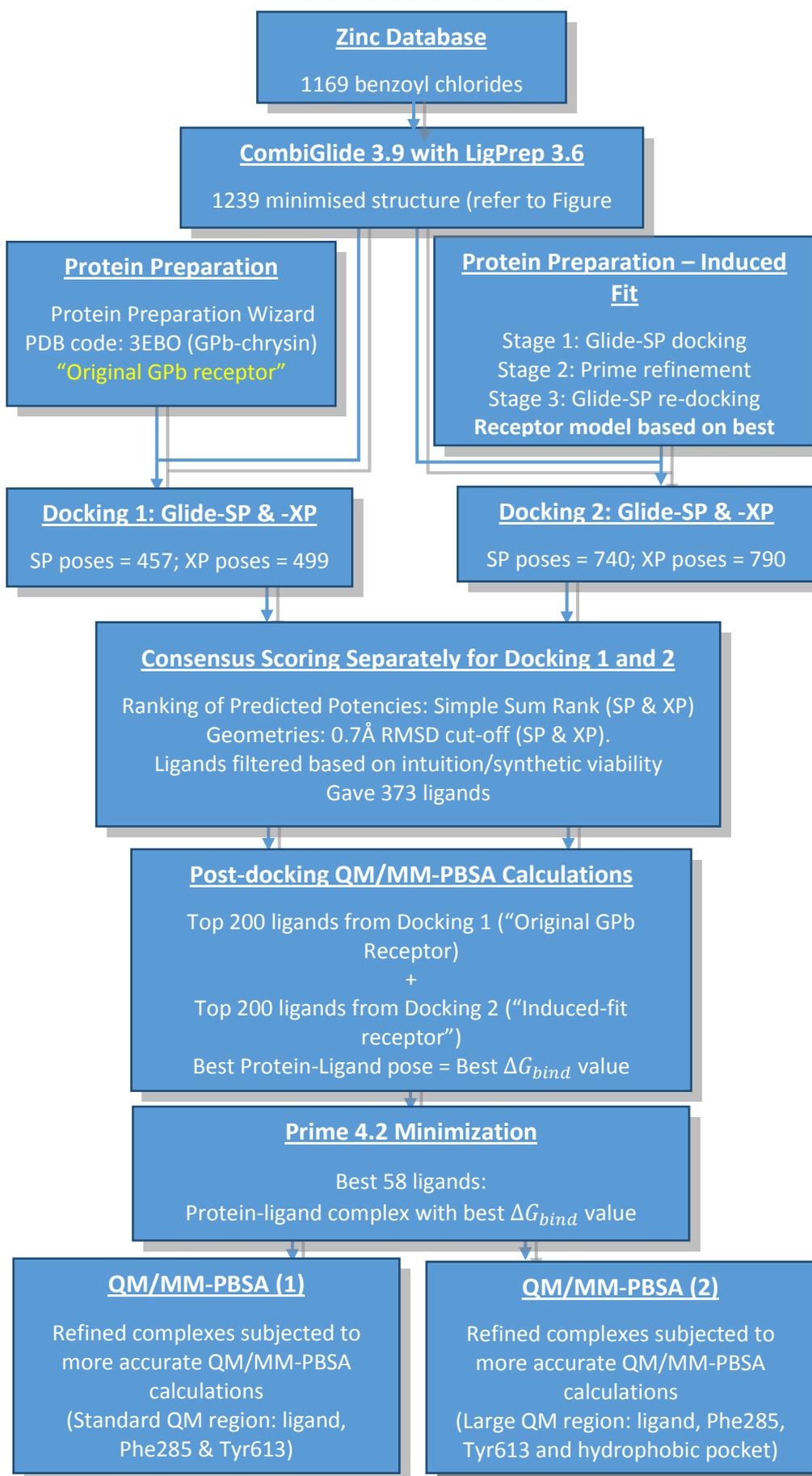
Figure 6.3 – Flowchart representing the computational process for identification of ligand candidates.

6.3 – Results and Discussions

6.3.1 – Docking Results

Docking to "Original GPb Receptor"

The aim of docking was to be able to screen a wide range of compounds to predict their relative binding affinities to the inhibitor site of GPb. Using the Zinc Database and CombiGlide we generated 1239 structures including tautomers, which were docked into the prepared GP protein (PDB code: 3EBO) using Glide –SP and –XP which gave 457 and 499 poses, respectively. The ligands were first ranked by best GlideScore (GS) of Glide-SP then –XP. This was to identify ligands with predicted binding free energies greater than chrysin. GSs revealed that 79 ligands using Glide-SP were better than chrysin and using Glide-XP, 131 ligands were better than chrysin. Then the ranks from Glide-SP and –XP were combined using the Simple Sum Rank (Eq. (6.1)). This consensus scoring was again used to identify ligands with predicted binding free energies greater than chrysin, but considered more accurate due to the known benefits of combining results from different dockings to reduce the chance of false positives [41]. This yielded 79 ligands predicted as more potent than chrysin. The ligands that ranked well (Table 6.1) had halogenated substituents on the B ring, especially fluorine and chlorine at the meta position (3' and 5', refer to Figure 6.1 for numbering Scheme). Hydroxyl groups in the ortho position (2' and 6') also appear to give extremely good binding affinities in these early docking studies.

Table 6.1 – Top ligands by simple sum rank from Glide-SP & -XP results from "Original

Simple Sum Rank: 3	Simple Sum Rank: 6	Simple Sum Rank: 12
ZINC34207034	ZINC39056429	ZINC39056876
GlideScore –SP: -9.311 (1)	GlideScore –SP: -8.889 (2)	GlideScore –SP: -8.756 (3)
GlideScore –XP: 9.590 (2)	GlideScore –XP: -9.168 (4)	GlideScore –XP: -8.610 (9)
Simple Sum Rank: 26	Simple Sum Rank: 29	Simple Sum Rank: 33
ZINC39056879	ZINC38350563	ZINC39057341
GlideScore –SP: -8.549 (11)	GlideScore –SP: -8.730 (4)	GlideScore –SP: -8.447 (19)
GlideScore –XP: -8.494 (15)	GlideScore –XP: -8.373 (25)	GlideScore –XP: -8.526 (14)
Simple Sum Rank: 36	Simple Sum Rank: 47	Simple Sum Rank: 49
ZINC39056727	ZINC34160328	ZINC39056853
GlideScore –SP: -8.372 (28)	GlideScore –SP: -8.705 (6)	GlideScore –SP: -8.314 (42)
GlideScore –XP: -8.643 (8)	GlideScore –XP: -7.847 (41)	GlideScore –XP: -8.749 (7)
Simple Sum Rank:52	Simple Sum Rank:56	Simple Sum Rank: 63
ZINC39056307	ZINC39056495	ZINC39712909
GlideScore –SP: -8.725 (5)	GlideScore –SP: -8.327 (36)	GlideScore –SP: -8.463 (18)
GlideScore –XP: -7.790 (47)	GlideScore –XP: -8.416 (20)	GlideScore –XP: -7.812 (45)
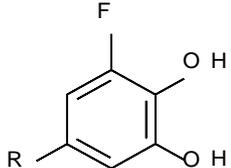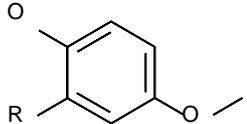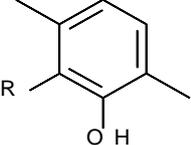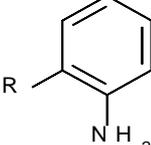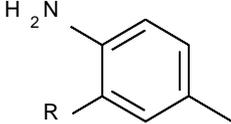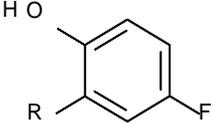
[a] Docking ranks shown in brackets.

Table 6.1(continued) – Top ligands by simple sum rank from Glide-SP & -XP results from

"Original GPb Receptor"[a]

Simple Sum Rank: 66

ZINC34207038
GlideScore –SP: -8.287 (47)
GlideScore –XP: -8.443 (19)

Simple Sum Rank: 67
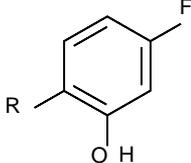
ZINC39056475
GlideScore –SP: -8.491 (13)
GlideScore –XP: -7.696 (54)

Simple Sum Rank: 71
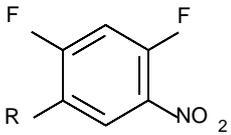
ZINC79387399
GlideScore –SP: -8.587 (10)
GlideScore –XP: -7.596 (61)

Simple Sum Rank: 77
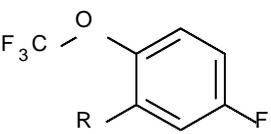
ZINC71867286
GlideScore –SP: -8.391 (25)
GlideScore –XP: -7.729 (52)

Simple Sum Rank: 83
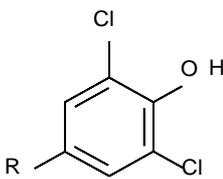
ZINC39375152
GlideScore –SP: -8.346 (32)
GlideScore –XP: -7.749 (51)

Simple Sum Rank: 84
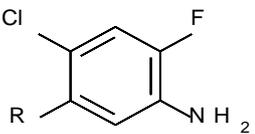
ZINC39057333
GlideScore –SP:
GlideScore –XP:

Simple Sum Rank: 86
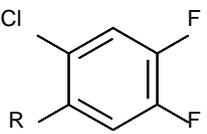
ZINC02540168
GlideScore –SP: -8.443 (21)
GlideScore –XP: -7.574 (65)

Simple Sum Rank: 87
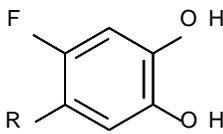
ZINC39056878
GlideScore –SP: -8.181 (82)
GlideScore –XP: -9.117 (5)

Simple Sum Rank:
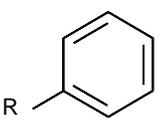
Chrysin
GlideScore –SP: -8.193 (79)
GlideScore –XP: -7.255 (128)

[a] Docking ranks shown in brackets.

Docking to "Induced-fit GPb Receptor"

An induced fit receptor was then created, by relaxing the residues inside the hydrophobic pocket to allow for larger substituents on the B-ring to be able to fit into the cavity. Using the same parameters as the original Glide-SP and –XP, all 1239 ligands were docked into this new receptor. This gave 740 poses and 790 poses for Glide-SP and –XP respectively. The docking GSs revealed that 398 ligands using Glide-SP were better than chrysin and 214 ligands using Glide-XP were better than chrysin. These poses were then rescored using the Simple Sum Rank of Glide-SP and –XP ranks

(Eq 6.1). 257 ligands were found to bind better than chrysin in the "Induced fit receptor". The ligands that gave good poses with Glide-SP and –XP were fluorine and chlorine in the para position (4'), halogens in the meta position (3' and 5') also appear to give greater binding affinities. Similar to the original GPb receptor, hydroxyl groups in the ortho (2' and 6') position seemed to favour binding. As anticipated, the Induced Fit Receptor allowed slightly larger substituents on the B ring to be docked into the hydrophobic pocket, due to the more relaxed open cavity, which is also why the "Induced Fit Receptor" yielded more results than the original. Chrysin, as somewhat expected, was found to go down the ranking for this receptor with the GPb conformation adjusted from the native GPb-chrysin complex conformation of the solved crystallographic complex (PDB code: 3EBO).

6.3.2 - QM/MM-PBSA Results

The best 200 Glide-SP poses for each receptor were taken forward and used as input for preliminary QM/MM-PBSA binding free energy calculations. More accurate QM/MM-PBSA binding free energy calculations were employed on ligands that gave values more negative than -11 kcal/mol in the original QM/MM-PBSA calculation (58 protein-ligand complexes with receptor conformation either from "Original GPb receptor" or "Induced-fit receptor"). Accuracy was improved by using Prime 4.2 minimisation to refine the protein-ligand complexes, so that the complexes were theoretically closer to the "true" complex structures. These QM/MM-PBSA(1) and QM/MM-PBSA(2) results should in theory give more accurate results than the preliminary QM/MM-PBSA calculations on the "Original GPb Receptor" and the "Induced Fit Receptor" due to the relaxed receptor from Prime, as the residues around the ligand were in more optimal positions to the specific ligands. The results from the faster QM/MM-PBSA(1) calculations (smaller QM region) after the Prime minimisation gave more negative ΔG_{bind} values than the original QM/MM-PBSA calculations, whereas the slower QM/MM-PBSA(2) calculations gave in general less negative ΔG_{bind} values.

The first set of QM/MM-PBSA(1) tests (small QM region, results shown in Table 6.2) employed the ligand and the two aromatic π-stacking residues Phe285 and Tyr613 as the QM region, leaving the rest of the receptor in the MM region. This has a greater number of more negative ΔG_{bind} values compared to the slower QM/MM-PBSA(2)

with a larger QM region. The QM/MM-PBSA(1) calculations showed that 36 ligands gave better ΔG_{bind} values than chrysin. Out of these 36 ligands, 26 also performed better in the preliminary QM/MM-PBSA calculations. Of the total 58 ligands, the QM/MM-PBSA(1) calculations did, however, give five ligands with values below the -11 kcal/mol cut off, which was the threshold value used for selection of candidates from the preliminary QM/MM-PBSA calculations. Encouragingly, the overall outcomes does suggest that multiple ligands are more potent than chrysin. From Table 6.1 considering all 58 ligands, generally the $\Delta E_{QM/MM}$ values dominate, with values between -79.51 and -54.46 kcal/mol. The ligand with a value of -79.51 kcal/mol has hydroxyl groups in the 2' and 3' position, whilst the ligand with a value of -54.46 kcal/mol has two fluorines at the ortho positions. This suggests that more polar substituents give better $\Delta E_{QM/MM}$ values. As for halogenated ligands, for example, Zinc code 39712934 (2',6'=F, 4'=Cl) gave a value of -54.50 kcal/mol and ligand 02381665 (2',6'=F, 3'=Cl) gave a value of -54.51 kcal/mol. Whilst total solvation free energies range from 21.23 and 45.98 kcal/mol, the ligand that gave the worst result of 45.98 kcal/mol has two hydroxyl groups on the meta positions of the B ring, whilst the best result of 23.35 kcal/mol has two flourines at the ortho positions, the same which gave the poor $\Delta E_{QM/MM}$ result. This highlights the sensitive balance between $\Delta E_{QM/MM}$ and ΔG_{solv} required to design an effective potent inhibitor. Entropy values range from -15.99 to -13.37 kcal/mol, whilst chrysin has a value of -13.64 kcal/mol, which means that the entropy loss of chrysin upon binding is less than all but one of the ligands, which has a fluorine in the meta position. This low entropy value is a contributing factor to its competitive ΔG_{bind} value. The greatest entropy loss value of -15.99 kcal/mol has hydroxyls on the 2' and 3' positions, which is the same ligand that gave the best $\Delta E_{QM/MM}$ result. This suggests that hydroxyl groups on the B ring have less favourable entropy contributions, as other ligands with hydroxyl groups have also poor TΔS values, whilst the more favourable entropy contributions tend to come from ligands that have only one substituent, of either fluorine or chlorine, the two smaller halogens, especially at the ortho position, such as ligand 01609560 (2'=Cl) with a value of -13.85 kcal/mol and 01699267 (2'=Br) with a value of -13.90 kcal/mol.

Table 6.2: Table illustrating the varied contributions to the terms in Eq. 6.3 to the QM/MM-PBSA binding free energies for four different ligands.[a]
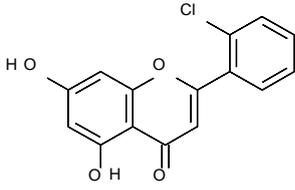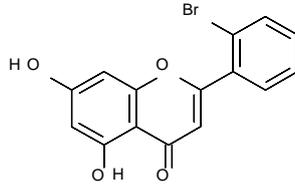
Structure		
Zinc Code	01609560	01699267
ΔG_{bind}	-13.60	-11.20
$\Delta E_{QM/MM}$	-51.95	-52.84
ΔG_{solv}	24.50	27.74
$T\Delta S$	-13.85	-13.91
Structure		
Zinc Code	02381665	39712934
ΔG_{bind}	-10.58	-4.93
$\Delta E_{QM/MM}$	-51.91	-51.70
ΔG_{solv}	27.13	31.85
$T\Delta S$	-14.20	-14.93
Structure		
Zinc Code	34207036	02041164
ΔG_{bind}	-5.92	-12.35
$\Delta E_{QM/MM}$	-66.76	-51.44
ΔG_{solv}	45.98	25.45
$T\Delta S$	-14.86	-13.64
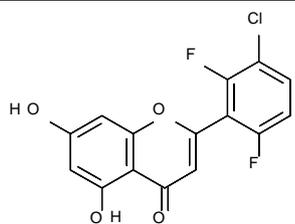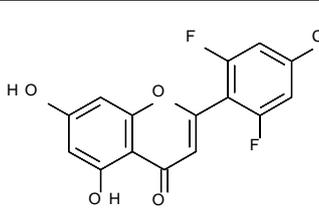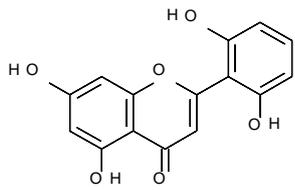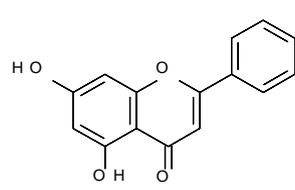
[a]All Values in table refer to results from the QM/MM-PBSA (2) in kcal/mol.

Table 6.3: QM/MM-PBSA(1) results showing the breakdown of ΔG_{bind}.

