IN SILICO DESIGN OF COMPOUNDS AS SGLT2 INHIBITORS

Research Project Submitted to the Central University of Punjab

For the Award of

Master of Science

In

Life Sciences with Specialization in Human Genetics

By

Kamaljyoti Chakravorty

Supervisor

Prof. (Dr.) Anjana Munshi

Department of Human Genetics and Molecular Medicine

School of Health Sciences

Central University of Punjab, Bathinda

May, 2018

CENTRAL UNIVERSITY OF PUNJAB, BATHINDA

DECLARATION

I declare that all the changes suggested by examiners in the research project entitled **"*In silico* design of compounds as SGLT2 inhibitors"** have been incorporated by me for the award of the degree of **M.Sc. in Life Sciences with specialization in Human Genetics** in the **Department of Human Genetics and Molecular Medicine** in this project work.

Kamaljyoti Chakravorty
Department of Human Genetics and Molecular Medicine
School of Health Sciences
Central University of Punjab, Bathinda- 151001.

DECLARATION

I declare that the project entitled "*IN SILICO* DESIGN OF COMPOUNDS AS SGLT2 INHIBITORS" has been prepared by me under the guidance of Prof. (Dr.) Anjana Munshi, H.O.D Department of Human genetics and Molecular medicine, Dean, School of Health Sciences, Central University of Punjab, Bathinda. No part of this dissertation has formed the basis for the award of any degree or fellowship previously.

[Kamaljyoti Chakravorty]

[16mslshg04]

[Department of Human genetics and Molecular medicine]

School of Health Sciences,

Central University of Punjab,

Bathinda – 151001

Date:

CERTIFICATE

I certify that Kamaljyoti Chakravorty has prepared his project entitled "*IN SILICO* DESIGN OF COMPOUNDS AS SGLT2 INHIBITORS", for the award of M.Sc. degree from Central University of Punjab, Bathinda under my guidance. He has carried out this work at the Department of Human Genetics and Molecular Medicine, School of Health Sciences, Central University of Punjab, Bathinda.

[Dr. Anjana Munshi]

[Department of Human Genetics and Molecular Medicine],

School of Health Sciences,

Central University of Punjab,

Bathinda – 151001

Date:

ABSTRACT

In silico design of compounds as SGLT2 inhibitors

Name of the student: Kamaljyoti Chakravorty

Registration number: 16mslshg04

Degree for which submitted: M.Sc Life Sciences (specialization in Human genetics)

Name of the supervisor: Prof. Anjana Munshi

Name of the department: Human genetics and Molecular medicine

Name of the School: Health Sciences

Key words: SGLT2 inhibitors, type 2 diabetes mellitus, gliflozin, structural insights, aglyconic, binding site.

SGLT2 inhibitors (SGLT2i) are the current novel therapeutic approach expected to control the growing threat of type 2 diabetes mellitus (T2DM). With comparatively least reported side effects, SGLT2i (gliflozins) have been identified as promising tools to tackle the threat of T2DM. However, studies have also reported these drugs to cause renal impairments and urinogenital infections among certain T2DM patients. The efficacy of an inhibitor is fundamentally determined by the stability of protein-inhibitor complex. Therefore, it is essential to study the binding site residues of SGLT2 in the light of inhibitors' interactions. Structural insights of SGLT2 suggested a competitive inhibition of the ligand glucose (agonist) by the gliflozins. The inhibitory effect of SGLT2i reduces the renal reabsorption of glucose and promotes glycosuria. Consequently, a reduced plasma glucose level prevents the risk of hyperglycemia and further T2DM. In order to design potent inhibitors the structures of the available gliflozins (empagliflozin, canagliflozin, dapagliflozin and ertugliflozin) were analysed. Accordingly, a library of 44 fragment molecules was generated for interactive enumeration. A core ligand structure was designed based on the structural analysis of the gliflozin. A total of 3250 ligand molecules were obtained using schrodinger maestro v11.3. Docking results have shown ligand molecules exhibiting two different modes of inhibition. Set A molecules exhibited glyconic mode of interaction while set B molecules displayed aglyconic mode of interaction at the glucose binding site. The interactions of set B molecules are similar to that of the standard gliflozins and are expected to be less stable. Lesser stability of the protein-

ligand complex perhaps lead to the reported side effects. On the other hand, set A molecules (with lower docking score) are expected to be much stable in terms of their interactions which differ significantly from that of the standard gliflozins. Therefore, set A molecules are the anticipated potent SGLT2 inhibitors expected to show reduced side effects.

(Kamaljyoti Chakravorty) Prof. Anjana Munshi

ACKNOWLEDGEMENT

I would like to express my gratitude to Prof. (Dr.) Anjana Munshi (supervisor), H.O.D Department of Human genetics and Molecular medicine, Dean, School of Health Sciences, Central University of Punjab, Bathinda for her advice, guidance and supervision.

I am grateful to Dr. Raj Kumar, H.O.D Department of Pharmaceutical Sciences and Natural products for granting me permission to use their computational tools.

I am deeply thankful to Saurav Kalra, Ph.D for his constant guidance and advices regarding the use of computational tools and the relevant work.

I am grateful to Prof. (Dr.) R.K. Kohli (Vice Chancellor) and Prof. (Dr.) P. Ramarao (Dean, School of Basic and Applied Sciences), Central University of Punjab, Bathinda.

I am thankful to all the faculty members of Department of Human genetics and Molecular medicine for their indispensable support and encouragement.

I am also thankful to Mr. Roshan Kumar and my seniors for their constant guidance in the laboratory.

My special thanks to my parents along with all my friends, classmates and lab mates for their precious support and encouragement required for the completion of this project work.

Table of Contents

List of tables

List of figures

Abbreviation

S. no.	Abbreviations	Full form
1.	WHO	World Health Organisation
2.	PCOS	Plycystic Ovarian Syndrome
3.	ER	Endoplasmic Reticulum
4.	T1DM	Type 1 diabetes mellitus
5.	T2DM	Type 2 diabetes mellitus
6.	GDM	Gestational diabetes mellitus
7.	IR	Insulin Receptor
8.	GLUT	Glucose Transporter
9.	NF-kB	nuclear factor kappa-light-chain-enhancer of activated B cells
10.	MODY	Maturity onset diabetes in youngs
11.	GCK	Glucokinase
12.	ROS	Reactive oxygen species
13.	AGE	Advanced glycation end product
14.	PKC	Protein kinase C
15.	CAD	Coronary artery disease
16.	PAD	Peripheral artery disease
17.	LDL	Low density lipo-protein
18.	MI	Myocardial Infarction
19.	NO	Nitric oxide
20.	CADD	Computer-aided drug design
21.	SBDD	Structure-based drug design
22.	LBDD	Ligand based drug design
23.	QSAR	Quantitative structure-activity relationship

24.	DPP-4	Dipeptidyl peptidase 4
25.	TZD	Thiazolidinedione
26.	SGLT	Sodium glucose co-transporters
27.	SLC	Solute carrier family
28.	PCT	Proximal convoluted tubule
29.	FRG	Familial renal glycosuria
30.	FDA	Food and Drug Administration
31.	GFR	Glomerular filtration rate
32.	CYP	Cytochrome
33.	UGT	UDP-glucuronosyltransferase
34.	SGLT2i	SGLT2 inhibitor
35.	PDB	Protein databank
36.	OPLS3	Optimised potential for Liquid Simulation version 3
37.	HTVS	High Throughput Virtual Screening
38.	SP	Standard Precision
39.	XP	Extra Precision

Chapter 01

INTRODUCTION

Diabetes mellitus is a metabolic disorder characterized by hyperglycemia and glucose intolerance as their hallmark (Association, 2014). As far as the prevalence is concern, Diabetes mellitus is expected to become the seventh leading cause of death by 2030 (Mathers and Loncar, 2006). According to WHO, the prevalence of diabetes has increased by four times since 1980 to 2014 and has reached its epidemic proportion by now (Organization, 2016). India is among the top ten countries with high diabetic prevalence and currently faces a great threat with 69.2 million diabetic individuals approximately (Collaboration, 2016; Prabhakaran *et al.*, 2016; Sandu *et al.*, 2016). In fact, it is expected to reach beyond 123.5 million by 2040 (Patnaik *et al.*, 2016). Therefore, it is essential to either improve the available therapeutics or develop new therapeutics with utmost treatment efficacy.

Diabetes mellitus is a multifactorial disorder caused by genetic factors, environmental factors as well as gene-environment interactions (Kaveeshwar and Cornwall, 2014). Obesity in association with lifestyle changes is one of the major risk factors for diabetes. Modern lifestyle along with overconsumption of energy-rich food leads to excess release of fat molecules in the body. The long saturated free fatty acids are believed to induce ER stress which leads to β cell dysfunction and apoptosis eventually (Back and Kaufman, 2012). In addition to this, age, ethnicity and family history are also among the essential risk factors (Gunasekaran and Gannon, 2011; Spanakis and Golden, 2013; Zhang *et al.*, 2006). Moreover, PCOS patients often demonstrate high insulin level conferring insulin resistance and further increasing the risk for type 2 diabetes mellitus (Gambineri *et al.*, 2012).

Diabetes mellitus is broadly classified into three types; Type 1 diabetes mellitus (formerly named Insulin Dependent Diabetes Mellitus), Type 2 diabetes mellitus (formerly named Non-Insulin Dependent Diabetes Mellitus) and Gestational diabetes (GDM). In type 1 diabetes mellitus (T1DM) the cellular uptake of plasma glucose decreases due to reduction in insulin secretion. Auto - antibodies formed against the pancreatic β cells cause β cell apoptosis and, thus, reduce the secretion of insulin (Deutschländer *et al.*, 2010). Type 2 diabetes mellitus (T2DM) is, however, mostly antibody independent and occurs due to insulin insensitivity. It alone accounts for

about 90% of the diabetes cases while T1DM accounts only for about 10% (Chao and Henry, 2010). Gestational diabetes mellitus is uniquely observed in pregnant women during their third trimester. The gravid females show varying degrees of hyperglycemia during pregnancy. Although the condition usually gets resolved within six weeks of delivery, yet it increases the risk of developing T2DM in the later part of their life (Laboratories, 2016).