Rank[c]	Ligand Zinc Code	2'	3'	4'	5'	6'	Simple Sum Rank a^a	b^b	$\Delta E_{QM/MM}$ [d]	ΔG_{solv}^{PB} [d]	ΔG_{solv}^{SA} [d]	ΔG_{solv}^{total} [d]	$T\Delta S$ [d]	ΔG_{bind} [c,d]
1(28)	01566597	H	H	Br	H	H	671 (346)	442 (146)	-59.29	26.09	-4.86	21.23	-14.69	-23.37 (-13.84)
2(15)	01997104	H	H	H	H	OH	137 (51)	573 (242)	-70.83	38.18	-4.78	33.40	-14.69	-22.74 (-15.37)
3(41)	01609560	H	H	H	H	Cl	229 (95)	759 (416)	-57.88	27.20	-4.98	22.21	-13.85	-21.82 (-12.32)
4(10)	20248556	F	H	H	Br	H	412 (198)	364 (103)	-59.96	28.66	-4.93	23.73	-14.73	-21.50 (-15.90)
5(26)	05782628	H	H	H	CH₃	OH	483 (232)	119 (23)	-65.72	34.82	-4.96	29.86	-14.96	-20.89 (-13.99)
6(18)	11804354	NH₂	H	H	Cl	H	89 (23)	20 (4)	-65.65	35.80	-4.58	31.22	-15.14	-19.29 (-15.06)
7(27)	01699267	H	H	H	H	Br	269 (119)	837 (464)	-59.07	30.53	-4.55	25.98	-13.90	-19.18 (-13.88)
8(50)	02506767	H	H	Cl	H	F	-	432 (140)	-57.08	28.14	-4.69	23.45	-14.49	-19.13 (-11.48)
9(4)	01995213	H	H	H	F	H	426 (207)	343 (95)	-56.08	28.45	-4.76	23.69	-13.37	-19.02 (-16.98)
10(14)	01845844	H	H	F	Br	H	127 (44)	157 (36)	-60.71	32.43	-5.14	27.29	-14.55	-18.87 (-15.45)
11(13)	02560120	H	H	F	Cl	H	212 (81)	104 (16)	-59.63	31.45	-4.79	26.66	-14.24	-18.73 (-15.55)
12(8)	02140809	H	H	F	H	F	188 (75)	297 (78)	-57.21	29.59	-4.47	25.13	-13.69	-18.39 (-16.15)
13(21)	02243102	F	H	H	I	H	591 (289)	500 (175)	-57.34	29.39	-5.06	24.33	-14.63	-18.38 (-14.79)
14(11)	02382083	F	H	H	Cl	H	458 (225)	212 (52)	-57.07	29.55	-4.95	24.60	-14.32	-18.15 (-15.89)
15(2)	20248548	OH	H	H	Cl	H	166 (67)	107 (18)	-71.17	43.26	-5.11	38.15	-15.24	-17.78 (-17.79)

[a]: "Original GPb Receptor". [b]: "Induced Fit Receptor". [c]: Values in brackets from preliminary QM/MM-PBSA calculations. [d]: (kcal/mol) Chrysin highlighted in red.

78

Table 6.3 (continued): QM/MM-PBSA(1) results showing the breakdown of ΔG_{bind}.

16(55)	13481417	H	H	H	OH	OH	95 (28)	299 (80)	-79.51	50.38	-4.61	45.77	-15.99	-17.74 (-11.14)
17(34)	02140813	H	H	Cl	Cl	H	258 (112)	201 (49)	-59.11	32.04	-4.91	27.13	-14.79	-17.19 (-12.84)
18(3)	03860261	H	H	H	Cl	H	544 (272)	479 (164)	-56.89	30.59	-4.94	25.66	-14.11	-17.12 (-17.70)
19(46)	34207038	H	H	H	NH$_2$	H	66 (13)	435 (141)	-59.52	32.78	-4.69	28.09	-14.42	-17.01 (-11.57)
20(16)	01847456	H	H	F	H	H	129 (45)	530 (211)	-59.04	32.80	-4.76	28.04	-14.07	-16.93 (-15.33)
21(9)	14989302	H	Br	H	F	H	284 (125)	389 (120)	-56.32	29.18	-4.42	24.76	-14.64	-16.93 (-15.91)
22(51)	01662293	H	H	H	Br	H	608 (300)	410 (131)	-57.31	31.07	-4.76	26.31	-14.11	-16.89 (-11.40)
23(7)	02022413	H	H	H	H	F	353 (168)	609 (269)	-56.43	30.63	-4.39	26.24	-13.64	-16.54 (-16.19)
24(42)	02539282	H	H	F	F	H	162 (66)	76 (12)	-58.70	33.86	-5.03	28.83	-14.00	-15.87 (-12.24)
25(30)	16159637	H	H	H	Br	F	557 (274)	366 (108)	-56.23	30.86	-4.73	26.13	-14.33	-15.76 (-13.66)
26(45)	20248536	H	H	Cl	H	OH	-	277 (69)	-71.56	45.10	-4.62	40.48	-15.43	-15.65 (-11.85)
27(52)	34207036	H	OH	H	OH	H	284 (126)	182 (42)	-71.88	46.16	-4.67	41.49	-14.86	-15.53 (-11.29)
28(29)	67800103	F	H	H	Br	F	428 (209)	360 (101)	-55.13	30.35	-4.92	25.43	-14.43	-15.27 (-13.75)
29(38)	02539342	H	F	H	F	H	144 (54)	403 (127)	-55.03	30.39	-4.60	25.79	-14.20	-15.03 (-12.41)
30(5)	60084043	F	H	F	CH$_3$	H	392 (189)	512 (188)	-59.81	34.58	-4.96	29.62	-15.33	-14.87 (-16.87)
31(43)	02539281	F	H	H	F	H	327 (154)	220 (53)	-56.20	32.22	-4.86	27.36	-13.98	-14.86 (-12.09)
32(25)	02390497	Cl	H	H	Br	H	198 (78)	414 (133)	-58.38	33.94	-4.85	29.09	-14.64	-14.65 (-14.14)
33(54)	14443449	H	H	H	OH	H	235 (99)	601 (266)	-70.06	45.21	-4.67	40.54	-15.07	-14.45 (-11.16)
34(6)	39056287	CHO	H	H	H	H	232 (98)	-	-58.52	34.48	-4.67	29.81	-14.36	-14.36 (-16.37)
35(22)	02382234	H	H	H	Cl	F	424 (206)	287 (76)	-55.62	31.98	-4.92	27.06	-14.27	-14.29 (-14.56)
36(12)	01604094	H	H	H	I	H	622 (308)	529 (210)	-57.39	33.50	-4.71	28.79	-14.34	-14.25 (-15.85)
37(40)	**02041164**	**H**	**H**	**H**	**H**	**H**	**207 (80)**	**590 (258)**	**-56.24**	**33.30**	**-4.82**	**28.48**	**-13.64**	**-14.12 (-12.35)**
38(58)	36533250	H	H	Br	H	CH$_3$	528 (259)	509 (187)	-60.45	35.98	-4.96	31.03	-15.36	-14.06 (-11.02)
39(31)	39056301	Et	H	H	H	H	286 (127)	995 (557)	-58.65	34.99	-4.74	30.25	-14.36	-14.04 (-13.32)

[a]: "Original GPb Receptor". [b]: "Induced Fit Receptor". [c]: Values in brackets from preliminary QM/MM-PBSA calculations. [d]: (kcal/mol) Chrysin highlighted in red.

Table 6.3 (continued): QM/MM-PBSA(1) results showing the breakdown of ΔG_{bind}

40(37)	02582858	H	NS	N	H	H		339 (93)	-62.86	38.77	-4.82	33.95	-14.88	-14.03 (-12.58)
41(57)	02140810	F	H	H	H	F	326 (153)	721 (374)	-54.46	31.04	-4.74	26.30	-14.22	-13.94 (-11.04)
42(56)	02527758	H	Cl	H	F	H	262 (115)	222 (54)	-55.62	32.12	-4.67	27.45	-14.46	-13.71 (-11.07)
43(17)	39056676	OH	H	H	CHO	H	122 (39)	524 (204)	-72.85	48.59	-4.64	43.95	-15.45	-13.44 (-15.18)
44(24)	02381665	F	H	H	Cl	F	400 (192)	271 (66)	-54.51	31.82	-4.93	26.90	-14.20	-13.41 (-14.25)
45(48)	04262390	F	H	F	F	H	136 (49)	63 (8)	-57.47	34.51	-4.59	29.93	-14.21	-13.34 (-11.52)
46(44)	39057325	CH$_2$Cl	H	F	H	H	179 (72)	284 (74)	-62.70	39.64	-4.52	35.12	-14.41	-13.17 (-11.97)
47(19)	Chrysin_BC1-1	H	H	H	H	H	511 (252)	750 (403)	-59.68	37.41	-5.15	32.26	-14.58	-12.84 (-14.84)
48(32)	02545260	F	H	F	H	F	272 (120)	226 (57)	-55.62	33.43	-5.01	28.43	-14.36	-12.83 (-12.89)
49(1)	39056475	H	H	F	H	OH	67 (14)	351 (97)	-71.96	47.17	-2.55	44.61	-14.86	-12.48 (-18.73)
50(23)	02584346	H	H	H	F	F	255 (110)	235 (61)	-54.52	33.06	-4.91	28.16	-14.11	-12.26 (-14.44)
51(53)	39057324	CH$_2$Cl	H	H	F	H	134 (48)	361 (102)	-61.60	39.94	-4.82	35.12	-14.37	-12.11 (-11.21)
52(47)	39056678	CH$_2$OCH$_3$	H	H	H	H	170 (69)	617 (278)	-61.37	39.98	-4.68	35.30	-14.43	-11.65 (-11.53)
53(36)	39056429	F	H	H	H	OH	6 (2)	371 (110)	-67.43	45.37	-4.38	40.98	-15.00	-11.45 (-12.67)
54(49)	02560325	CH$_3$	H	H	F	H	230 (96)	348 (96)	-56.80	36.89	-5.03	31.86	-14.22	-10.72 (-11.51)
55(39)	39712934	F	H	Cl	H	F	-	398 (125)	-54.50	33.85	-4.73	29.11	-14.93	-10.45 (-12.37)
56(20)	39056495	OH	H	H	F	H	56 (11)	270 (65)	-69.97	49.81	-5.16	44.65	-15.08	-10.23 (-14.80)
57(33)	02600074	H	H	H	CF$_3$	F	-	467 (154)	-56.54	35.70	-4.74	30.96	-15.56	-10.03 (-12.87)
58(35)	02382119	H	H	F	Cl	F	304 (145)	113 (20)	-58.24	40.40	-4.60	35.80	-14.37	-8.07 (-12.76)
-	01627231	H	CH$_3$	H	H	H	-	-	-57.77	34.64	-4.86	29.78	-14.52	-13.47
-	02040256	H	H	CH$_3$	H	H	-	-	-50.53	28.42	-4.59	23.83	-14.68	-12.01

[a]: "Original GPb Receptor". [b]: "Induced Fit Receptor". [c]: Values in brackets from preliminary QM/MM-PBSA calculations. [d]: (kcal/mol) Chrysin highlighted in red.

The second set of QM/MM-PBSA(2) tests (large QM region, results shown in Table 6.4) employed the ligand, the two aromatic π-stacking residues Phe285 and Tyr613 and the residues inside the hydrophobic pocket (Asp283, Asn284, Glu382, Leu380, Hie571, Glu572, Tyr573, Arg770 and Phe771) as the QM region, with the remainder of the receptor being the MM region. This has a greater number of less negative ΔG_{bind} values compared to the fast QM/MM-PBSA(1) with a smaller QM region. The QM/MM-PBSA(2) calculations revealed that 20 ligands gave better ΔG_{bind} values than chrysin. Out of these 20 ligands, 14 also performed better in the preliminary QM/MM-PBSA calculations. Of the total 58 ligands, the QM/MM-PBSA(2) calculations did however give 28 ligands with values worse than the -11 kcal/mol cut-off, which was the threshold value used for selection of candidates from the preliminary QM/MM-PBSA calculations. Generally, the $\Delta E_{QM/MM}$ values dominate, with values between -76.98 and -50.11 kcal/mol, the ligand with a value of -76.98 kcal/mol has hydroxyl groups in the 2' and 3' position, whilst the ligand with a value of -50.11 kcal/mol has two ortho flourines. An increase in polarity correlates with better $\Delta E_{QM/MM}$ values. Whilst total solvation free energies range from 23.35 and 45.98 kcal/mol, the ligand that gave the least favourable value of 45.98 kcal/mol has two hydroxyl groups on the meta positions of the B ring, whilst the best result of 23.35 kcal/mol has two flourines at the ortho positions, consistent with the values from the QM/MM-PBSA(1) calculations. The non-polar contributions to solvation are similar as well, consistently around -4 to -5 kcal/mol, with the polar contributions dominating. Entropy values for QM/MM-PBSA(2) in Table 6.3 are identical to the values for QM/MM-PBSA(1) in Table 6.1 due to the nature of the calculations using the same input pose, which favour smaller halogenated derivatives and the ortho position on the B ring. In summary and in general, therefore, the results from QM/MM-PBSA(1) and QM/MM-PBSA(2) appear to be quite consistent, adding credibility to the predictive nature of these calculations.

Table 6.4: QM/MM-PBSA(2) results showing the breakdown of ΔG_{bind}
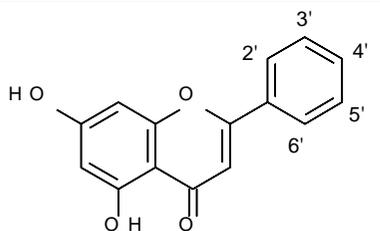
Rank[c]	Ligand Zinc Code	2′	3′	4′	5′	6′	a[a]	b[b]	$\Delta E_{QM/MM}$[d]	ΔG_{solv}^{PB}[d]	ΔG_{solv}^{SA}[d]	ΔG_{solv}^{total}[d]	$T\Delta S$[d]	ΔG_{bind}[c,d]
							Simple Sum Rank		QM/MM-PBSA (kcal/mol)					
1(26)	05782628	H	H	H	CH$_3$	OH	483 (232)	119 (23)	-62.93	31.70	-4.81	26.88	-14.96	-21.09 (-13.99)
2(15)	01997104	H	H	H	H	OH	137 (51)	573 (242)	-65.59	37.47	-4.55	32.93	-14.69	-17.98 (-15.37)
3(55)	13481417	H	H	H	OH	OH	95 (28)	299 (80)	-76.98	48.44	-4.72	43.72	-15.99	-17.26 (-11.14)
4(14)	01845844	H	H	F	Br	H	127 (44)	157 (36)	-55.75	29.18	-4.94	24.24	-14.55	-16.97 (-15.45)
5(51)	01662293	H	H	H	Br	H	608 (300)	410 (131)	-53.73	28.26	-4.59	23.67	-14.11	-15.95 (-11.40)
6(25)	02390497	Cl	H	H	Br	H	198 (78)	414 (133)	-53.78	28.44	-4.79	23.64	-14.64	-15.49 (-14.14)
7(28)	01566597	H	H	Br	H	H	671 (346)	442 (146)	-55.03	29.99	-4.65	25.34	-14.69	-15.00 (-13.84)
8(16)	01847456	H	H	F	H	H	129 (45)	530 (211)	-53.45	29.30	-4.76	24.54	-14.07	-14.84 (-15.33)
9(22)	02382234	H	H	H	Cl	F	424 (206)	287 (76)	-52.91	29.29	-4.65	24.64	-14.27	-13.99 (-14.56)
10(29)	67800103	F	H	H	Br	F	428 (209)	360 (101)	-52.78	29.56	-5.05	24.51	-14.43	-13.84 (-13.75)
11(9)	14989302	H	Br	H	F	H	284 (125)	389 (120)	-52.57	28.65	-4.42	24.22	-14.64	-13.71 (-15.91)
12(41)	01609560	H	H	H	H	Cl	229 (95)	759 (416)	-51.95	28.85	-4.35	24.50	-13.85	-13.60 (-12.32)
13(45)	20248536	H	H	Cl	H	OH	-	277 (69)	-67.98	43.23	-4.24	38.99	-15.43	-13.56 (-11.85)
14(2)	20248548	OH	H	H	Cl	H	166 (67)	107 (18)	-67.39	43.69	-4.93	38.76	-15.24	-13.40 (-17.79)
15(21)	02243102	F	H	H	I	H	591 (289)	500 (175)	-52.39	29.20	-4.61	24.59	-14.63	-13.17 (-14.79)

[a]: "Original GPb Receptor". [b]: "Induced Fit Receptor". [c]: Values in brackets from preliminary QM/MM-PBSA calculations. [d]: (kcal/mol) Chrysin highlighted in red.