The current focus is on T2DM along addressing the current pharmaco-therapeutic approaches available in the market.

1.1. **Pathophysiology of T2DM**

The human body has certain level of glucose tolerance due to the antagonistic balanced metabolism of plasma glucose by insulin and glucagon in their optimum level (normoglycemic condition). However, hyperglycemia is one of the outcomes of a change in this equilibrium. In T2DM, hyperglycemia is known to be the first manifestation due to insulin resistance. Insulin receptors (IR) expressed by the cells (especially muscle cells and fat cells; also known as the cells of periphery) do not respond to the domain-bound insulin molecules. Consequently, the activation of glucose transporters (GLUT) fails to occur and halts the uptake of glucose increasing the plasma glucose concentration. Such resistance to insulin causes a transformation of the β cell mass and further increasing the plasma insulin level for compensating the demand. The increased level of insulin (hyperinsulinemia) along with insulin resistance leads to impaired glucose tolerance eventually (Cernea and Dobreanu, 2013; DeFronzo *et al.*, 2015). Insulin resistance also impairs the suppression of hepatic glycolysis and gluconeogenesis and also weakens the uptake of free fatty acids by the adipocytes (causing hyperlipidemia) (Otero *et al.*, 2014). Chronic hyperglycemia (glucotoxicity) stimulates multiple stress signals (e.g. ER stress) and pathways (e.g. NF-κB pathway) which eventually leads to β cell apoptosis and, hence, reduces insulin secretion (Back and Kaufman, 2012; Cernea and Dobreanu, 2013; Fiorentino *et al.*, 2013).

The mode of Inheritance of T2DM is yet unclear except for MODY. Maturity onset diabetes in young (MODY) is an autosomal dominant heterogenous disorder reported to occur from the mutation of glucokinase (GCK) gene (7p13).

1.2. **Complications of T2DM**

Chronic hyperglycemia has been reported to cause long-term damage and failure of various organ systems. Such adverse complications of T2DM are broadly divided into two types; Microvascular complications and Macrovascular complications. Microvascular complications arise due to damage of microvasculatures (small blood vessels for intratissue blood circulation) while macrovascular complications are caused due to damage of macrovasculatures (large blood vessels for intertissue blood circulation) (Organisation, 2018). As shown in table 1, microvascular complications include diabetic retinopathy, nephropathy and neuropathy while macrovascular complications include coronary artery disease, peripheral arterial disease and stroke (Fowler, 2011). Studies have reported that reactive oxygen species (ROS), advanced glycation end product (AGE) and intracellular signaling molecules such as protein kinase C (PKC) are certain key modulators responsible for the various complications of T2DM (Chawla *et al.*, 2016). Currently most of the studies are concerned with macrovascular complications. Drug therapeutics is, however, available to delay the onset and further halt the progression of these complications.

Table 1. Complications of type 2 diabetes mellitus (T2DM).

Microvascular complications	**Features**	**References**
Diabetic retinopathy	Hyperglycemia causes damage to the blood retinal barrier. Consequently leakage of fluid into the tissue, causing edema, leads to progressive loss of vision.	(Organisation, 2018)

Diabetic nephropathy	High blood glucose level damages the small blood vessels of kidney leading to microalbuminuria. Progressively proteinuria is observed which eventually overt diabetic kidney disease.	(Fowler, 2011; Organisation, 2018)
Diabetic neuropathy	Hyperglycemia mediated damage to microvasculatures reduces blood flow to the nerves. As a result, the diminished neural activity leads to sensory loss in extremities, limb damage (amputation in extreme cases) and even impotency in some individuals.	(Organisation, 2018)
Macrovascular complications	**Features**	**References**
Coronary artery disease (CAD)	Diabetic patients are diagnosed with a large concentration of small LDL molecules (due to hyperlipidemia). Due to high atherogenicity of small LDL, they deposit in the base of large bood vessels of the heart (coronary arteries). The deposition obstructs the blood flow (atherosclerosis) and increases the risk of myocardial infarction (MI).	(Vergès, 2015)
Peripheral artery disease (PAD)	Diabetic patients are reported with a decreased bioavailability of NO (a potent vasodilator) and increased secretion of endothelin-1 (vasoconstrictor) leading to endothelial–smooth muscle dysfunction.	(Sena *et al.*, 2013)

	This along with atherosclerosis and arterial stenosis increases the risk of PAD, especially claudications.	
Stroke	Anti–coagulant mechanisms are observed to be diminished in diabetic patients along with an increase level of blood clotting factors and platelets. Therefore, arteries ruptured by atherosclerotic plaque results in increased platelet aggregation (hypercoagulation). Progressive hypercoagulation and occlusion in blood vessels increases the risk of ischemic stroke.	(Jacobson *et al.*, 2014)

1.3. **Computer-aided drug design**

Computer-aided drug design (CADD) is a method to design, enhance and discover efficient drug molecules using computational chemistry and analyze their biological activity. Advanced computational tools are used to design library of ligand molecules using the knowledge from the target molecule. CADD is also referred to as rational drug design in some context. More fundamentally, CADD deals with estimating the affinity of a given molecule to bind with the target and henceforth estimating the strength of their bond. Although computational tools are essential for drug designing, not all designing methods necessarily depend on computer modeling techniques.

CADD, ideally, allows a researcher to identify the potent ligand molecule prior to its synthesis. As a result, the required molecule can be synthesized making judicious utilization of the available limited resources. Although CADD has accelerated drug discovery by reducing the number of required iterations, yet the estimates of binding affinity are still qualitative in practice. Lack of perfect computational methods perhaps

is the reason for the recurrent need of several iterations before the discovery of the drug molecule.

Drug design techniques are of two major types; Structure-based drug design and Ligand-based drug design.

1.3.1. **Structure-based drug design (SBDD)**

SBDD is concerned with using the 3D structure of a protein to create a potent drug molecule. Homology modeling is a well-known technique in SBDD. X-ray crystallography, Nuclear Magnetic Resonance, and Mass Spectrometry are the techniques used to obtain a 3D structure of a protein with optimum resolution. Resolution of a few angstroms can favor researchers to precisely examine the various atomic interactions within the target molecule as well as between the target and the drug molecule. Drug targets, especially protein, appear to be a key molecule for the occurrence of certain diseases. Therefore, the creation of a potent drug molecule for the target can help in counteracting and reducing the prevalence of the disease.

1.3.2. **Ligand-based drug design (LBDD)**

In case the structure of target molecule is not known but the interacting ligand molecules are known, molecules with structural similarities can be created. The resulting molecules have a structure-activity relationship (SAR) with the standard ligand molecule and expected to exhibit similar interactions. The widely used techniques such as QSAR and Pharmacophore modeling take into account the chemistry, shape, and interaction points to estimate the similarity. Pharmacophore modeling analyses the functional part of the standard ligand molecule and assesses in the creation of a ligand molecule database. However, QSAR requires a large extent of standard molecules and their activity data.

1.4. **Docking**

Docking is a computational method to predict the preferred orientation of a molecule when bound to another molecule. Specially designed computer software (such as

Maestro, AutoDock etc.) are employed for docking and analyzing the various interactions within a complex. Knowledge about their interaction could unveil the biological activity of one molecule with respect to the other. For an effective drug molecule, the interactions could reveal the strength of its association with the residues of the active site. The procedures of docking can be performed much before any (wet) experiments with a benefit of saving resources and time eventually. Therefore, Docking is a crucial method of drug design and computational biology.

1.4.1. **Theories behind molecular docking**

The process of molecular docking has concepts of two sets of theories; *Lock and key* and *Induced fit*. According to the *Lock and key* theory, both the target and the ligand molecule are assumed to be *rigid*. The ligand molecule can only show the *best fit* if it has the preferred conformation as needed by the active site. However, in practice, the phenomenon of docking resembles more like the *hand-in-a-glove*. According to this concept, both the target and the ligand molecules are assumed to be *flexible*. In fact, both the molecules undergo certain degrees of change in their conformation to achieve a preferred orientation for binding. This theory is often referred to as the theory of *induced fit*.

The software of molecular docking provide a molecular modeling environment to simulate the process of docking. They also identify the optimized conformations of each ligand molecules and the target molecule required during the complex formation. The aim of the whole process is to minimize the free energy of the system and obtain an optimized protein-ligand complex with utmost stability.

1.4.2. **Types of docking**

Molecular docking broadly has two main types of docking; *rigid ligand docking* and *flexible ligand docking*. In both the docking types the target molecule is usually kept rigid so that the docking of the ligand molecule can be assessed. In rigid docking, the ligand molecule does not undergo any significant change in its conformation during the process. However, flexible docking enables the ligand molecule to exhibit

changes in bond angles, bond length and torsion angles at certain stages of simulation. Changes in conformation prolong the simulation time in flexible docking.

1.4.3. **Mechanics of docking**

In docking, the flexibility of a ligand molecule is determined by the number of rotatable bonds available. The flexibility of molecule enables it to change conformation and attain the best fit. However, higher the flexibility longer will be the simulation time.

The conformational changes bring up optimized ligand molecule with expected potency of drug action. Docking involves sets of algorithm to search and find the optimized ligand conformations. Furthermore, the process involves mathematical methods which provide each conformation with a score with respect to their interactions. Therefore, a docking program includes four main components to ensure accuracy, reliability, and reproducibility of the result. They are Search algorithm, Ligand flexibility, Receptor flexibility and scoring function.