Table 6.4 (continued): QM/MM-PBSA(2) results showing the breakdown of ΔG_{bind}

16(35)	02382119	H	H	F	Cl	F	304 (145)	113 (20)	-54.62	31.67	-4.58	27.09	-14.37	-13.17 (-12.76)
17(57)	02140810	F	H	H	H	F	326 (153)	721 (374)	-50.11	28.26	-4.91	23.35	-14.22	-12.54 (-11.04)
18(46)	34207038	H	H	H	NH$_2$	H	66 (13)	435 (141)	-55.90	33.63	-4.61	29.03	-14.42	-12.46 (-11.57)
19(8)	02140809	H	H	F	H	F	188 (75)	297 (78)	-51.24	29.36	-4.22	25.14	-13.69	-12.41 (-16.15)
20(10)	20248556	F	H	H	Br	H	412 (198)	364 (103)	-55.35	33.39	-5.13	28.26	-14.73	-12.36 (-15.90)
21(40)	**02041164**	**H**	**H**	**H**	**H**	**H**	**207 (80)**	**590 (258)**	**-51.44**	**29.92**	**-4.47**	**25.45**	**-13.64**	**-12.35 (-12.35)**
22(13)	02560120	H	H	F	Cl	H	212 (81)	104 (16)	-54.57	33.00	-4.81	28.19	-14.24	-12.14 (-15.55)
23(54)	14443449	H	H	H	OH	H	235 (99)	601 (266)	-66.42	43.80	-4.53	39.27	-15.07	-12.07 (-11.16)
24(42)	02539282	H	H	F	F	H	162 (65)	76 (12)	-52.80	31.45	-4.65	26.80	-14.00	-12.00 (-12.24)
25(34)	02140813	H	H	Cl	Cl	H	258 (112)	201 (49)	-56.07	34.44	-5.15	29.29	-14.79	-11.99 (-12.84)
26(5)	60084043	F	H	F	CH$_3$	H	392 (189)	512 (188)	-55.03	32.76	-4.53	28.23	-15.33	-11.47 (-16.87)
27(37)	02582858	H	NS	N	H	H	-	339 (93)	-59.89	38.03	-4.45	33.58	-14.88	-11.43 (-12.58)
28(27)	01699267	H	H	H	H	Br	269 (119)	837 (464)	-52.84	32.19	-4.46	27.74	-13.91	-11.20 (-13.88)
29(36)	39056429	F	H	H	H	OH	6 (2)	371 (110)	-63.56	42.05	-4.60	37.45	-15.00	-11.11 (-12.67)
30(48)	04262390	F	H	F	F	H	136 (49)	63 (8)	-52.21	31.24	-4.33	26.91	-14.21	-11.10 (-11.52)
31(32)	02545260	F	H	F	H	F	272 (120)	226 (57)	-51.28	30.37	-4.36	26.01	-14.36	-10.91 (-12.89)
32(4)	01995213	H	H	H	F	H	426 (207)	343 (95)	-50.65	31.38	-4.87	26.51	-13.37	-10.77 (-16.98)
33(11)	02382083	F	H	H	Cl	H	458 (225)	212 (52)	-52.87	32.46	-4.62	27.84	-14.32	-10.72 (-15.89)
34(1)	39056475	H	H	F	H	OH	67 (14)	351 (97)	-66.13	44.80	-4.14	40.66	-14.86	-10.60 (-18.73)
35(24)	02381665	F	H	H	Cl	F	400 (192)	271 (66)	-51.91	31.93	-4.80	27.13	-14.20	-10.58 (-14.25)
36(50)	02506767	H	H	Cl	H	F	-	432 (140)	-52.69	32.26	-4.30	27.96	-14.49	-10.23 (-11.48)
37(12)	01604094	H	H	H	I	H	622 (308)	529 (210)	-53.08	33.67	-4.91	28.77	-14.34	-9.97 (-15.85)
38(7)	02022413	H	H	H	H	F	353 (168)	609 (269)	-51.04	31.95	-4.48	27.47	-13.64	-9.92 (-16.19)
39(58)	36533250	H	H	Br	H	CH$_3$	528 (259)	509 (187)	-56.48	35.95	-4.68	31.27	-15.36	-9.85 (-11.02)

[a]: "Original GPb Receptor". [b]: "Induced Fit Receptor". [c]: Values in brackets from preliminary QM/MM-PBSA calculations. [d]: (kcal/mol) Chrysin highlighted in red.

Table 6.4 (continued): QM/MM-PBSA(2) results showing the breakdown of ΔG_{bind}

40(47)	39056678	CH$_2$O CH$_3$	H	H	H	H	170 (69)	617 (278)	-57.14	37.83	-4.75	33.08	-14.43	-9.63 (-11.53)
41(30)	16159637	H	H	H	Br	F	557 (274)	366 (108)	-53.78	35.00	-4.92	30.08	-14.33	-9.36 (-13.66)
42(18)	11804354	NH$_2$	H	H	Cl	H	89 (23)	20 (4)	-60.41	40.39	-4.41	35.99	-15.14	-9.28 (-15.06)
43(17)	39056676	OH	H	H	CHO	H	122 (39)	524 (204)	-70.22	50.72	-4.81	45.91	-15.45	-8.86 (-15.18)
44(33)	02600074	H	H	H	CF$_3$	F	-	467 (154)	-55.13	35.34	-4.59	30.75	-15.56	-8.83 (-12.87)
45(3)	03860261	H	H	H	Cl	H	544 (272)	479 (164)	-53.07	35.16	-4.91	30.25	-14.11	-8.71 (-17.70)
46(23)	02584346	H	H	H	F	F	255 (110)	235 (61)	-50.63	32.55	-4.62	27.93	-14.11	-8.59 (-14.44)
47(31)	39056301	Et	H	H	H	H	286 (127)	995 (557)	-53.55	35.40	-4.79	30.61	-14.36	-8.57 (-13.32)
48(43)	02539281	F	H	H	F	H	327 (154)	220 (53)	-51.14	33.30	-4.31	28.99	-13.98	-8.18 (-12.09)
49(20)	39056495	OH	H	H	F	H	56 (11)	270 (65)	-65.26	47.21	-4.68	42.53	-15.08	-7.65 (-14.80)
50(38)	02539342	H	F	H	F	H	144 (54)	403 (127)	-50.61	33.66	-4.86	28.80	-14.20	-7.62 (-12.41)
51(56)	02527758	H	Cl	H	F	H	262 (115)	222 (54)	-51.76	34.44	-4.57	29.88	-14.46	-7.43 (-11.07)
52(49)	02560325	CH$_3$	H	H	F	H	230 (96)	348 (96)	-51.51	34.63	-4.64	29.98	-14.22	-7.31 (-11.51)
53(52)	34207036	H	OH	H	OH	H	284 (126)	182 (42)	-66.76	50.80	-4.83	45.98	-14.86	-5.92 (-11.29)
54(53)	39057324	CH$_2$Cl	H	H	F	H	134 (48)	361 (102)	-56.65	41.70	-4.71	37.00	-14.37	-5.29 (-11.21)
55(39)	39712934	F	H	Cl	H	F	-	398 (125)	-51.70	36.78	-4.93	31.85	-14.93	-4.93 (-12.37)
56(6)	39056287	CHO	H	H	H	H	232 (98)	-	-53.46	39.32	-4.91	34.41	-14.36	-4.69 (-16.37)
57(44)	39057325	CH$_2$Cl	H	F	H	H	179 (72)	284 (74)	-57.29	42.95	-4.72	38.23	-14.41	-4.65 (-11.97)
58(19)	Chrysin_BC 1-1	H	H	H	H	H	511 (252)	750 (403)	-54.32	41.35	-5.33	36.01	-14.58	-3.73 (-14.84)
-	01627231	H	CH$_3$	H	H	H	-	-	-54.05	27.50	-4.55	30.61	-14.52	-12.04
-	02040256	H	H	CH$_3$	H	H	-	-	-48.52	25.96	-4.64	32.05	-14.68	-7.87

[a]: "Original GPb Receptor". [b]: "Induced Fit Receptor". [c]: Values in brackets from preliminary QM/MM-PBSA calculations. [d]: (kcal/mol) Chrysin highlighted in red.

Consensus QM/MM-PBSA ranking

Using ΔG_{bind} as a new rank system for the preliminary QM/MM-PBSA, the faster QM/MM-PBSA(1) and the slower QM/MM-PBSA(2), we combined the results using a Simple Sum Rank of the ligands to decide which ligands performed the best across the board [144]. This revealed that 13 ligands achieved a higher Simple Sum Rank than the benchmark chrysin, as shown in Table 6.5. Out of the 13 ligands which achieved a higher Simple Sum Rank than the benchmark chrysin, two ligands scored better than chrysin in every QM/MM-PBSA calculation, both of which have a single hydroxyl group in the ortho position. Out of the top 13 ligands, 7 of them contain fluorine atoms, mostly in ortho and para positions. Generally, the position that seems to dominate the most for substitutions is the meta position, with 8 of the 13 ligands displaying substituents at this position. For example, ligand with Zinc code 4989302 has a fluorine and bromine at the meta positions. The only other substituent to appear here is a non-polar methyl group. The lack of polarity at the meta position suggests that polarity here leads to poor binding due to the hydrophobic pocket, consistent with previous suggestions [83]. The highest scoring ligand (Zinc code 13481417) that displays polarity (hydroxyl) at the meta position has a sum rank of 74, which is 9 places lower than the benchmark chrysin. This lack of polarity inside the pocket is even more apparent for the para position, as none of the ligands that made it past the filtering stages has para polar substituents such as hydroxyls. This is also consistent with previous work [83]. For the top 13 compounds, apart from ligand Zinc code 01566597 with a para-bromine substituent and whose binding is shown in Fig 6.4(a), only fluorine appears in the para position. This may be due to it's size; being the smallest hydrophobic substituent present in these results. The ortho position shows different preferences. This allows substituents to be more polar, as they do not project as deeply into the hydrophobic pocket. Three of the best 13 ligands display hydroxyl groups at the ortho position, and all three give high ΔG_{bind} values. Ligand Zinc code 01997104 which has only one substituent, a hydroxyl at the ortho position, scored the highest in regards to the Simple Sum Rank. One reason that the hydroxyls bind well when at the ortho position is that the hydrogen forms a hydrogen bond to the oxygen on the Glu572, which can be seen with ligands Zinc codes 01997104 (ortho = OH) and 20248548 (ortho = OH; meta = Cl), and whose binding is shown in Figures 6.4(c) and 6.4(d), respectively. The ligand 01997104 (Fig 6.4(c)) hydroxyl forms a hydrogen bond with a donor angle of 155° and a distance of 1.94 Å; these values imply the hydrogen bond between the two is strong. The same trend is observed for Zinc code 20248548 (Fig

6.4(d)), with values of 1.81 Å and 161°. Ligand Zinc code 01609560 (Fig 6.4(b)) with a chloro ortho substituent is also predicted to be favourable. When this ligand binds, the chlorine may form halogen bonding to the same Glu572 oxygen; the distance between the chlorine and oxygen is 3.22 Å, which is close to the optimal value of 3.12 Å [150]. This halogen bond is a key contributor factor to the favourable ΔG_{bind} value. Ligands with halogen substituents show a trend with less favorable $\Delta E_{QM/MM}$ contributions, whilst polar groups have more negative and favourable $\Delta E_{QM/MM}$ values. However, due to the polarity, these ligands require more energy to desolvate, which is shown in the ΔG_{solv}^{PB} values; polar ligands have much higher values, which brings down the overall ΔG_{bind} of the ligand. On the other hand, non-polar halogen substituents with less $\Delta E_{QM/MM}$ contributions to binding do not get as heavily penalised by high ΔG_{solv}^{PB} results, due to them not interacting with solvents as strongly.

Table 6.5 – Predicting the best ligands for GPb inhibition by exploiting consensus scoring the values obtained for QM/MM-PBSA calculations.

Rank By Simple Sum Rank	Ligand	2'	3'	4'	5'	6'	ΔG_{bind} Preliminary (Kcal/mol)	ΔG_{bind} Preliminary rank	ΔG_{bind} QM/MM-PBSA(1) (Kcal/mol)	ΔG_{bind} QM/MM-PBSA(1) rank	ΔG_{bind} QM/MM-PBSA(2) (Kcal/mol)	ΔG_{bind} QM/MM-PBSA(2) rank	Simple Sum Rank
1	01997104	H	H	H	H	OH	-16.98	4	-22.74	2	-17.98	2	8
2	01566597	H	H	Br	H	H	-14.85	19	-23.37	1	-15.00	7	27
3	01609560	H	H	H	H	Cl	-15.85	12	-21.82	3	-13.60	12	27
4	20248548	OH	H	H	Cl	H	-17.79	2	-17.78	15	-13.40	14	31
5	05782628	H	H	H	CH₃	OH	-13.99	27	-20.89	5	-21.09	1	33
6	20248556	F	H	H	Br	H	-15.90	10	-21.50	4	-12.36	20	34
7	14989302	H	Br	H	F	H	-15.91	9	-16.93	21	-13.71	11	41
8	01845844	H	H	F	Br	H	-13.88	28	-18.87	10	-16.97	4	42
9	01847456	H	H	F	H	H	-15.45	14	-16.93	20	-14.84	8	42
10	01995213	H	H	H	F	H	-15.33	16	-19.02	9	-10.77	32	57
11	02390497	Cl	H	H	Br	H	-14.56	23	-14.65	32	-15.49	6	61
12	60084043	F	H	F	CH₃	H	-16.87	5	-14.87	30	-11.47	26	61
13	02243102	F	H	H	I	H	-12.84	35	-18.38	13	-13.17	15	63
14	**02041164**	**H**	**H**	**H**	**H**	**H**	**-16.19**	**7**	**-14.12**	**37**	**-12.35**	**21**	**65**
15	02140810	F	H	H	H	F	-16.15	8	-13.94	41	-12.54	17	66
16	02560120	H	H	F	Cl	H	-12.89	33	-18.73	11	-12.14	22	66
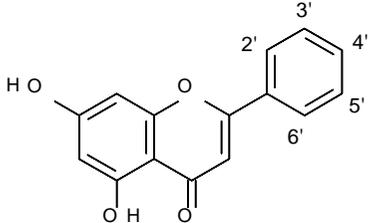
Table 6.5 (continued) - Predicting the best ligands for GPb inhibition by exploiting consensus scoring the values obtained for QM/MM-PBSA calculations.

17	101804354	NH₂	H	H	Cl	H	-15.06	18	-19.29	6	-9.28	42	66
18	67800103	F	H	H	Br	F	-13.75	30	-15.27	28	-13.84	10	68
19	01662293	H	H	H	Br	H	-12.32	42	-16.89	22	-15.95	5	69
20	02506767	H	H	Cl	H	F	-14.14	26	-19.13	8	-10.23	36	70
21	02140809	H	H	F	H	F	-12.35	41	-18.39	12	-12.41	19	72
22	02382083	F	H	H	Cl	H	-14.25	25	-18.15	14	-10.72	33	72
23	13481417	H	H	H	OH	OH	-11.14	55	-17.74	16	-17.26	3	74
24	02022413	H	H	H	H	F	-15.37	15	-16.54	23	-9.92	38	76
25	04262390	F	H	F	F	H	-17.70	3	-13.34	45	-11.10	30	78
26	02382234	H	H	H	Cl	F	-12.76	36	-14.29	35	-13.99	9	80
27	34207038	H	H	H	NH₂	H	-11.57	47	-17.01	19	-12.46	18	84
28	39056475	H	H	F	H	OH	-18.73	1	-12.48	49	-10.60	34	84
29	20248536	H	H	Cl	H	OH	-11.85	46	-15.65	26	-13.56	13	85
30	02382119	H	H	F	Cl	F	-15.89	11	-8.07	58	-13.17	16	85
31	01699267	H	H	H	H	Br	-11.40	51	-19.18	7	-11.20	28	86
32	02539282	H	H	F	F	H	-12.09	44	-15.87	24	-12.00	24	92
33	39056287	CHO	H	H	H	H	-16.37	6	-14.36	34	-4.69	56	96
34	16159637	H	H	H	Br	F	-13.66	31	-15.76	25	-9.36	41	97
35	03860261	H	H	H	Cl	H	-12.87	34	-17.12	18	-8.71	45	97
36	02140813	H	H	Cl	Cl	H	-11.04	57	-17.19	17	-11.99	25	99
37	02381665	F	H	H	Cl	F	-14.79	22	-13.41	44	-10.58	35	101
38	01604094	H	H	H	I	H	-13.84	29	-14.25	36	-9.97	37	102
39	39056676	OH	H	H	CHO	H	-15.18	17	-13.44	43	-8.86	43	103
40	14443449	H	H	H	OH	H	-11.16	54	-14.45	33	-12.07	23	110

Table 6.5 (continued) - Predicting the best ligands for GPb inhibition by exploiting consensus scoring the values obtained for QM/MM-PBSA calculations.