1.4.3.1. **Search algorithm**

The process of molecular docking produces optimized ligand molecules with conformation of best fit in 3D space. However, in practice, exploring the entire space for the molecules is impossible. Therefore, the search algorithm program helps in identifying the ligand molecules in the space. It enumerates almost every possible distortion undergone by the ligand molecule and further calculates the energy needed for the complex formation in terms of *Snapshots*. Each Snapshot represents a *pose* defining kind and distance of interaction between the ligand molecule and target residues.

1.4.3.2. **Ligand flexibility**

Flexibility of a ligand is determined by the number of rotatable bonds. Although ligand flexibility is essential for docking yet higher the flexibility of the ligand lower is the potency of the molecule as a drug. Therefore, a potent ligand molecule ought to have a minimum number of rotatable bonds (atmost 5 allowed rotatable bonds) within certain defined range favorable to docking.

1.4.3.3. **Receptor flexibility**

The flexibility of a receptor is determined by the number of rotamers. Rotamers are low energy isomers with conformations due to restricted rotation around the rotatable bonds (usually single bond). Higher the number of rotamers lower is the stability of the receptor molecule.

1.4.3.4. **Scoring function**

Scoring function uses mathematical methods to estimate the binding affinity between a ligand and the receptor molecule. Docking involves non covalent interactions between the receptor and ligand. However, some ligands bound to the receptor weakly in comparison to the other molecules. The purpose of scoring function is to evaluate the docked ligands and estimate their binding affinity. Theoretically, binding affinity is defined as the ratio of the rate at which the complex associates to that of the rate at which it dissociates. Based on the estimations the program ranks the molecules and subsequently defines the extent of their interaction. There are three groups of scoring function in case of the docking algorithms (table 2);

a) Force field
b) Empirical
c) Knowledge-based

Table 2. Classes of Scoring function along with their descriptive features and example.

Scoring function	Feature	Example
Force field	➢ Binding affinity is estimated by evaluating the Intramolecular interactions, that is, van der waals force and electrostatic force. ➢ Intramolecular energy and desolvation energy of the ligand	Autodock, G-score, D-score.

	are also evaluated.	
Empirical	➢ Binding affinity is estimated either by evaluating the number of different interactions or by assessing the *solvent accessible surface area* before and after the complex formation. ➢ The interactions and entities mainly assessed by this program are hydrophobic, hydrophilic and hydrogen bonds along with rotatable bonds.	ChemScore, F-Score, X-Score and Fresno.
Knowledge-based	➢ Certain atoms or entities occur more frequently compared to the others in random. Such atoms or entities have higher binding affinity in terms of energy. ➢ Intermolecular interactions are evaluated statistically and the binding poses are identified.	PMF, DrugScore, and SMoG.

1.5. **Drug-Receptor Interaction**

The Ligand or Drug molecules interact with the receptor making reversible or irreversible bonds. Covalent bonds are irreversible bonds due to their utmost stability and high bond enthalpy. On the other hand the non–covalent bonds are weak and reversible. Docking studies are mostly based on drugs with non–covalent interaction for a short term termination of receptor activity. However, covalent interactions are studied rarely in concerned with drug–receptor interactions (e.g. poison molecules.). Non–covalent interactions can further be classified as;

- Van der waals Interactions
- Electrostatic Interactions
- Dipole Interactions
- Hydrogen Bonds (hydrophilic interactions)
- Hydrophobic interactions
- π-π stacking

1.6. Limitations with current docking methodologies

- Higher the number of rotatable bonds, higher will be the flexibility of the ligand and lower will be the stability.
- Higher numbers of rotatable bonds increases entropic effect. Consequently, formation of Intramolecular bonds reduces the stability of the ligand.
- Higher the flexibility of receptor lower is the stability and hence the ability to make best fit with ligand molecule.
- Docking occurs in a space devoid of water molecules and ions while the human body has a fluid based internal environment.
- Larger the size of ligand lower is the specificity to bind to the active site of the receptor molecule.
- Pharmacokinetic effects, allosteric effects, biomolecule-biomolecule interactions cannot be understood using docking.

1.7. Applications of docking

- Identification of Hit
- Optimization of Lead
- Identification of the binding site
- Protein–Protein or Protein–nucleic acid interaction
- Structure–function relationship studies
- Bio–remediation

1.8. Anti-diabetic drugs in the market

Although plenty of drugs are available for diabetes yet very few have the potential to reduce the adverse effects of T2DM with fewer side effects. Antidiabetic drugs such as biguanides, insulin or insulin analogues, thiazolidinedione, sulfonylureas, incretin mimetics, α-glucosidase inhibitors, DPP-4 inhibitors and Meglitinides are available with different brand names in the market. Metformin (biguanide) is an orally administered antihyperglycaemic agent which reduces the fasting blood glucose level by lowering gluconeogenesis and enhancing cellular insulin sensitivity. Phenformin and Buformin were other biguanides which were withdrawn from the market as they increased the risk of lactic acidosis and cardiac mortality (Foretz *et al.*, 2014). However, Metformin with very few cases of adverse side effects is currently the first line of treatment for T2DM (Foretz *et al.*, 2014). Thiazolidinedione (TZD) is a class of drugs that binds to certain receptors of adipocytes and promotes adipogenesis. Consequently, the circulating fatty acids concentration and availability of lipids decreases (prevents hyperlipidemia) and insulin sensitivity improves with advent of time. Rosiglitazone and Pioglitazone are the TZDs available in the market (Soccio *et al.*, 2014). Troglitazone was another TZD removed from the market in 2000 for increasing the risk of liver failure among the patients (Kohlroser *et al.*, 2000; Soccio *et al.*, 2014). Sulfonylureas drugs are insulin secretagogues which follow a mechanism to increase plasma insulin level by blocking the inflow of K^+ into the β cells. As a result, insulin secretory cells get depolarized promoting the influx of Ca^{2+} into the cytosol. The rise in the cytosolic Ca^{2+} concentration causes actomyosin contraction and hence exocytosis of insulin (Sola *et al.*, 2015). Carbutamide was the first sulfonylurea to be introduced into the market, which was later withdrawn due to its bone marrow related adverse effects (Sola *et al.*, 2015). The first generation drugs (tolbutamide, chlorpropamide, acetohexamide and tolazamide) are recommended rarely at present while the second generation drugs (glyburide, glipizide, gliclazide) are often prescribed in combination with Metformin and Insulin (Basit *et al.*, 2012; Thulé and Umpierrez, 2014). Glimepiride is the most recently developed second generation sulfonylurea (Basit *et al.*, 2012; Thulé and Umpierrez, 2014). Meglitinides are the other insulin secretagogues available in the market (Repaglinide and Nateglinide). Their mechanism of action resembles that of the sulfonylureas although

their binding site is different on the “sulfonylurea receptor” of the pancreatic β cell (Luna and Feinglos, 2001). Incretins are peptide gut hormones which increases the secretion of Insulin in response to ingestion of glucose containing meal (Incretin effect) (Hansen *et al.*, 2010; Holst, 2007). However, T2DM patients show a reduction in the level of Incretin secretion (due to DPP-4 mediated enzymatic degradation of Incretins), contributing towards hyperglycaemia (Deacon *et al.*, 1995; Hansen *et al.*, 2010). Therefore, Incretin mimetics are certain analogues which tend to re-establish the incretin effect and subsequently recovering normoglycaemic state (Hansen *et al.*, 2010). Exenatide, Liraglutide, Albiglutide, Dulaglutide and Lixisenatide are the well known GLP-1 agonists (or incretin mimetics) available at present (Hansen *et al.*, 2010; Trujillo *et al.*, 2015). DPP-4 inhibitors, such as sitagliptin, prevent the enzymatic degradation of the incretin hormones and prevent hyperglycaemia (Pathak and Bridgeman, 2010). These drugs, however, often come up with side effects such as diarrhea, hypoglycemia, weight gain, edema, nausea, vomiting etc. Although combinations of multiple agents have shown positive outcomes yet their mechanisms involve lots of side effects. Therefore, it is essential to develop therapeutic agents with unique mechanisms and characteristics showing reduced side effects.

The above mentioned drugs are being summarized and described in table 3 along with their respective side effects.

Table 3. Anti-diabetic drugs and their side effects

Drug types	Available drugs	Side effects	References
Biguanides	Metformin	Long term use can cause Lactic acidosis, stomach pain, nausea, weight loss etc.	(Foretz *et al.*, 2014)

Insulin and its analogues	Peptides with different brand names, such as Humalog, NovoRapid, Lantus etc.	Weight gain is usually common. However, Skin reactions are also observed in some. Improper dose might cause low or high blood sugar.	(Drugs.com, 2018)
Thiazolidinedione (Glitazones)	Rosiglitazone, Pioglitazone	Increased risk of bone fracture and myocardial infarction, weight gain, dark – colored urine, upset stomach etc.	(Rizos *et al.*, 2009)
Sulfonylureas	1st Generation: Carbutamide, Acetohexamide, Tolbutamide, Chlorpropamide, Tolazamide. 2nd Generation: Glipizide, Glibenclamide, Gliclazide, Glyclopyramide, Glibornuride, Glyburide 3rd Generation: Glimepiride	Weight gain and allergic reactions are usually observed. Excessive dosage might lead to hypoglycemia.	(Sola *et al.*, 2015)