41	02582858	H	NS	N	H	H	-11.51	49	-14.03	40	-11.43	27	116
42	02545260	F	H	F	H	F	-12.41	39	-12.83	48	-10.91	31	118
43	39056301	Et	H	H	H	H	-13.32	32	-14.04	39	-8.57	47	118
44	39056429	F	H	H	H	OH	-12.67	37	-11.45	53	-11.11	29	119
45	02560325	CH$_3$	H	H	F	H	-15.55	13	-10.72	54	-7.31	52	119
46	02539342	H	F	H	F	H	-12.24	43	-15.03	29	-7.62	50	122
47	02600074	H	H	H	CF$_3$	F	-14.44	24	-10.03	57	-8.83	44	125
48	Chrysin_BC1-1	H	H	H	H	H	-14.85	20	-12.84	47	-3.73	58	125
49	39056495	OH	H	H	F	H	-14.80	21	-10.23	56	-7.65	49	126
50	34207036	H	OH	H	OH	H	-11.29	52	-15.53	27	-5.92	53	132
51	02584346	H	H	H	F	F	-12.58	38	-12.26	50	-8.59	46	134
52	36533250	H	H	Br	H	CH$_3$	-11.02	58	-14.06	38	-9.85	39	135
53	02539281	F	H	H	F	H	-11.07	56	-14.86	31	-8.18	48	135
54	39056678	CH$_2$O CH$_3$	H	H	H	H	-11.53	48	-11.65	52	-9.63	40	140
55	2527758	H	Cl	H	F	H	-11.48	50	-13.71	42	-7.43	51	143
56	39057325	CH$_2$Cl	H	F	H	H	-11.97	45	-13.17	46	-4.65	57	148
57	39712934	F	H	Cl	H	F	-12.37	40	-10.45	55	-4.93	55	150
58	39057324	CH2Cl	H	H	F	H	-11.21	53	-12.11	51	-5.29	54	158
-	01627231	H	CH$_3$	H	H	H	-8.90	-	-13.47		-12.04		
-	02040256	H	H	CH$_3$	H	H	3.18	-	-12.01		-7.87		
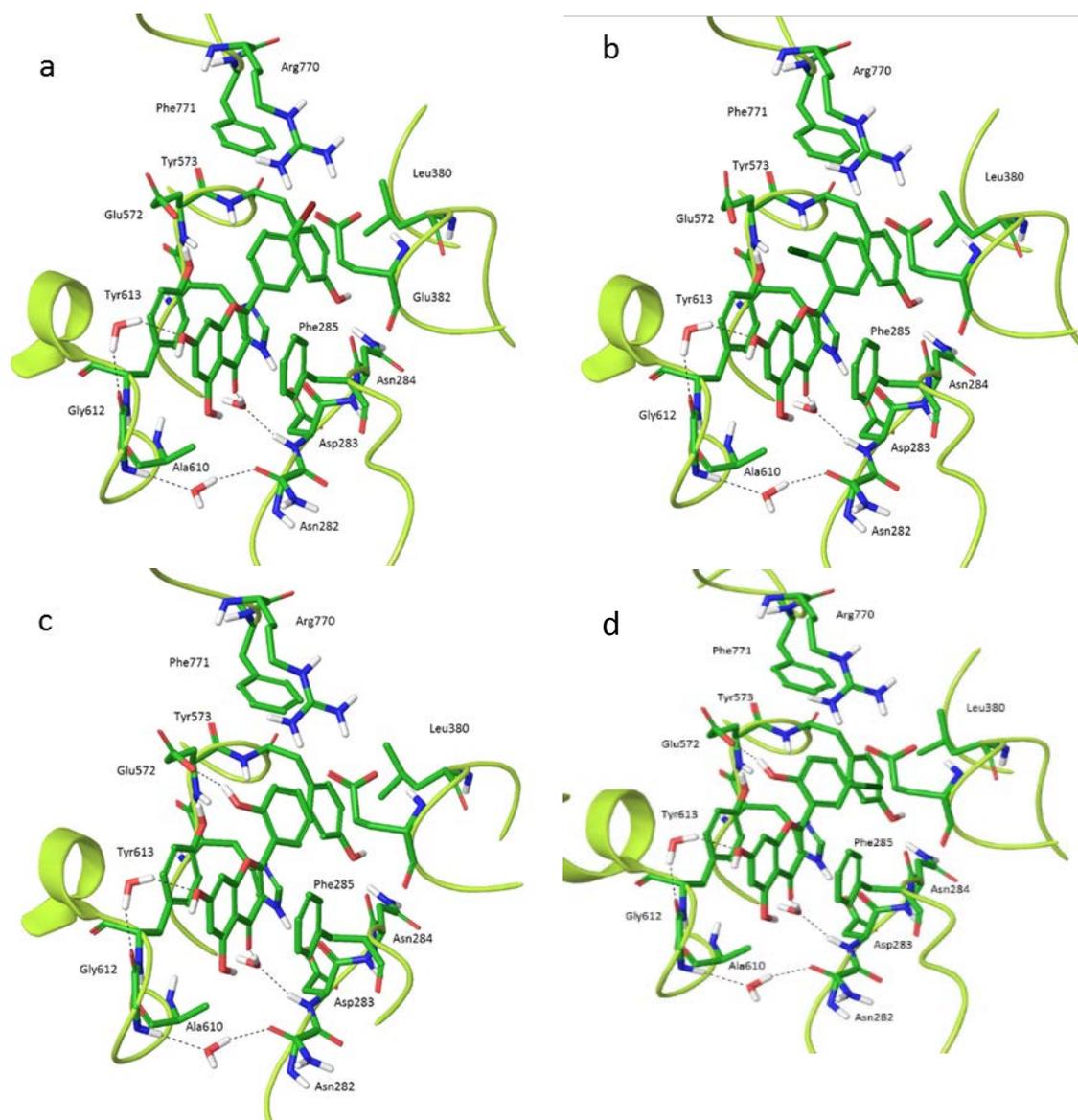
Figure 6.4 – Predicted binding of four of the best candidates at the inhibitor site of GPb. With displayed ribbons. Hydrogen bonds are shown in black, non-polar hydrogens are hidden for clarity. Zinc Codes (a) 01566597, (b) 01609560, (c) 01997104 and (d) 20248548.

6.4 - Conclusion

Computational drug design has extremely useful applications regarding the inhibition of GP for treatment of T2D. By aligning the core of the flavone structures between the two aromatic residues Phe285 and Tyr613 and modelling binding of different ligand groups occupying the hydrophobic cavity formed by key residues Leu380, Tyr573 and Phe771, we can clearly see potential for more effective inhibitor candidates. The computational predictions favour substituted phenyl B rings with ortho-substituted hydroxyl groups

and halogenated substituents at the meta and para positions, whilst polar groups which enter the hydrophobic pocket give much poorer results. Also, even when they are non-polar, bulky groups appear to be too sterically demanding when entering the hydrophobic cavity. Of the top 13 ligands based on consensus scoring of QM/MM-PBSA results, five, seven and one were mono-, di- and tri-substituted, respectively. These results can help guide the decision which ligands would be the most ideal for synthesis; the four ligands: 01566597 (Fig 6.4(a)), 01609560 (Fig 6.4 (b)), 01997104 (Fig 6.4(c)) and 20248548 (Fig 6.4(d)) would be the most ideal and provide a range of substituents. Ligand Zinc code 01566597 has a halogen in the para position, which would show whether non-polar groups do bind well in the hydrophobic pocket. Whilst the ligand with Zinc code 01997104 would be able to confirm that hydroxyl groups at the ortho position interacting with Glu572 gives strong binding. Likewise, the ligand with Zinc code 01609560 has a chlorine in the ortho position, to see whether halogen bonds can compete with hydrogen bonds in this position. The ligand with Zinc code 20248548 has both a hydroxyl in the ortho position and a halogen inside the hydrophobic pocket, which would be able to show whether these two interactions can work in synergy with each other, to give strong binding to GPb. A methyl substitution in the meta position also appears twice (Zinc codes 05782628 and 60084043), indicating that ligands of this type are worthy of synthetic consideration, to start with, however, meta and para methyl substitutions are considered due to likely less synthetic complications and to optimise a synthetic protocol. These ligands demonstrated competitive ΔG_{bind} values in the QM/MM-PBSA calculations here compared to those calculated for chrysin.

Chapter 7 – Synthetic Methods

7.1 - Introduction

Type-2 Diabetes (T2D) is a chronic illness which is characterised by three factors: 1) a resistance of insulin action on glucose uptake, 2) relatively low production of insulin in the pancreas, and 3) impaired insulin action to inhibit the hepatic glucose production [123]. One way of combatting T2D is by inhibiting the glycogen phosphorylase (GP) enzyme, which is a key enzyme in glycogenolysis (breakdown of glycogen to glucose), and is responsible for 70% of the glucose production in the liver. It is well reported that flavonoids can bind to multiple sites. Flavonoids are polyphenols and ubiquitous components in both food of plant origin, found throughout nature and consumed daily as part of a healthy balanced diet. Studies in diabetic and non-diabetic rats have shown that flavonoids do have anti-hyperglycemic properties due to their effect on glycogen metabolism,[151-153] and have been shown to bind at multiple sites on GP (*c.f.* Chapter 3.4). The synthesis of a variety of flavonoids has been reported. But the focus of this work is on the specific group of flavonoids known as flavones (Fig 7.1), and the effects that different substituents at the B ring have on the inhibition of GP at the caffeine binding site.
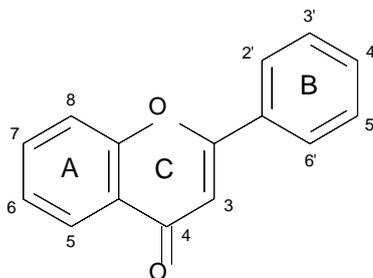
Figure 7.1 – Basic structure of flavones.

There are multiple synthetic routes and methods to yield the flavone structure with the desired substituents. For inhibition of GP, hydroxyl groups on the 5 and 7 position (see Fig 7.1) are required to increase bioavailability, which is important to remember when designing synthetic routes to these compounds so that the hydroxyl groups do not interfer with the desired reactions. As such, multiple synthetic methods were considered, (*c.f.* Chapter 5) including the Baker-Venkataraman rearrangement,

microwave methods, synthesis via 2'-hydroxychalcone, Claisen-Schmidt and the use of metal catalysts.

7.2 - Synthetic Approach to Chrysin (6a).

General procedure:
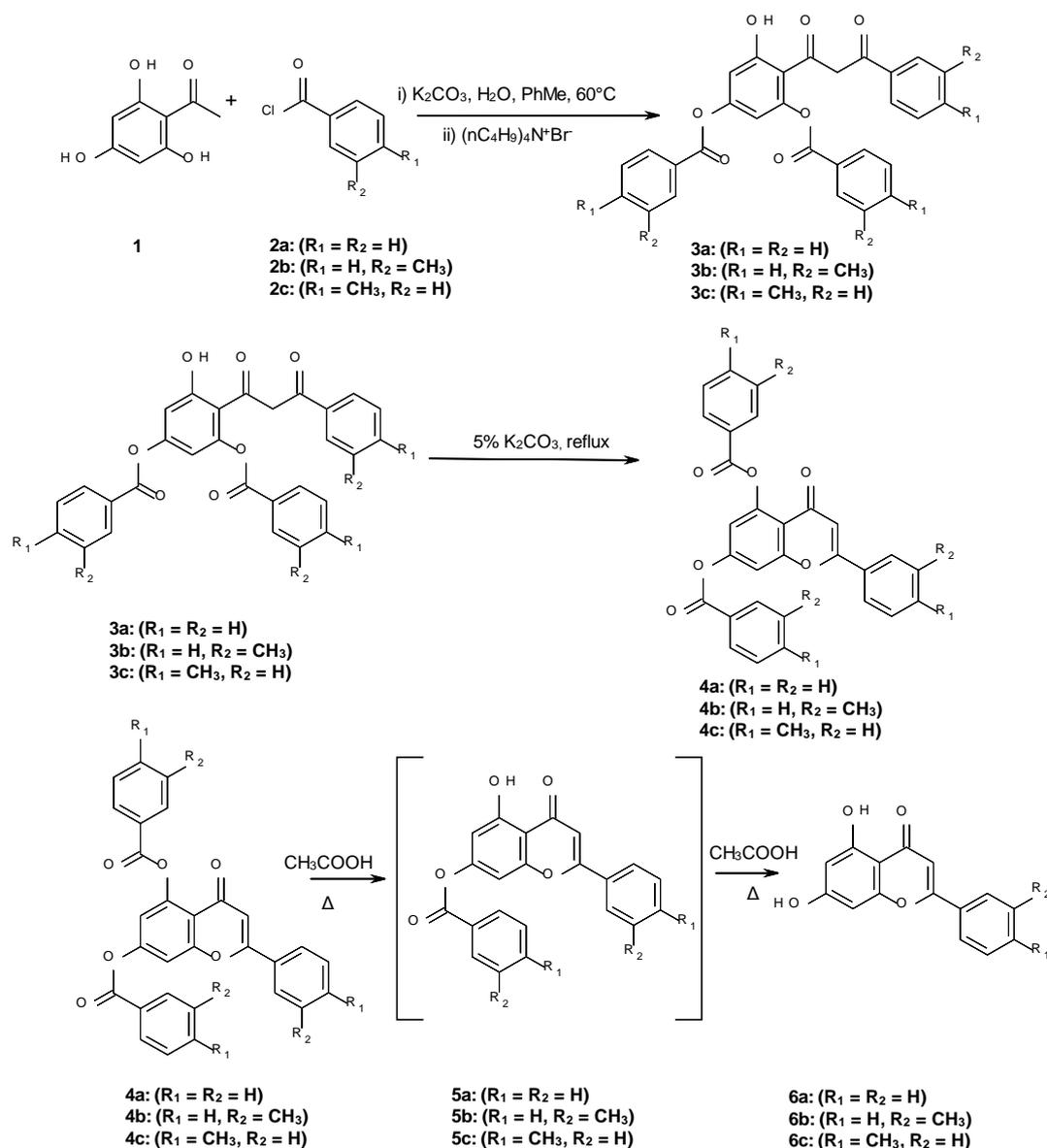
Scheme 7.1 – Synthesis of 5,7-dihydroxyflavone analogues (**6a-c**) via the Baker-Venkataraman rearrangement.

2,4,6-Trihydroxyacetophenone (**1**) was added to benzoyl chloride (**2a**) in a mixture of toluene, saturated potassium carbonate and tetrabutylammonium bromide and stirred, upon which an orange solution of (**3a**) was formed. To this solution, 5% potassium

carbonate was added and the reaction stirred at reflux. A dark orange/brown sticky solid formed in the orange liquid, which was decanted off and added to glacial acetic acid with heat and stirring. When added to ice, a cream coloured solid was yielded and assigned to be the only partially hydrolysed ester, **5a**.

The ^{13}C DEPT NMR showed that cyclisation had occurred, as no CH_2 peak was present at approximately 58 ppm. However, confusingly, ^{1}H and ^{13}C NMR, along with LCMS, confirmed that there was only trace amounts of (**6a**) and that the majority of the isolated compound was an esterified analogue: shown to be 5-hydroxy-4-oxo-2-phenyl-4H-chromen-7-yl benzoate (**5a**) as seen in Fig 7.2.

5a
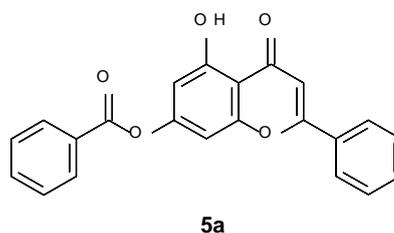
Figure 7.2 – The unexpected partially hydrolysed benzoyl ester of chrysin isolated.

The structure of the ester was realised, since the ^{1}H NMR showed that only one hydroxyl peak had formed in the polar region at 12.24 ppm. This was further evidenced by the aromatic region, which showed that there were 13 aromatic protons. To confirm its identity, LCMS analysis showed a major peak at m/z = 359 which is consistent with the esterified compound, **5a**. Nevertheless, despite not fully hydrolysing, the method did show that the Baker-Venkataraman method was effective for forming the flavone structure. After some time, and further attempts at its synthesis, due to the commercial availability of chrysin, and the trace amounts of compound left after analysis, the decision was made to purchase chrysin and move onto the synthesis of simple hydrophobic chrysin analogues.

7.3 - Synthesis of 5,7-dihydroxy-2-(3-methylphenyl)chromen-4-one (6b)

2,4,6-Trihydroxyacetophenone (**1**) was added to 3-methylbenzoyl chloride (**2b**) in a mixture of toluene, saturated potassium carbonate and tetrabutylammonium bromide and stirred; an orange solution of (**3b**) was formed. The reaction was left until completion, as determined by TLC. The organic layer was separated and washed, then added to potassium carbonate and heated at reflux with stirring to form (**4b**), which was heated at reflux in glacial acetic acid to yield the final compound (**4b**), which was purified by flash column chromatography. Again, despite attempting alternative conditions for ester hydrolysis (increased time, base and temperature), the same problem occurred, with small amounts of the desired product (**6b**) being isolated, but forming large quantities of (**5b**). Furthermore, extra attempts at ester hydrolysis of the isolated product (**5b**) also failed, such as, a screen of bases, the use of nucleophilic solvents, such as methanol, and extended reaction times. Depsite these attempts, the presence of the ester was confirmed by ^{1}H NMR and MS in each case, demonstrating the surprising resilience in the hydrolysis of this ester.

7.4 - Synthesis of 5,7-dihydroxy-2-(4-methylphenyl)chromen-4-one (6c)

This second synthesis aimed to create another methyl-chrysin analogue with the methyl in the para-position of the B ring, using slight changes to the synthetic method as above, in the hope that the para-substituted analogue would be less resilient to hydrolysis. In the event, (**1**) and (**2c**) were mixed together using the same molar ratios as used previously, but more time was given for each step to complete, in an attempt to achieve a higher yield and purity. Once the reaction was deemed complete by a TLC of 1:5 ethyl actete: petroleum ether, it was isolated and analysed by ^{1}H NMR to check for impurities and unreacted starting materials. Unfortunately, the ^{1}H NMR was extremely complex, with no identifiable products having been formed. Due to the lack of success with the previous synthetic attempts, and again here, it was decided to try a different approach, to achieve higher quality compounds for biological evaluation.

7.5 - Second Attempt at Synthesis of 5,7-dihydroxy-2-(4-methylphenyl)chromen-4-one (6c)

General procedure:
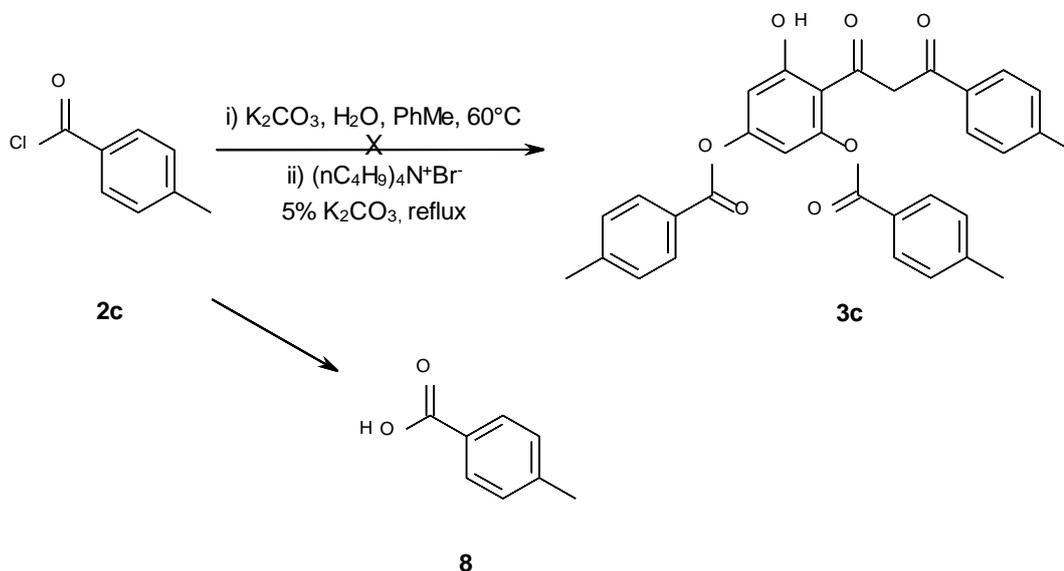
Scheme 7.2 – Hydrolysis of p-toluoyl chloride (**2c**) to form p-toluic acid (**8**)

This reaction was an attempt at the previous synthesis with the para-positioned methyl group on the B ring, but under different conditions, such as longer reaction times, extra catalyst and higher concentrations of base. However the benzoyl chloride reacted with water to form the carboxylic acid, not the desired product (**3c**).