Incretin mimetics (GLP-1)	Exenatide, Liraglutide, Albiglutide, Dulaglutide, Lixisenatide	Diarrhea, Constipation, Nausea, Vomiting, Indigestion etc.	(Hansen *et al.*, 2010)
α-glucosidase inhibitors	Acarbose, Miglitol, Voglibose	Abdominal bloating, flatulence, diarrhea and other gastrointestinal adverse effects.	(Ghosh and Collier, 2012)
DPP-4 inhibitors	sitagliptin, vildagliptin, saxagliptin, linagliptin, alogliptin	respiratory tract infection, nasopharyngitis, urinary tract infection, and headache.	(Pathak and Bridgeman, 2010)
Meglitinides	repaglinide, nateglinide	In some patients weight gain is reported. Back pain, GI side effects and upper respiratory infections are also observed in other patients.	(Ryder *et al.*, 2010)
SGLT2 inhibitors	dapagliflozin, canagliflozin, empagliflozin, ertugliflozin	Urinary tract infection in some females and weight loss in most of the individuals.	(Madaan *et al.*, 2016)

The anti-diabetic drugs listed above have proven their immense efficacy in lowering the incidence of T2DM to certain extent. However, the reported rapid growth in T2DM

prevalence clearly indicates the necessity of more efficient drugs. So far SGLT2 inhibitors are the current drug therapeutics available with less reported side effects among T2DM patients. In this context I have focused on designing new potent SGLT2 inhibitors taking the structural insights of the protein and binding conformations into consideration.

Chapter 02

AIMS AND OBJECTIVES

The objectives of the study are as follows:

- Selection of SGLT2 protein with good resolution
- Designing a library of compounds using interactive enumeration in Schrodinger.
- Virtual screening of designed library (database) of compounds as SGLT2 inhibitors.

Chapter 03

REVIEW OF LITERATURE

2.1 Sodium glucose co-transporters (SGLTs)

SGLTs are membrane transport proteins mostly associated with transportation of glucose molecules. The solute carrier family is one of the largest families of membrane transporters with about 400 members distributed among 65 sub-families (Hediger *et al.*, 2004; Perland and Fredriksson, 2017). There are 12 members in solute carrier family 5 (SLC5) responsible for Na+ coupled co-transportation of sugars, myo-inositol, iodide, short-chain fatty acids, and choline (Wright *et al.*, 2011). Among these six members are named as sodium glucose co-transporters and were reported to show varied preferences for glucose binding (table 4). However, only two members of SGLTs namely SGLT1 and SGLT2 are extensively studied (Gallo *et al.*, 2015). The association of SGLT2 and type 2 diabetes mellitus has recently been established. At present, SGLT2 inhibitors have evolved as promising therapeutics for the treatment of T2DM.

Table 4. Tabular representation of the SGLTs along with their discriminatory characteristics.

Protein (Gene)	Chromosomal location	Organ of expression	Natural Substrate	Biological role
SGLT1 (*SLC5A1*)	22q12.3	Mostly in Small intestine, but also in kidney and heart	Glucose and Glucose	Absorption of dietary glucose and glucose
SGLT2 (*SLC5A2*)	16p11.2	Mostly in kidney	Glucose	Reabsorption of renal glucose
SGLT3 (*SLC5A4*)	22q12.3	Mostly in small intestine	Gets activated by imino sugars	Functions as a glucose sensor. Also mediates cellular influx of Na+

SGLT4 (*SLC5A9*)	1p33	Mostly in small intestine	Mannose, glucose and fructose	Absorption of mannose, glucose and fructose
SGLT5 (*SLC5A10*)	17p11.2	Mostly in kidney and retina	Mannose, fructose, glucose and glucose	Mostly reabsorb mannose and fructose and show lesser affinity for glucose and glucose
SGLT6 (*SLC5A11*)	16p12.1	Small intestine, brain and kidney	Myo-inositol and D-chiro-inositol	Mostly involved in sodium dependent co-transportation of myo-inositol and D-chiro-inositol

SGLT1 mostly contributes to the intestinal absorption of dietary glucose and glucose. However, it also plays role in renal reabsorption of glucose. SGLT2 is exclusively expressed in the renal tubule and regulates renal reabsorption of glucose. It has recently evolved as a drug target for regulating plasma glucose levels in type 2 diabetes mellitus. On the other hand, human SGLT3 is a glucose sensor without any role in co–transportation of glucose. It is highly expressed in the enteric nervous system and muscles with a high affinity for imino sugars. SGLT4 and SGLT5, however, were reported to exhibit lower affinity towards glucose and higher affinity for mannose transportation. Moreover, SGLT6 has been reported to prefer D-chiro-inositol as its substrate for transportation (Gallo *et al.*, 2015).

2.2 Sodium glucose co-transporters (SGLTs) and their association with type 2 diabetes mellitus (T2DM)

In the recent past kidney has been identified as an essential organ to improve glycemic control in T2DM patients. Researchers began studying the renal reabsorbtion of glucose in association with T2DM. It is a homeostatic process of

retaining calories and maintaining a sustained elevated plasma glucose level. However, inhibiting glucose reabsorbtion and promoting glycosuria is thought to be a potential approach towards T2DM prevention. Glucose excretion coupled with reabsorbtion blockade can prevent hyperglycemia and thereby T2DM. Under normal conditions, about 180g of plasma glucose gets filtered and reabsorbed completely in the proximal convoluted tubule (PCT) of nephron (Chao and Henry, 2010). This reabsorbtion is facilitated by certain carrier proteins; sodium glucose co-transporters (SGLTs) (Chao and Henry, 2010; Zou *et al.*, 2017). Therefore, a pharmacological blockade of SGLT has evolved as a potential therapeutic means for treating T2DM patients (Zou *et al.*, 2017).