7.6 - Second attempt of synthesis of chrysin (6a).

Due to the trace amounts of desired methylated-chrysin found in the second synthesis above, attempts turned back to create chrysin itself with further alterations to the literature conditions, in the hope of achieving higher yields. This involved using dichloromethane as solvent instead of the previously used toluene, since Shabnam Saxena and co-workers had found that doing so could produce higher yields [154]. In the event, the reaction was evaluated by ^{1}H NMR of the crude mixture to determine how successful the solvent change had been. However, ^{1}H NMR of what was expected to be the intermediate (**3a**) showed a very complex mixture, and so it was concluded that dichloromethane was not suitable here and that toluene would be used for all remaining attempts.

7.7 - Third Attempt of Synthesis of 5,7-dihydroxy-2-(4-methylphenyl)chromen-4-one (6c).

Returning to toluene as solvent, the first step was performed the same as previously attempted, except that less catalyst was used and the reaction was left for a longer time. It was hoped that with less catalyst, the desired intermediate would be easier to isolate pure with improve yield. It was also decided not to wash the organic layer with aqueous solutions, but instead, isolate the organic compound directly from the organic layer, in the hope that there would be less chance of the desired product being lost. Performing this altered work-up yielded the purified esterified product (**5c**) in 15% yield. Once isolated, separate attempts to cleave the esters using acetonitrile/H_2O and potassium hydroxide, were performed and the purified product again revealed hydrolysed benzoic acid and unhydrolysed ester.

7.8 - Fourth Attempt of Synthesis of 5,7-dihydroxy-2-(4-methylphenyl)chromen-4-one (6c).
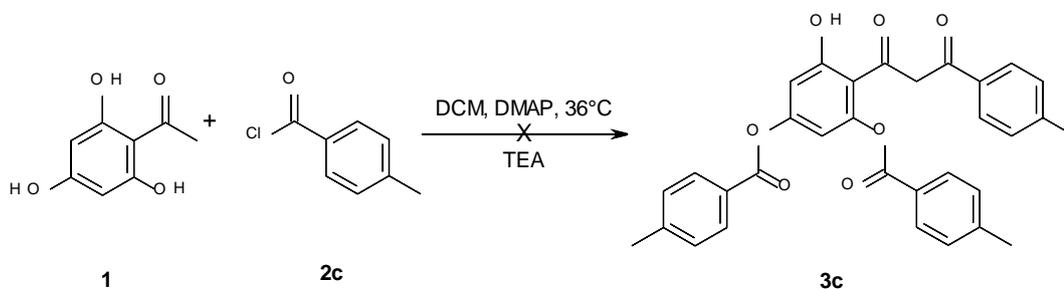
Scheme 7.3 - Unsuccesful synthesis of (**6c**) using alternated conditions.

As all attempts with the method proposed by Shabnam Saxena *et al* [154] were unsuccessful, a step-wise approach was taken, using dichloromethane, 4-dimethylaminopyridine and trimethylamine to form and isolate the ester intermediate, rather than telescope all steps into one process. However, the initial reaction (DCM, DMAP, TEA, rt→reflux) showed no change by TLC, and when pyridine was added and heat was applied, there was still no significant change, so this method was not taken

further, despite such conditions being successful on other compounds in the Snape group.

7.9 - Fifth Attempt of Synthesis of 5,7-dihydroxy-2-(4-methylphenyl)chromen-4-one (6c).

Successive failures at this reaction encouraged a new approach to be adopted whereby each step was closely monitored, isolated and purified where possible to determine where and why the literature procedure was failing. Therefore, reverting back to the method using dichloromethane as a solvent instead of toluene, once the reactants had fully reacted by TLC, the first product was analysed by ^{1}H NMR, which revealed that there was a large mixture of aromatics and methyl groups, which were expected due to the three arylmethyl groups on the intermediate, and the excess unreacted p-toluoyl chloride. So, the mixture was purified via flash column chromatography, which revealed six products. The fractions that were separable were isolated and analysed, revealing that one of the products was the ester (**5c**) seen previously, as identified by a single hydroxyl peak at 13.06 ppm, instead of the expected two peaks. The other two fractions gave inseparable mixtures and were not taken forward.

7.10 - Sixth Synthesis of 2-(4-Methylphenyl)-4H-chromen-4-one (6c).
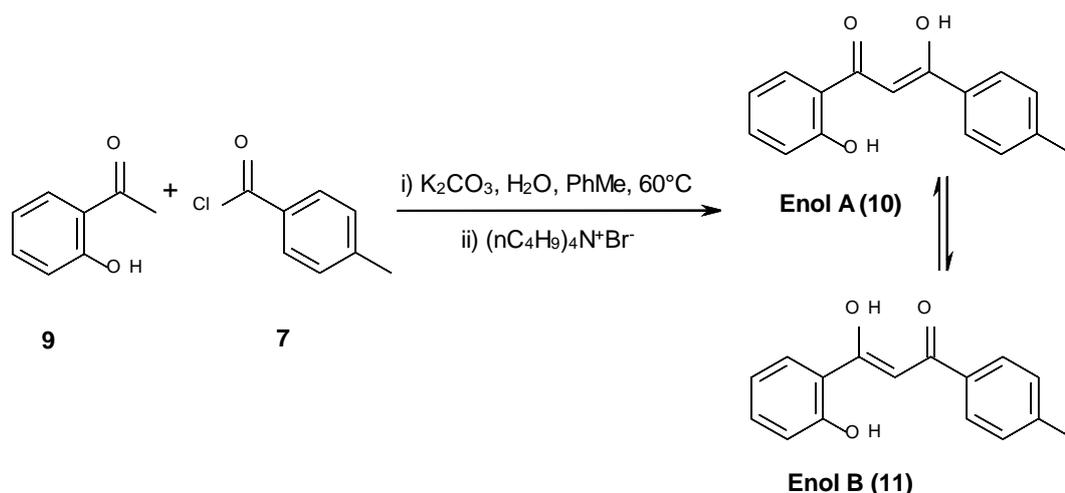
Scheme 7.4 - Synthesis of 5,7-dihydroxy-2-(4-methylphenyl)chromen-4-one via the Baker-Venkataraman rearrangement forming the two stabilised enols.

To confirm that the continued reaction reaction failure was due to the two extra hydroxyl groups in the starting matieral, and not due to human error, the synthesis was attempted again, but without the hydroxyl groups at positions 5 and 7 (Scheme 7.4), to form 2-(4-Methylphenyl)-4H-chromen-4-one (*c.f.* Figure 7.3). By using the same conditions, the diketone intermediate was formed, which stabilised itself as two enolic forms (enol A and enol B, see Scheme 7.4), which were confirmed by ^{1}H NMR. Pleasingly, these products confirm that these reaction conditions do work, and that it is the two hydroxyls that are interfering with the synthesis. Unfortunatley, the additional hydroxyl groups are essential for biological activity (*c.f.* Chapter 3), meaning that it was necessary to persevere with the failed syntheses. As such, the synthesis in Scheme 7.4 was not taken further as it lacked the essential hydroxyls, and much more promising progress had been made in parallel experiments, the details of which are shown in the following section.
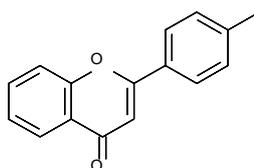
Figure 7.3 – Structure of 2-(4-Methylphenyl)-4H-chromen-4-one

7.11 - Alternative Synthesis of 5,7-dihydroxy-2-(4-methylphenyl)chromen-4-one (6c) Using Methyl Protecting Groups.
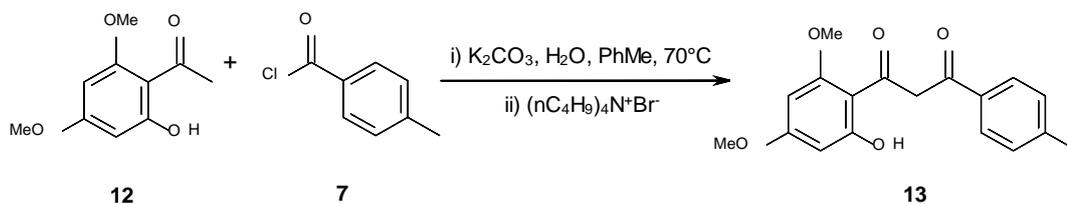
Scheme 7.5 - Synthesis of 5,7-dihydroxy-2-(4-methylphenyl)chromen-4-one via the Baker-Venkataraman rearrangement.

Unlike the previous methods that exploited an *in situ* protecting group strategy, which subsequently proved difficult to remove, this new method used methyl protecting groups on the hydroxyls, to prevent unnecessary esters forming on the 5 and 7 position, an approach that had been shown to be successful by Khanapur [155]. Using similar methods as previously performed, with toluene as solvent, the method did show promise. However, these trial reactions were only performed on a test scale meaning that sufficient material was not obtained at first attempt. Nevertheless, after analysis of the small scale reaction, promising results were observed. Analysis showed a mixture of desired intermediate (**13**) and unreacted starting material. Nevertheless, these promising results for the methylated starting material, meant we were encouraged to test alternative conditions using the methyl protected starting material.

7.12 - Alternative Synthesis of 5,7-dihydroxy-2-(3-methylphenyl)chromen-4-one (6b,6c) Using Methyl Protecting Groups.
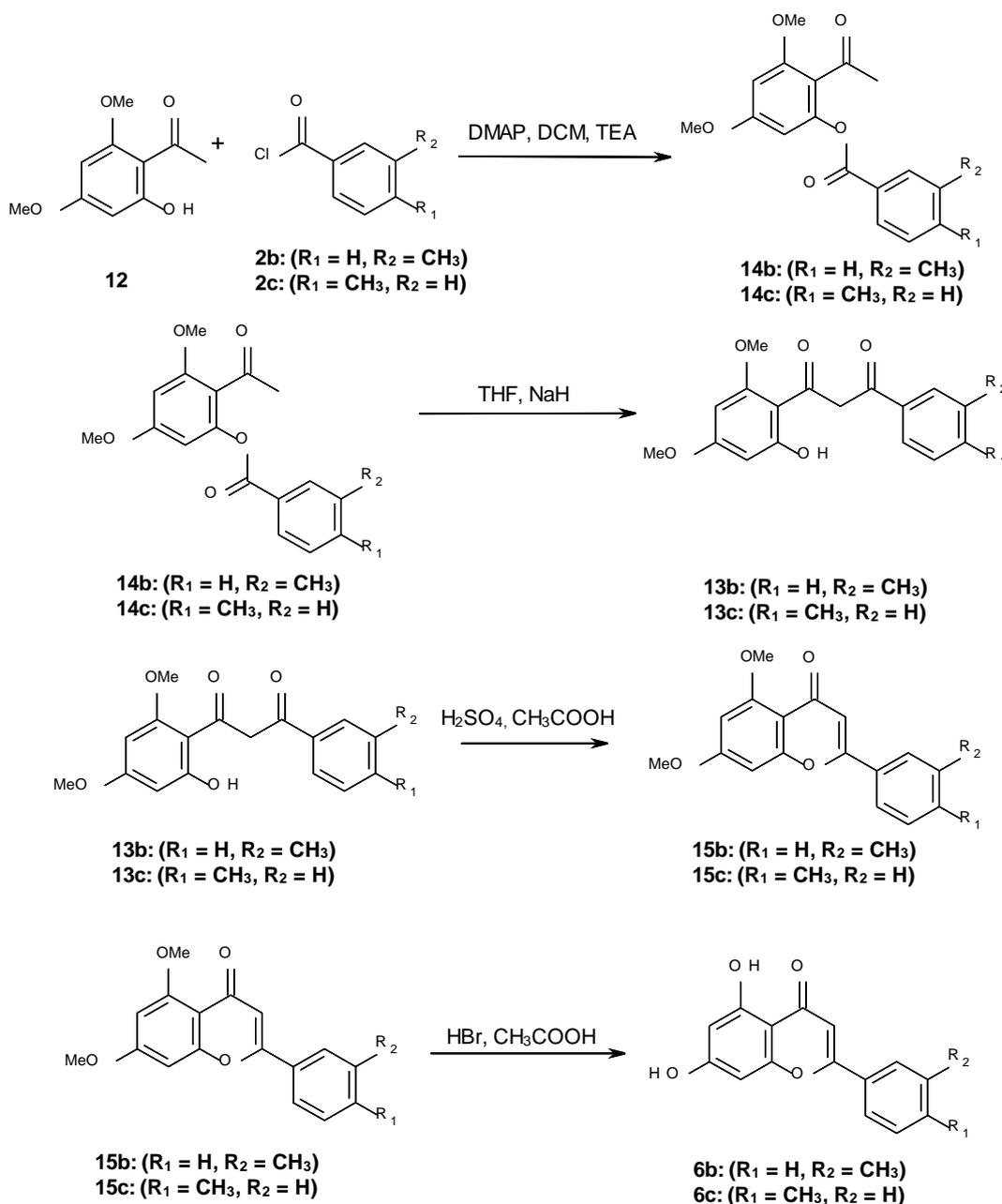
Scheme 7.6 – Synthesis of 5,7-dihydroxy-2-(3-methylphenyl)chromen-4-one (**6b-c**) via the Baker-Venkataraman rearrangement using the methyl protecting groups.

During the alternative method, it was decided we would stop at each stage of the process to isolate, purify and characterise the products, in attempts to confirm reactions and optomise the process, to later achieve higher yields, as seen in Scheme 7.6. This method of using protecting groups from the start, not *in situ* bypassed the problem of de-esterification that arose from the previous methods. However, methoxy groups are

notoriously difficult to remove and require harsh chemicals to do so, but by carrying out this step, the final compound was reached. This method also gave better results throughout, determined by in-process analysis, possibly due to the lack of *in situ* protecting groups as intermediates. Unfortunately, after column chromatography and also purification with analysis at each step, the yield for this metod was low, and did not leave enough material for initial de-methylation attempts.

7.13 - Second Synthesis of 5,7-dihydroxy-2-(3-methylphenyl)chromen-4-one (6b) Using Methyl Protecting Groups.

Using the same conditions as the previous attempt, the methylated flavone was created with high purity in 35% yield. Literature procedures stated that removal of methyl groups from a flavone skeleton could be achieved by the use of boron tribromide, [156] as such, this method was attempted with small amounts of the purified compound in anhydrous dichloromethane, (boron tribromide is extremely reactive in the presence of water, becoming deactivated). Multiple attempts showed little to no activity, so alternative methods were investigated. Ranu B. *et al* explained that a mixture of hydrobromic acid and glacial acetic acid can remove the methyl group on such a flavonoid core, [157] so this procedure was tested and showed extremely good results, which resulted in the highly pure desired compound in its fully deprotected state. Although the overall yield was low, it was decided to try this method again, as the analysis and purification methods may have affected the yield. Due to the success of the meta-substituted analogue we also decided to synthesise the para substituent.

7.14 - Conclusion

The synthesis of chrysin analogues proved more complicated than first expected, despite the recorded literature. However, a viable synthetic route for flavones which contain hydroxyl groups at the 5 and 7 position was finally discovered by exploiting methyl protecting groups, which proved to be more effective than using *in situ* esters protecting groups as first thought – despite the notorious difficulty in removing methyl protecting groups during natural product synthesis. The newly discovered route did succesfully lead to the synthesis of two analogues with hydrophobic methyl groups on

ring B, and theoretically, the same methods can be used for hydroxyl and halogenated groups, as these show promise from the results of the computational studies (*c.f.* Chapter 6). These synthesised compounds will now be taken forward for kinetic experimental analysis to compare the computational results to the actual activity of the compounds. Then using this data, a refined search for the ideal ligand can be initiated, using the synthetic methods described and optimised above to create more potent compounds that bind to GP.

7.15 – Experimental

7.15.1 – Overview of Materials and Methods

Reactions were followed by analytical thin layer chromatography (TLC) using plastic-backed TLC plates coated in silica G/UV254, run in a variety of solvent systems and visualised with a UV light at 254 nm, p-anisaldehyde stain and/or potassium permanganate stain. Commercially available solvents and reagents were purchased from Fisher, Sigma Aldrich, TCI and Fluorochem and were used without further purification unless specified in the syntheses. Flash column chromatography was carried out on Davisil silica 60 Å (40 – 63 µm) under bellows pressure. ^{1}H and ^{13}C NMR were carried out on a Bruker Fourier 300 (300 MHz). GCs ran with DSQ II, Trace GC Ultra, Triplus, trace1800 GC & ITQ700. LCs ran with Finnigan LCQ Advantage Max, Finnigan PDA Plus detector. Infrared spectra were recorded on a solid sample using a Thermo Nicolet IR-200 FT-IR.

7.15.2 - 5-Hydroxy-4-oxo-2-phenyl-4H-chromen-7-yl benzoate
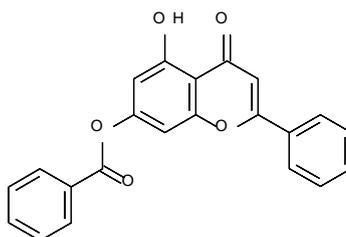
5a

2,4,6-Trihydroxyacetophenone (**1**) (1 g, 5.37 mmol), was added to benzoyl chloride (**2**) (2.49 ml, 21.49 mmol) in a mixture of toluene (40 ml), saturated potassium carbonate (40 ml) and the reaction stirred at 60 °C for 1 hour, and monitered by TLC of 1:3 ethyl acetate: petroleoum ether. Tetrabutylammonium bromide (2.6 g, 8.06 mmol) was added and the mixture stirred at 60 °C for 3 hours, upon which an orange solution of (**3a**) was formed. The reaction was left to cool. To this solution, 5% K_2CO_3 (25 ml) was added and the mixture stirred at reflux (110 °C) for 4 hours. A dark orange solid formed in the orange liquid. This liquid was decanted off and added to glacial acetic acid (50 ml) with heat and stirring for 10 minutes. The solution was added to ice and a cream coloured solid was formed, which was washed with petroleum ether and water. The product was purified by flash column chromatography (1:3 Ethyl Acetate: petroleum ether) to yield compound (**5a**) (*c.f.* Chapter 7, Fig 7.2) (0.20g, 10.4%). ^{1}H NMR (300 MHz, DMSO): δ_H= 12.36 (1H, s, OH), 7.95 (1H, s, Ar CH), 7.92 (1H, s, Ar CH), 6.47 (1H, d, J = 3, Ar CH), 6.31 (1H, d, J = 3, Ar-CH), 7.66-7.54 (3H, m, Ar CH), 7.48-7.33 (5H, m, Ar CH). ^{13}C NMR (75 MHz, DMSO): δ_C = 192.91 (OC=C), 180.25 (C=O), 165.66 (Ar CO), 161.86 (Ar C-C), 158.06 (Ar C-O), 136.97 (Ar C-C), 134.70 (Ar C-C), 132.14 (Ar CH), 131.59 (Ar CH), 129.72 (Ar CH), 129.50 (Ar CH), 129.32 (Ar CH), 128.75 (Ar CH), 120.62 (Ar C-C), 103.74 (Ar CH), 99.90 (Ar CH), 94.88 (Ar CH). MS (ESI): *m/z* 359.1 ([M+H]$^+$).

7.15.3 - 5,7-dihydroxy-2-(3-methylphenyl)chromen-4-one
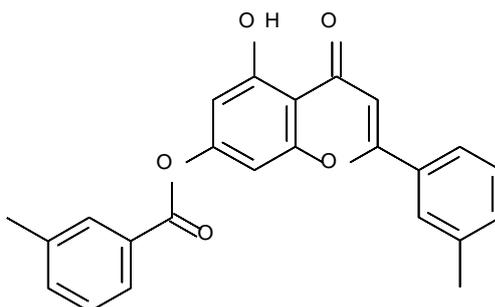
5b

2,4,6-Trihydroxyacetophenone (**1**) (1 g, 5.37 mmol) was added to 3-methylbenzoyl chloride (**2b**) (2.84 ml, 21.49 mmol) in a mixture of toluene (40 ml) and saturated potassium carbonate (40 ml), and the reaction stirred at 60 °C overnight. Tetrabutylammonium bromide (2.6 g, 8.06 mmol) was then added and the mixture stirred at 60 °C for 6 hours, and monitered by TLC, upon which an orange solution of (**7**) was formed. The reaction was left to cool over-night. The organic layer was separated and washed with 2 x 20 ml of distilled water. The interface and aqueous layer was washed with 2 x 20 ml of ethyl acetate, 2 x 20 ml of distilled water and 20 ml of brine. The organic layers evaporated, and 5% K_2CO_3 (50 ml) was added and the reaction heated at reflux with stirring for 2 hours to form (**8**), which was decanted off under suction filtration. The solid was transferred to a conical flask and boiled in glacial acetic acid (50 ml) for 30 minutes. Water (50 ml) was added, turning solution creamy and creating a precipitate, which was suction filtered and dried over night, then washed with ethyl acetate to yield the final compound (**5b**) (1.31 g, 97.58%), which was purified by flash column chromatography in 5:2 ethyl acetate: petroleum ether to yield **5b**. 5ml of MeOH with K_2CO_3 (100 mg) was added to **8** and the attempted hydrolysis stirred at 30 °C for 1 hour. TLC showed no change, so more K_2CO_3 (100 mg) was added and left for another hour. TLC showed no change. (**5b**) = [^1H] NMR (300 MHz, CDCl$_3$): δ_H= 12.37 (1H, s, OH), 7.73 (2H, Ar CH), 7.37 (4H, Ar CH), 7.22 (1H, Ar CH), 6.47 (1H, d, J = 3, Ar CH), 6.30 (1H, d, J = 3, Ar CH), 6.03 (1H, Ar CH), 2.37 (3H, s, CH$_3$), 2.31 (3H, s, CH$_3$). [^13C] NMR (75 MHz, DMSO): δ_C = 192.52 (OC=C), 179.89 (C=O), 165.03 (Ar C-C), 162.25 (Ar C-O), 161.46 (Ar C-O), 157.61 (Ar C-O), 138.54 (Ar C-C), 138.29 (Ar CH), 136.69 (Ar C-C), 134.90 (Ar C-C), 132.32 (Ar-CH), 131.16 (Ar CH), 129.37 (Ar CH), 128.92 (Ar CH), 128.73 (Ar CH), 128.70 (Ar CH), 126.63 (Ar CH), 125.53 (Ar C-

C) 120.31 (Ar CH), 103.40 (Ar C-C), 103.40 (Ar CH), 99.40 (Ar CH), 94.40 (Ar C-C), 20.90 (Ar-CH$_3$), 20.75 (Ar-CH$_3$).

7.15.4 - p-Toluic acid

General procedure:
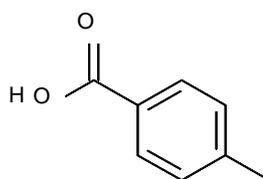
8

2,4,6-Trihydroxyacetophenone (**1**) (1.8 g, 9.67 mmol) and (**2c**) p-toluoyl chloride (2 ml, 15 mmol) were added to 50 ml of pyridine. The solution was stirred for 2 hours and poured over crushed ice with 3% HCl (20 ml) and stirred for 1 hour. The reaction was suction filtered and washed with methanol (2 x 20ml) and distilled water (2 x 20 ml) resulting in a white slurry. The aqueous layer was also washed with diethyl ether (2 x 20 ml). These washing were combined with the white slurry and washed with 1 M HCl (2 x 50 ml), distilled water (2 x 50 ml) and brine (50 ml) then dried with MgSO$_4$. The resulting white powder was dissolved in pyridine with stirring at 55 °C, then 1 g of potassium hydroxide was added which turned the solution yellow. 10% acetic acid (10 ml) was added with stirring for 1 hour, turning the solution dark. Rotary evaporation if the reaction yielded a yellow solid. Sulphuric acid (25 ml) was added with glacial acetic acid (25 ml) and the reaction stirred for 1 hour. The reaction was washed with water and dried with suction filtration, leaving tan coloured compound. The product was isolated by flash column chromatography to yield p-toluic acid (**8**) (40 %) ^{1}H NMR (300 MHz, CDCl$_3$): δ_H= 12.55 (1H, s, OH), 8.03 (2H, d, J = 9, Ar-H), 7.29 (2H, d, J = 9, Ar-H), 2.45 (3H, s, CH$_3$). Mp: 176-180 °C (lit: 177-180 °C)

7.15.5 – General procedure for the synthesis of B-ring analogues of chrysin
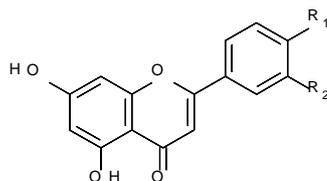
6b: (R₁ = H, R₂ = CH₃)
6c: (R₁ = CH₃, R₂ = H)

7.15.5.1 – Procedure for the synthesis of 14b and 14c.
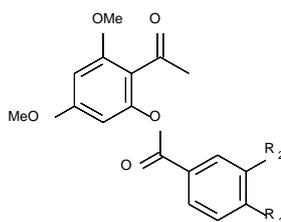
14b: (R₁ = H, R₂ = CH₃)
14c: (R₁ = CH₃, R₂ = H)

2′-Hydroxy-4′,6′-dimethoxyacetophenone (1.0 eq.) was added to dichloromethane. m-Toluoyl chloride (b) and p-toluoyl chloride (c) (1.1 eq.) was then added to the mixture with 4-(dimethylamino)pyridine (0.1 eq.) and triethylamine (1.5 eq.). This solution was stirred vigorously at room temperature until deemed complete by a TLC of petroleum, ether and ethyl acetate (3:1). The product was washed with hydrochloric acid (3 x 25 ml) and extracted with dichloromethane (2 x 25 ml). The organic layers were combined and washed with distilled water (2 x 25 ml) and brine. The solution was evaporated *in vacuo* leaving an oil, which was dried *in vacuo* for 48 hours.

14b

Chromatography solvent, 1:3 ethyl acetate: petroleum ether. ^{1}H NMR (300 MHz, CDCl$_3$): δ_H= 7.94 (2H, Ar-H), 7.39 (2H, Ar-H), 6.40 (1H, AR-H), 6.35 (1H, Ar-H), 3.87 (3H, O-CH$_3$), 3.82 (3H, O-CH$_3$), 2.48 (3H, C-CH$_3$), 2.42 (3H, Ar-CH$_3$). MS (ESI): *m/z* 315.2 ([M+H]$^+$).

14c

Chromatography solvent, 1:3 ethyl acetate: petroleum ether. ^{1}H NMR (300 MHz, CDCl$_3$): δ_H= 7.48 (2H, Ar-H), 7.26 (2H, Ar-H), 6.38 (1H, AR-H), 6.36 (1H, Ar-H), 3.81 (3H, O-CH$_3$), 3.77 (3H, O-CH$_3$), 2.47 (3H, C-CH$_3$), 2.40 (3H, Ar-CH$_3$). MS (ESI): *m/z* 315.3 ([M+H]$^+$).

7.15.5.2 – General procedure for 13b and 13c
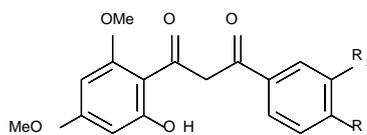
13b: (R$_1$ = H, R$_2$ = CH$_3$)
13c: (R$_1$ = CH$_3$, R$_2$ = H)

Tetrahydrofuran was added with 4 Å molecular sieves to a dry round bottomed flask and stirred for 2 hours at 80 °C. 14b/14c was then added. Sodium hydride (4.0 eq.) was added and heated to 90 °C for 3 hours with stirring and left to cool over night. The mixture was then filtered and washed with hydrochloric acid until neutral and ethyl acetate (2 x 25 ml). The organic layer was washed with distilled water (2 x 25 ml) and brine (25 ml) and dried with MgSO$_4$. This was then evaporated *in vacuo* resulting in a black oil.

13b

Chromatography solvent, 1:5 ethyl acetate: petroleum ether. ^{1}H NMR (300 MHz, CDCl$_3$): δ$_H$= 15.54 (1H, Ar-OH), 6.47 (1H, AR-H), 6.44 (1H, Ar-H), 7.98-7.28 (4H, m, Ar-H), 3.92 (3H, O-CH$_3$), 3.81 (3H, O-CH$_3$), 3.45 (2H, C-CH$_3$), 2.43 (Ar-CH$_3$). MS (ESI): *m/z* 314.3 ([M+H]$^+$).

13c

Chromatography solvent, 1:5 ethyl acetate: petrololeum ether. ^{1}H NMR (300 MHz, CDCl$_3$): δ$_H$= 15.61 (1H, Ar-OH), 8.04-7.19 (4H, m, Ar-H), 6.47 (1H, AR-H), 6.45 (1H, Ar-H), 3.88 (3H, O-CH$_3$), 3.82 (3H, O-CH$_3$), 3.47 (2H, C-CH$_3$), 2.42 (Ar-CH$_3$). MS (ESI): *m/z* 314.2 ([M+H]$^+$).

7.15.5.3– General procedure for 15b and 15c

15b: (R$_1$ = H, R$_2$ = CH$_3$)
15c: (R$_1$ = CH$_3$, R$_2$ = H)
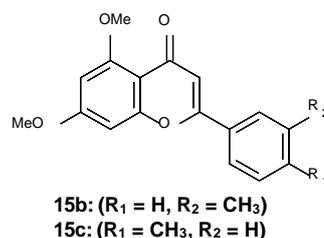
Glacial acetic acid (100 ml per 1g of 2′-hydroxy-4′,6′-dimethoxyacetophenone) was added to 13b/13c along with concentrated sulphuric acid (2 ml per 1g of 2′-hydroxy-4′,6′-dimethoxyacetophenone) and stirred at 110 °C for 4 hours. The hot solution was then poured over crushed ice and poured into a separating funnel. The solution was washed with sodium hydroxide (1 M) and ethyl acetate (2 x 25 ml), then water (2 x 25ml) and brine (50ml). The resulting solid was purified by flash column chromatography using a solution of 1:9 ethyl acetate: petroleum ether.

15b

Chromatography solvent, 1:3 ethyl acetate: petroleum ether. ^{1}H NMR (300 MHz, CDCl$_3$): δ_H= 7.67 (2H, Ar-H), 7.34 (2H, Ar-H), 6.67 (1H, AR-H), 6.58 (1H, Ar-H), 6.38 (1H, Ar-H), 3.96 (3H, O-CH$_3$), 3.92 (3H, O-CH$_3$), 2.44 (3H, Ar-CH$_3$). MS (ESI): *m/z* 297.1 ([M+H]$^+$).

15c

Chromatography solvent, 1:3 ethyl acetate: petroleum ether. ^{1}H NMR (300 MHz, CDCl$_3$): δ_H=7.98 (1H, Ar-H), 7.77 (1H, Ar-H), 7.30 (2H, Ar-H), 6.69 (1H, AR-H), 6.57 (1H, Ar-H), 6.38 (1H, Ar-H), 3.95 (3H, O-CH$_3$), 3.91 (3H, O-CH$_3$), 2.43 (3H, Ar-CH$_3$). MS (ESI): *m/z* 297.2 ([M+H]$^+$).

7.15.5.4 – General procedure for synthesis of 6b and 6c
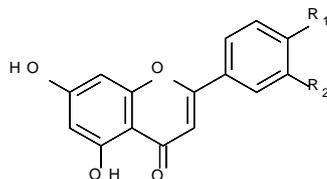
6b: (R₁ = H, R₂ = CH₃)
6c: (R₁ = CH₃, R₂ = H)

The purified solid of 15b/15c was added to a round bottomed flask with hydrobromic acid (15 ml per 1g of 2′-hydroxy-4′,6′-dimethoxyacetophenone) and glacial acetic acid (15 ml per 1g of 2′-hydroxy-4′,6′-dimethoxyacetophenone) and stirred at 180 °C for 18 hours to yield the final demethylated product, which was washed with sodium hydroxide until neutral, ethyl acetate (2 x 25 ml), 50 ml water and 50 ml brine, then dried with MgSO₄. Rotary evaporation yielded the title compound as a pure product.

5,7-Dihydroxy-2-(3-methylphenyl)chromen-4-one (6b).

(0.14 g, 10.26 %) Chromatography solvent, 1:9 ethyl acetate: petroleum ether. ^{1}H NMR (300 MHz, CDCl₃): δ_H= 12.85 (1H, s, OH), 10.96 (1H, s, OH), 7.88 (2H, d, $J = 9$, Ar-H), 7.45 (2H, d, $J = 6$, Ar-H), 6.95 (1H, s, Ar-H), 6.52 (1H, s, Ar-H), 6.22 (1H, s, Ar-H), 2.41 (3H, s, CH₃). ^{13}C NMR (75 MHz, DMSO): δ_C = 181.90 (C=O), 164.45 (Ar, C-C), 163.39 (Ar, C-OH), 161.48 (Ar, C-OH), 157.49 (Ar, C-O), 138.64 (Ar, C-C), 132.74 (Ar, C-C), 130.67 (Ar, CH), 129.07 (Ar, CH), 126.84 (Ar, CH), 123.65 (Ar, CH), 105.13 (Ar, C-C), 103.98 (Ar, CH), 99.03 (Ar, CH), 94.16 (Ar, CH), 20.98 (CH₃). MS (ESI): m/z 269.2 ([M+H]⁺). IR (neat, cm⁻¹) v = 1649 (C=O), 1100 (O-H), 1117 (O-H), 1182 (C-O-C). mp: 231-234 °C

5,7-Dihydroxy-2-(4-methylphenyl)chromen-4-one. (6c)

(0.27 g, 19.79 %) Chromatography solvent, 1:7 ethyl acetate: petroleum ether. ^{1}H NMR (300 MHz, CDCl₃): δ_H= 12.27 (1H, s, OH), 10.91 (1H, s, OH), 7.97 (2H, d, $J = 8$, Ar-H), 7.38 (2H, d, $J = 8$, Ar-H), 6.93 (1H, s, Ar-H), 6.51 (1H, s, Ar-H), 6.21 (1H, s, Ar-

H), 2.39 (3H, s, CH$_3$). ^{13}C NMR (75 MHz, DMSO): δ_C = 181.89 (C=O), 164.36 (Ar, C-C), 163.36 (Ar, C-OH), 161.47 (Ar, C-OH), 157.42 (Ar, C-O), 142.37 (Ar, C-C), 129.76 (Ar, C-C), 127.91 (Ar, CH), 126.39 (Ar, CH), 104.49 (Ar, C-C), 103.93 (Ar, CH), 98.98 (Ar, CH), 94.12 (Ar, CH), 21.12 (CH$_3$). MS (ESI): *m/z* 269.1 ([M+H]$^+$). IR (neat, cm^{-1}) v = 1653 (C=O), 1128 (O-H), 1117 (O-H), 1188 (C-O-C). mp: 238-243 °C

Chapter 8 - Conclusion and Future Work

Diabetes is one of the largest global health problems of the 21st century, with over 400 million people suffering from hyperglycemia in some way, and with type 2 diabetes being the most prelevant (91% of diabetes in adults in high income countries [158]). Although there are currently drugs on the market which aim to make the disease manageable, such as metformin, glibenlamide, glimepiride and repaglinide, alongside a healthy life style, there are considerable side-effects associated with the current drugs including a risk of hypoglycaemia. Hence, there is an immediate need for new, more effective treatments to make the lives of patients easier and healthier. Inhibition of glycogen phosphorylase, especially at the caffeine binding site, is an extremely promising approach to finding a better treatment, as this site can prevent glucose production when blood glucose levels are too high, with inhibition at this site being more potent under high blood glucose concentrations, whilst being less potent at lower blood glucose levels. This balance can prevent hypoglycaemia in patients, which is a common risk inolved in current treatments.

We have exploited computational methods such as docking with Glide SP and XP, which generated poses for further claculations. These were then followed by post docking calculations such as Prime and QM/MM-PBSA calculations, which minimised the poses, and accurately scores the ligands and predicts their free binding energy, as described throughout this project. By using these, we have efficiently screened a large library of possibly active and novel compounds, which have the potential to bind at the inhibitor site of glycogen phosphorylase. After multiple screening, filtering and consensus scoring techniques have been used, the computational simulations yielded a total of 13 ligands which are predicted to perform better at GP inhibition than the benchmark ligand chrysin. All 13 ligands can now be taken forward for synthesis via the Baker-Venkataraman re-aarrangement by manipulating methyl protected hydroxyl groups. The synthesised ligands will then undergoe further testing of their inhibitory function in Greece via kinetic eperiments. The original synthetic process did unfortunately take longer to optimise than expected, due to complications with de-esterification on the A ring of the flavone in the final stages, which were formed *in situ*. Nonetheless, the method has finally been altered and refined to create the desired flavone structure with the hydroxyls left intact on the 5 and 7 position of the A ring by

using methoxy protecting groups, rather than *in situ* protection methods. This procedure can now be taken forward and adapted to synthesise many of the investigated chrysin analogues which have been discovered via the computational methods. It is expected that the conditions throughout the refined synthesis will allow halogens and alcohols to be incorporated with minimal adaptation to the method. This is important, as the post docking QM/MM-PBSA calculations seemed to favour halogens and hydroxyls at the ortho position of the ring, due to binding interactions with the residue glutamate 572. The synthetic procedure will, in theory, be viable for all ligands produced, as this was taken into consideration during the initial screening of ligands. The final compounds should also be viable as the halogens will not be reactive and the hydroxyls can be synthesised via methyl protecting groups in the same way as the hydroxyls in the 5 and 7 position of the A ring. Though there will be epected alterations made to the synthetic route, such as the need to possibly allow more time for certain stages, changes to TLC and flash column chromatography solvent systems and potentially minor ratio changes, which can all be discovered in the lab, using simple observational methods and intuition.

Combining the methods used throughout this project has considerable potential for the discovery of novel and active compounds in the treatment of type 2 diabetes, as thousands of ligands can be rapidly screened for their and potential, then selective synthesis of the best ligands can be implemented. Extension of this work has an exciting and promising future. With the refinement of the synthetic process, it will now be far easier to create the desired ligands which are predicted to have high binding affinities by the computational predictions. The computational results have suggested ortho-substituted halogens and hydroxyl groups in the B ring, which would be considered first. Then, using kinetic experimental data from the synthesised compounds, we can then feed the information about what works well and what does not, back into the computational data to refine the results even further until we can successfully design more potent and drug-like inhibitors of glycogen phosphorylase. This structure activity relationship process can additionally be used to investigate how changes to the A ring affect binding in combination with the results we have for the B ring, until eventually we find the most optimal substitution combinations for the flavone to bind strongest. Further work would then involve pharmacokinetic and *in vivo* studies with the ultimate aim to achieve a drug that performs better than currently available treatments for this worldwide healthcare problem affecting the lives of many.

References

1.   Seuring, T., O. Archangelidi, and M. Suhrcke, *The Economic Costs of Type 2 Diabetes: A Global Systematic Review.* Pharmacoeconomics, 2015. **33**(8): p. 811-31.

2.   Prevention, C.f.D.C.a., *National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States, 2014*, in *National Diabetes Statistics Report*. 2014, US Department of Health and Human Services: Atlanta, GA.

3.   Aguiree, F., et al., *IDF Diabetes Atlas.* 2013: p. 155.

4.   Kahn, B.B., *Type 2 diabetes: when insulin secretion fails to compensate for insulin resistance.* Cell, 1998. **92**: p. 593-596.

5.   NHS. *Hypoglycaemia (low blood sugar)*. 2015 [cited 2015 22/09/2015]; Available from: http://www.nhs.uk/conditions/hypoglycaemia/Pages/Introduction.aspx

6.   Saltiel, A.R. and C.R. Kahn, *Insulin signalling and the regulation of glucose and lipid metabolism.* Nature, 2001. **414**(December): p. 799-806.

7.   Alberti, K.G. and P.Z. Zimmet, *Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation.* Diabetic medicine : a journal of the British Diabetic Association, 1998. **15**: p. 539-553.

8.   UK, D. *Differences Between Type 1 and Type 2*. 2014 30/09/15]; Available from: http://www.diabetes.co.uk/difference-between-type1-and-type2-diabetes.html

9.   UK, D. *Gestational Diabetes*. 2014 07/08/2014 30/09/15]; Available from: http://www.nhs.uk/Conditions/gestational-diabetes/Pages/Introduction.aspx.

10. Bellamy, L., et al., *Type 2 diabetes mellitus after gestational diabetes: a systematic review and meta-analysis.* Lancet, 2009. **373**(9677): p. 1773-1779.

11. Holt, T.K., S., *ABC of Diabetes*. 6th ed. 2013: Wiley-Blackwell.

12. NHS. *Type 2 Diabetes- Symptoms*. 2014 18/06/14 24/09/15]; Available from: http://www.nhs.uk/Conditions/Diabetes-type2/Pages/Symptoms.aspx

13. Guide, D.T. *Four Common Signs And Symptoms Of Progressive Diabetes*. 2014 30/09/15]; Available from: http://www.diabetestreatmentguide.org/four-common-signs-and-symptoms-of-progressive-diabetes/

14. UK, D., *State Of The Nation (England): Challenges For 2015 And Beyond.* 2008: p. 36.

15. William, G.P., J. C., *Handbook of Diabetes*. 3rd ed. 2004: Blackwell Publishing Ltd. 275.

16. *National Diabetes Audit Mortality Analysis.* 2015: p. 37.

17. Group, D.P.P.R., *The Diabetes Prevention Program: Description of lifestyle intervention.* Diabetes Care, 2002. **25**: p. 2165-2171.

18. Kemp, D.M., *Type 2 Diabetes: Disease Overview*, in *New Therapeutic Strategies for Type 2 Diabetes*, R.M. Jones, Editor. 2012, Royal Society of Chemistry Publishing. p. 1-14.

19. Rubino, F., et al., *The early effect of the Roux-en-Y gastric bypass on hormones involved in body weight regulation and glucose metabolism.* Annals of Surgery, 2004. **240**(2): p. 236-242.

20. Rubino, F.A. and J. Marescaux, *Effect of duodenal-jejunal exclusion in a non-obese animal model of type 2 diabetes - A new perspective for an old disease.* Annals of Surgery, 2004. **239**(1): p. 1-11.

21. Viollet, B., et al., *Cellular and molecular mechanisms of metformin: an overview.* Clinical Science, 2012. **122**(5-6): p. 253-270.

22. Sundriyal, S., et al., *Metformin and glitazones: does similarity in biomolecular mechanism originate from tautomerism in these drugs?* Journal of Physical Organic Chemistry, 2008. **21**: p. 30-33.

23. Chemspider. *Glibenclamide*. 2015 05/10/15]; Available from: http://www.chemspider.com/Chemical-Structure.3368.html.

24. PubChem. *gliclazide*. 2015 05/10/15]; Available from: https://pubchem.ncbi.nlm.nih.gov/compound/3475#section=Top.

25. PubChem. *Glimepiride*. 2015 05/10/15]; Available from: https://pubchem.ncbi.nlm.nih.gov/compound/3476.

26. PubChem. *Repaglinide*. 2015 05/10/15]; Available from: https://pubchem.ncbi.nlm.nih.gov/compound/65981.

27. Landgraf, R., *Meglitinide analogues in the treatment of type 2 diabetes mellitus.* Drugs & Aging, 2000. **17**(5): p. 411-425.

28. Black, C., et al., *Meglitinide analogues for type 2 diabetes mellitus (Review).* Cochrane Database of Systematic Reviews, 2007(2).

29. McCulloch, D.K. *Patient information: Diabetes mellitus type 2: Insulin treatment (Beyond the Basics)*. 2015 05/10/15 07/10/15]; Available from: http://www.uptodate.com/contents/diabetes-mellitus-type-2-insulin-treatment-beyond-the-basics

30. Jedinak, A., et al., *Approaches to flavonoid production in plant tissue cultures.* Biologia, 2004. **59**(6): p. 697-710.

31. Diabetes, C. *Glycogenesis*. 2014 [cited 2015 28/09/15]; Available from: http://www.checkdiabetes.org/glycogenesis/.

32. Diabetes, C. *Gluconeogenesis*. 2014 28/09/15]; Available from: http://www.checkdiabetes.org/gluconeogenesis/

33. Szablewski, L., *Glucose Homeostasis and Insulin Resistance.* Glucose Homeostasis and Insulin Resistance, 2011: p. 211.

34. Beamer, L., *Mutations in hereditary phosphoglucomutase 1 deficiency map to key regions of enzyme structure and function.* Journal of Inherited Metabolic Disease, 2015. **38**(2): p. 243-256.

35. Barford, D., S.-H. Hu, and L.N. Johnson, *Structural mechanism for glycogen phosphorylase control by phosphorylation and AMP.* Journal of Molecular Biology, 1991. **218**: p. 233-260.

36. Hayes, J.M., A.L. Kantsadi, and D.D. Leonidas, *Natural products and their derivatives as inhibitors of glycogen phosphorylase: Potential treatment for type 2 diabetes.* Phytochemistry Reviews, 2014. **13**: p. 471-498.

37. Obel, L.F., et al., *Brain glycogen—new perspectives on its metabolic function and regulation at the subcellular level.* Frontiers in Neuroenergetics, 2012. **4**: p. 1-15.

38. Agius, L., *Physiological Control of Liver Glycogen Metabolism: Lessons from Novel Glycogen Phosphorylase Inhibitors.* Mini-Reviews in Medicinal Chemistry, 2010. **10**(12): p. 1175-1187.

39. Kantsadi, A.L., et al., *The s-Hole Phenomenon of Halogen Atoms Forms the Structural Basis of the Strong Inhibitory Potency of C5 Halogen Substituted Glucopyranosyl Nucleosides towards Glycogen Phosphorylase b.* Chemmedchem, 2012. **7**(4): p. 722-732.

40. Johnson, L.N., *GLYCOGEN-PHOSPHORYLASE - CONTROL BY PHOSPHORYLATION AND ALLOSTERIC EFFECTORS.* Faseb Journal, 1992. **6**(6): p. 2274-2282.

41. Somsak, L., et al., *New Inhibitors of Glycogen Phosphorylase as Potential Antidiabetic Agents.* Current Medicinal Chemistry, 2008. **15**(28): p. 2933-2983.

42. Hayes, J.M., A.L. Kantsadi, and D.D. Leonidas, *Natural products and their derivatives as inhibitors of glycogen phosphorylase: potential treatment for type 2 diabetes.* Phytochemistry Reviews, 2014. **13**(2): p. 471-498.

43. Hayes, J.M. and D.D. Leonidas, *Computation as a Tool for Glycogen Phosphorylase Inhibitor Design.* Mini-Reviews in Medicinal Chemistry, 2010. **10**(12): p. 1156-1174.

44. Zimmet, P.Z., et al., *Diabetes: a 21st century challenge.* Lancet Diabetes Endocrinol, 2014. **2**(1): p. 56-64.

45. Pasupuleti, V.K. and J.W. Anderson, *Nutraceuticals, glycemic health and type 2 diabetes*. 1st ed. ed. 2008, Oxford: Wiley-Blackwell.

46. Hellerstein, M.K., et al., *Hepatic gluconeogenic fluxes and glycogen turnover during fasting in humans. A stable isotope study.* J Clin Invest, 1997. **100**(5): p. 1305-19.

47.   Wagman, A.S. and J.M. Nuss, *Current therapies and emerging targets for the treatment of diabetes.* Curr Pharm Des, 2001. **7**(6): p. 417-50.

48.   Somsak, L., et al., *New inhibitors of glycogen phosphorylase as potential antidiabetic agents.* Curr Med Chem, 2008. **15**(28): p. 2933-83.

49.   Oikonomakos, N.G., *Glycogen phosphorylase as a molecular target for type 2 diabetes therapy.* Curr. Protein. Pept. Sci., 2002. **3**(6): p. 561-86.

50.   Treadway, J.L., P. Mendys, and D.J. Hoover, *Glycogen phosphorylase inhibitors for treatment of type 2 diabetes mellitus.* Expert Opin. Investig. Drugs, 2001. **10**(3): p. 439-54.

51.   Agius, L., *Physiological control of liver glycogen metabolism: lessons from novel glycogen phosphorylase inhibitors.* Mini Rev Med Chem, 2010. **10**(12): p. 1175-87.

52.   Furukawa, S., et al., *FR258900, a novel glycogen phosphorylase inhibitor isolated from Fungus No. 138354. II. Anti-hyperglycemic effects in diabetic animal models.* J Antibiot (Tokyo), 2005. **58**(8): p. 503-6.

53.   Hoover, D.J., et al., *Indole-2-carboxamide inhibitors of human liver glycogen phosphorylase.* J.Med. Chem., 1998. **41**: p. 2934-2938.

54.   Martin, W.H., et al., *Discovery of a human liver glycogen phosphorylase inhibitor that lowers blood glucose in vivo.* Proc. Natl. Acad. Sci. U.S.A., 1998. **95**: p. 1776-1781.

55.   Ong, K.C. and H.E. Khoo, *Effects of myricetin on glycemia and glycogen metabolism in diabetic rats.* Life Sci, 2000. **67**(14): p. 1695-705.

56.   Ahmad, M., et al., *Hypoglycaemic action of the flavonoid fraction of Cuminum nigrum seeds.* Phytother Res, 2000. **14**(2): p. 103-6.

57.   Anila, L. and N.R. Vijayalakshmi, *Beneficial effects of flavonoids from Sesamum indicum, Emblica officinalis and Momordica charantia.* Phytother Res, 2000. **14**(8): p. 592-5.

58.   Middleton, E., Jr., C. Kandaswami, and T.C. Theoharides, *The effects of plant flavonoids on mammalian cells: implications for inflammation, heart disease, and cancer.* Pharmacol Rev, 2000. **52**(4): p. 673-751.

59.   Nijveldt, R.J., et al., *Flavonoids: a review of probable mechanisms of action and potential applications.* American Journal of Clinical Nutrition, 2001. **74**(4): p. 418-425.

60.   Rimm, E.B., et al., *Relation between intake of flavonoids and risk for coronary heart disease in male health professionals.* Annals of Internal Medicine, 1996. **125**(5): p. 384-389.

61.   Rossi, M., et al., *Flavonoids and colorectal cancer in Italy.* Cancer Epidemiology Biomarkers & Prevention, 2006. **15**(8): p. 1555-1558.

62.   Scalbert, A. and G. Williamson, *Dietary intake and bioavailability of polyphenols.* Journal of Nutrition, 2000. **130**(8): p. 2073s-2085s.

63.   Kinoshita, T., Z. Lepp, and H. Chuman, *Construction of a novel database for flavonoids.* J Med Invest, 2005. **52 Suppl**: p. 291-2.

64.   Peterson, J. and J. Dwyer, *An informatics approach to flavonoid database development.* J. Food Compos. Anal., 2000. **13**: p. 441-454.

65.   Beecher, G.R., *Overview of dietary flavonoids: nomenclature, occurrence and intake.* J Nutr, 2003. **133**(10): p. 3248S-3254S.

66.   Cao, J., et al., *Content of Selected Flavonoids in 100 Edible Vegetables and Fruits.* Food Science and Technology Research, 2010. **16**(5): p. 395-402.

67.   Kantsadi, A.L., 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, 2014. **67**: p. 35-43.

68.   Kamiyama, O., et al., *In vitro inhibition of α-glucosidases and glycogen phosphorylase by catechin gallates in green tea.* Food Chemistry, 2010. **122**(4): p. 1061-1066.

69.   Kato, A., et al., *Protective effects of dietary chamomile tea on diabetic complications.* Journal of Agricultural and Food Chemistry, 2008. **56**(17): p. 8206-8211.

70. Irwin, J.J., et al., *ZINC: A Free Tool to Discover Chemistry for Biology.* J Chem Inf Model, 2012.

71. Jakobs, S., et al., *Natural flavonoids are potent inhibitors of glycogen phosphorylase.* Mol Nutr Food Res, 2006. **50**(1): p. 52-7.

72. Kato, A., et al., *In vitro inhibition of alpha-glucosidases and glycogen phosphorylase by catechin gallates in green tea.* Food Chemistry, 2010. **122**(4): p. 1061-1066.

73. Kato, A., et al., *Structure-activity relationships of flavonoids as potential inhibitors of glycogen phosphorylase.* Journal of Agricultural and Food Chemistry, 2008. **56**(12): p. 4469-4473.

74. Oikonomakos, N.G., et al., *Flavopiridol inhibits glycogen phosphorylase by binding at the inhibitor site.* J Biol Chem, 2000. **275**(44): p. 34566-73.

75. Tsitsanou, K.E., et al., *Sourcing the affinity of flavonoids for the glycogen phosphorylase inhibitor site via crystallography, kinetics and QM/MM-PBSA binding studies: Comparison of chrysin and flavopiridol.* Food Chem Toxicol, 2013. **61**: p. 14-27.

76. Kyriakis, E., et al., *Natural flavonoids as antidiabetic agents. The binding of gallic and ellagic acids to glycogen phosphorylase b.* FEBS Lett, 2015. **589**(15): p. 1787-94.

77. Papageorgiou, A.C., N.G. Oikonomakos, and D.D. Leonidas, *Inhibition of rabbit muscle glycogen phosphorylase by D- gluconohydroximo-1,5-lactone-N-phenylurethane.* Archives of Biochemistry and Biophysics, 1989. **272**(2): p. 376-385.

78. Kasvinsky, P.J., et al., *Regulation of Glycogen Phosphorylase-a by Nucleotide Derivatives - Kinetic and X-Ray Crystallographic Studies.* Journal of Biological Chemistry, 1978. **253**(9): p. 3343-3351.

79. Madsen, N.B., S. Shechosky, and R.J. Fletterick, *Site-site interactions in glycogen phosphorylase b probed by ligands specific for each site.* Biochemistry, 1983. **22**(19): p. 4460-5.

80. Kosmopoulou, M.N., et al., *Binding of the potential antitumour agent indirubin-5-sulphonate at the inhibitor site of rabbit muscle glycogen phosphorylase b - Comparison with ligand binding to pCDK2-cyclin A complex.* European Journal of Biochemistry, 2004. **271**(11): p. 2280-2290.

81. Kosmopoulou, M.N., et al., *Indirubin-3 '-aminooxy-acetate inhibits glycogen phosphorylase by binding at the inhibitor and the allosteric site. Broad specificities of the two sites.* Letters in Drug Design & Discovery, 2005. **2**(5): p. 377-390.

82. Oikonomakos, N.G., et al., *Indirubin and indigo analogues as potential inhibitors of glycogenolysis: structural basis of glycogen phosphorylase inhibition*, in *Indirubin, the Red Shade of Indigo*, L. Meijer, et al., Editors. 2006, Life in Progress Editions: Roscoff, France. p. 177-189.

83. Tsitsanou, K.E., et al., *Sourcing the affinity of flavonoids for the glycogen phosphorylase inhibitor site via crystallography, kinetics and QM/MM-PBSA binding studies: Comparison of chrysin and flavopiridol.* Food and Chemical Toxicology, 2013. **61**: p. 14-27.

84. Hampson, L.J., et al., *Bioactivity of glycogen phosphorylase inhibitors that bind to the purine nucleoside site.* Bioorg Med Chem, 2006. **14**(23): p. 7835-45.

85. Wilcken, R., et al., *Principles and applications of halogen bonding in medicinal chemistry and chemical biology.* J Med Chem, 2013. **56**(4): p. 1363-88.

86. Valko, M., et al., *Role of oxygen radicals in DNA damage and cancer incidence.* Mol Cell Biochem, 2004. **266**(1-2): p. 37-56.

87. Schoonover, L.L., *Oxidative stress and the role of antioxidants in cardiovascular risk reduction.* Prog Cardiovasc Nurs, 2001. **16**(1): p. 30-2.

88. Halliwell, B., *Free radicals, antioxidants, and human disease: curiosity, cause, or consequence?* Lancet, 1994. **344**(8924): p. 721-4.

89. Baynes, J.W., *Role of oxidative stress in development of complications in diabetes.* Diabetes, 1991. **40**(4): p. 405-12.

90. Wolff, S.P., *Diabetes mellitus and free radicals. Free radicals, transition metals and oxidative stress in the aetiology of diabetes mellitus and complications.* Br Med Bull, 1993. **49**(3): p. 642-52.

91. Giugliano, D., A. Ceriello, and G. Paolisso, *Oxidative stress and diabetic vascular complications.* Diabetes Care, 1996. **19**(3): p. 257-267.

92. Nishikawa, T., et al., *Normalizing mitochondrial superoxide production blocks three pathways of hyperglycaemic damage.* Nature, 2000. **404**(6779): p. 787-790.

93. Singh, D.K., P. Winocour, and K. Farrington, *Oxidative stress in early diabetic nephropathy: fueling the fire.* Nat Rev Endocrinol, 2011. **7**(3): p. 176-84.

94. Takac, I., et al., *The E-loop is involved in hydrogen peroxide formation by the NADPH oxidase Nox4.* J Biol Chem, 2011. **286**(15): p. 13304-13.

95. Wolff, S.P., M.J.C. Crabbe, and P.J. Thornalley, *The autooxidation of simple monosaccharides.* Experientia, 1984. **40**(244-246).

96. Wolff, S.P., Z.Y. Jiang, and J.V. Hunt, *Protein glycation and oxidative stress in diabetes mellitus and ageing.* Free Radic Biol Med, 1991. **10**(5): p. 339-52.

97. Yan, S.D., et al., *Enhanced cellular oxidant stress by the interaction of advanced glycation end products with their receptors/binding proteins.* J Biol Chem, 1994. **269**(13): p. 9889-97.

98. Maziere, C., et al., *Glucose-enriched medium enhances cell-mediated low density lipoprotein peroxidation.* FEBS Lett, 1995. **363**(3): p. 277-9.

99. Burda, S. and W. Oleszek, *Antioxidant and antiradical activities of flavonoids.* J Agric Food Chem, 2001. **49**(6): p. 2774-9.

100. Blum, S., et al., *Vitamin E reduces cardiovascular disease in individuals with diabetes mellitus and the haptoglobin 2-2 genotype.* Pharmacogenomics, 2010. **11**(5): p. 675-84.

101. Cameron, N.E., M.A. Cotter, and E.K. Maxfield, *Anti-oxidant treatment prevents the development of peripheral nerve dysfunction in streptozotocin-diabetic rats.* Diabetologia, 1993. **36**(4): p. 299-304.

102. Kowluru, R.A., et al., *Beneficial effects of the nutritional supplements on the development of diabetic retinopathy.* Nutr Metab (Lond), 2014. **11**(1): p. 8.

103. Siman, C.M. and U.J. Eriksson, *Vitamin E decreases the occurrence of malformations in the offspring of diabetic rats.* Diabetes, 1997. **46**(6): p. 1054-61.

104. Ceriello, A., *New insights on oxidative stress and diabetic complications may lead to a "causal" antioxidant therapy.* Diabetes Care, 2003. **26**(5): p. 1589-96.

105. Jovanovic, S.V., et al., *Flavonoids as Antioxidants.* Journal of the American Chemical Society, 1994. **116**(11): p. 4846-4851.

106. Rice-Evans, C., *Screening of phenolics and flavonoids for antioxidant activity*, in *Antioxidant Food Supplements in Human Health*, L. Packer, M. Hiramatsu, and T. Yoshikawa, Editors. 1999, Academic Press: San Diego, CA. p. 239-253.

107. Bors, W., et al., *Flavonoids as antioxidants: determination of radical-scavenging efficiencies.* Methods Enzymol, 1990. **186**: p. 343-55.

108. Bors, W., C. Michel, and K. Stettmaier, *Antioxidant effects of flavonoids.* Biofactors, 1997. **6**(4): p. 399-402.

109. Cao, G., E. Sofic, and R.L. Prior, *Antioxidant and prooxidant behavior of flavonoids: structure-activity relationships.* Free Radic Biol Med, 1997. **22**(5): p. 749-60.

110. Chung, H.S., et al., *Flavonoid constituents of Chorizanthe diffusa with potential cancer chemopreventive activity.* J Agric Food Chem, 1999. **47**(1): p. 36-41.

111. Donnier-Marechal, M. and S. Vidal, *Glycogen phosphorylase inhibitors: a patent review (2013 - 2015).* Expert Opin Ther Pat, 2016. **26**(2): p. 199-212.

112. *Zinc Docking Database*. 2016 [cited 2015; Available from: http://zinc.docking.org/.

113. Begum, J., et al., *Computationally motivated synthesis and enzyme kinetic evaluation of N-(beta-D-glucopyranosyl)-1,2,4-triazolecarboxamides as glycogen phosphorylase inhibitors.* Medchemcomm, 2015. **6**(1): p. 80-89.

114. Lorber, D.M. and B.K. Shoichet, *Flexible ligand docking using conformational ensembles.* Protein Science, 1998. **7**(4): p. 938-950.

115. Leach, A.R., *Molecular Modelling Principles and Applications.* 2 ed. Vol. 2. 2001: Pearson Education.

116. Korb, O., T. Stutzle, and T.E. Exner, *Empirical Scoring Functions for Advanced Protein-Ligand Docking with PLANTS.* Journal of Chemical Information and Modeling, 2009. **49**(1): p. 84-96.

117. Huang, S.Y., S.Z. Grinter, and X.Q. Zou, *Scoring functions and their evaluation methods for protein-ligand docking: recent advances and future directions.* Physical Chemistry Chemical Physics, 2010. **12**(40): p. 12899-12908.

118. Peter Atkins, J.D.P., *Atkins' Physical Chemistry.* 9 ed. 2010, Oxford: Oxford University Press.

119. Peter Atkins, R.F., *Molecular Quantum Mechanics.* 2005, Oxford: Oxford University Press.

120. Peter Atkins, R.F., *Molecular Quantum Mechanics.* 5 ed. 2010, Oxford: Oxford University Press.

121. Parmenopoulou, V., et al., *Structure based inhibitor design targeting glycogen phosphorylase b. Virtual screening, synthesis, biochemical and biological assessment of novel N-acyl-beta-D-glucopyranosylamines.* Bioorganic & Medicinal Chemistry, 2014. **22**(17): p. 4810-4825.

122. Lorenzo, C., et al., *Insulin Resistance, beta-Cell Dysfunction, and Conversion to Type 2 Diabetes in a Multiethnic Population The Insulin Resistance Atherosclerosis Study.* Diabetes Care, 2010. **33**(1): p. 67-72.

123. B., K., *Type 2 Diabetes. When Insulin Secretion Fails to Compensate for Insulin Resistance.* Cell, 1998. **92**(5): p. 593-596.

124. Middleton, E., C. Kandaswami, and T.C. Theoharides, *The effects of plant flavonoids on mammalian cells: Implications for inflammation, heart disease, and cancer.* Pharmacological Reviews, 2000. **52**(4): p. 673-751.

125. Chen, Z.W., et al., *Synthesis and biological evaluation of flavonoids as vasorelaxant agents.* Bioorganic & Medicinal Chemistry Letters, 2004. **14**(15): p. 3949-3952.

126. Vyas, B., et al., *Aldose reductase inhibitors for diabetic complications: Receptor induced atom-based 3D-QSAR analysis, synthesis and biological evaluation.* Journal of Molecular Graphics & Modelling, 2015. **59**: p. 59-71.

127. Zheng, X., et al., *Synthesis and Preliminary Biological Evaluation of Chrysin Derivatives as Potential Anticancer Drugs.* Medicinal Chemistry, 2010. **6**(1): p. 6-8.

128. Kabalka, G.W. and A.R. Mereddy, *Microwave-assisted synthesis of functionalized flavones and chromones.* Tetrahedron Letters, 2005. **46**(37): p. 6315-6317.

129. Seijas, J.A., M.P. Vazquez-Tato, and R. Carballido-Reboredo, *Solvent-free synthesis of functionalized flavones under microwave irradiation.* Journal of Organic Chemistry, 2005. **70**(7): p. 2855-2858.

130. Zambare, A.S., et al., *Development of mild and efficient method for synthesis of substituted flavones using oxalic acid catalyst.* Chinese Chemical Letters, 2009. **20**(2): p. 171-174.

131. Sashidhara, K.V., M. Kumar, and A. Kumar, *A novel route to synthesis of flavones from salicylaldehyde and acetophenone derivatives.* Tetrahedron Letters, 2012. **53**(18): p. 2355-2359.

132. Kumar, S., Sharma, D., *Oxidative Cyclisation of 2'-Hydroxychalcones using Sodium Tellurite: Synthesis of Flavones*, in *Oriental Journal of Chemistry*. 2011. p. 761-763.

133. Kumar, K.H. and P.T. Perumal, *A novel one-pot oxidative cyclization of 2'-amino and 2'-hydroxychalcones employing FeCl3 center dot 6H(2)O-methanol. Synthesis of 4-alkoxy-2-aryl-quinolines and flavones.* Tetrahedron, 2007. **63**(38): p. 9531-9535.

134. Khan, A.T. and P. Goswami, *A highly efficient and environmentally benign synthesis of 6,8-dibromoflavones, 8-bromoflavones, 5,7-dibromoaurones and 7-bromoaurones.* Tetrahedron Letters, 2005. **46**(30): p. 4937-4940.

135. Lee, A.F., et al., *Hydrodebromination of bromobenzene over pt(111).* Journal of Physical Chemistry C, 2007. **111**(28): p. 10455-10460.

136. Yenjai, C., et al., *Structural Modification of 5,7-Dimethoxyflavone from Kaempferia parviflora and Biological Activities.* Archives of Pharmacal Research, 2009. **32**(9): p. 1179-1184.

137. Miao, H. and Z. Yang, *Regiospecific carbonylative annulation of iodophenol acetates and acetylenes to construct the flavones by a new catalyst of palladium-thiourea-dppp complex.* Organic Letters, 2000. **2**(12): p. 1765-1768.

138. Defronzo, R.A., R.C. Bonadonna, and E. Ferrannini, *Pathogenesis of Niddm - a Balanced Overview.* Diabetes Care, 1992. **15**(3): p. 318-368.

139. Oikonomakos, N.G. and L. Somsak, *Advances in glycogen phosphorylase inhibitor design.* Current Opinion in Investigational Drugs, 2008. **9**(4): p. 379-395.

140. Schrodinger, *L.L.C: New York.* 2016.

141. Lipinski, C.A., et al., *Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings.* Advanced Drug Delivery Reviews, 2001. **46**(1-3): p. 3-26.

142. Friesner, R.A., et al., *Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy.* Journal of Medicinal Chemistry, 2004. **47**(7): p. 1739-1749.

143. Sherman, W., H.S. Beard, and R. Farid, *Use of an induced fit receptor structure in virtual screening.* Chemical Biology & Drug Design, 2006. **67**(1): p. 83-84.

144. Feher, M., *Consensus scoring for protein-ligand interactions.* Drug Discovery Today, 2006. **11**(9-10): p. 421-428.

145. Houston, D.R. and M.D. Walkinshaw, *Consensus Docking: Improving the Reliability of Docking in a Virtual Screening Context.* Journal of Chemical Information and Modeling, 2013. **53**(2): p. 384-390.

146. Zhao, Y. and D.G. Truhlar, *The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals.* Theoretical Chemistry Accounts, 2008. **120**(1-3): p. 215-241.

147. Kaminski, G.A., et al., *Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides.* Journal of Physical Chemistry B, 2001. **105**(28): p. 6474-6487.

148. Harder, E., et al., *OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins.* Journal of Chemical Theory and Computation, 2016. **12**(1): p. 281-296.

149. Li, J.N., et al., *The VSGB 2.0 model: A next generation energy model for high resolution protein structure modeling.* Proteins-Structure Function and Bioinformatics, 2011. **79**(10): p. 2794-2812.

150. Wilcken, R., et al., *Principles and Applications of Halogen Bonding in Medicinal Chemistry and Chemical Biology.* Journal of Medicinal Chemistry, 2013. **56**(4): p. 1363-1388.

151. Ong, K.C. and H.E. Khoo, *Effects of myricetin on glycemia and glycogen metabolism in diabetic rats.* Life Sciences, 2000. **67**(14): p. 1695-1705.

152. Ahmad, M., et al., *Hypoglycaemic action of the flavonoid fraction of Cuminum nigrum seeds.* Phytotherapy Research, 2000. **14**(2): p. 103-106.

153. Anila, L. and N.R. Vijayalakshmi, *Beneficial effects of flavonoids from Sesamum indicum, Emblica officinalis and Momordica charantia.* Phytotherapy Research, 2000. **14**(8): p. 592-595.

154.   Saxena, S., J.K. Makrandi, and S.K. Grover, *Synthesis of 5-Hydroxyflavones and or 7-Hydroxyflavones Using a Modified Phase Transfer-Catalyzed Baker-Venkataraman Transformation.* Synthesis-Stuttgart, 1985(6-7): p. 696-696.

155.   Khanapur, M.P., Nishal K.; Badiger, Jaishree *Synthesis and anti-inflammatory in vitro, in silico, and in vivo studies of flavone analogues.* Medicinal Chemistry Research, 2015. **24**(6): p. 2656 - 2669.

156.   Ravishankar, D., et al., *Novel synthesised flavone derivatives provide significant insight into the structural features required for enhanced anti-proliferative activity.* Rsc Advances, 2016. **6**(69): p. 64544-64556.

157.   Ranu, B.C. and S. Bhar, *Dealkylation of ethers. A review.* Organic Preparations and Procedures International, 1996. **28**(4): p. 371-409.

158.   Federation, I.D., *IDF Diabetes Atlas.* International Diabetes Federation, 2015. **7**.