SGLT2 and SGLT1 are the two well known isoforms responsible for active transportation of glucose coupled with sodium ion (Harada and Inagaki, 2012). The reabsorbed glucose molecules pass through the intracellular space and reach the bloodstream via glucose transporter GLUT2 (also called SLC2A2) (Figure 1). Studies have shown elevated expression of SGLT2 and GLUT2 in the isolated PCT cells of T2DM patients (Bautista *et al.*, 2004; de Leeuw and de Boer, 2016; Rahmoune *et al.*, 2005).

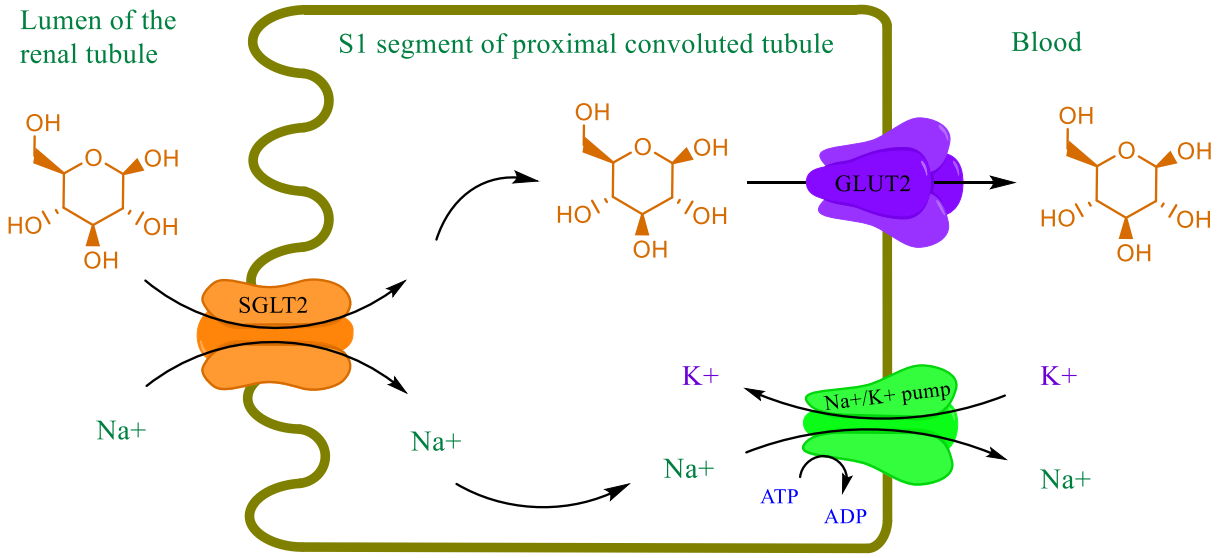


Figure 1. The figure represents the mechanism of action of SGLT2. Renal glucose is uptaken by the epithelial cells of Proximal Convoluted Tubule (PCT)

coupled with sodium ion against the concentration gradient (active glucose transportation). The sodium ion is, further, expelled into blood stream through ATP sensitive sodium–potassium transport channel with a simultaneous inflow of potassium ion. In a similar manner the reabsorbed glucose is also released into the blood stream by means of GLUT2 (facilitative glucose transportation).

About 90% of the glucose reabsorbtion is attributed to SGLT2 (Chao and Henry, 2010). However, the SGLT transport system has its own maximum transportation capacity known as transport maximum (T_{max}). The filtered load get excreted out along with urine when exceeds the threshold value; transport maximum (T_{max}) due to saturation of SGLTs. Therefore, an SGLT2 inhibitor can act as a potential drug for T2DM by lowering the T_{max} of the nephron and promoting glycosuria (Chao and Henry, 2010). Furthermore, SGLT2 inhibitors were expected to cause no renal adverse effects as individuals with familial renal glycosuria remained normal even after loss of glucose at the rate of 50g per day. Following this, in the recent past, several SGLT2 inhibitors have been developed (Chao and Henry, 2010). However, the unexpected renal side effects were yet reported among certain patients.

SGLT1 is situated at the distal end of PCT and accounts for only about 10% glucose reabsorbtion (Figure 2). However, SGLT2 situated at the proximal end of PCT accounts for about 90% of the reabsorbtion (Chao and Henry, 2010). Therefore, SGLT2 is considered to be a transport protein with high capacity yet low affinity for glucose (Chao and Henry, 2010) (table 5). This review primarily focuses on the essential role played by SGLT2 in controlling T2DM.

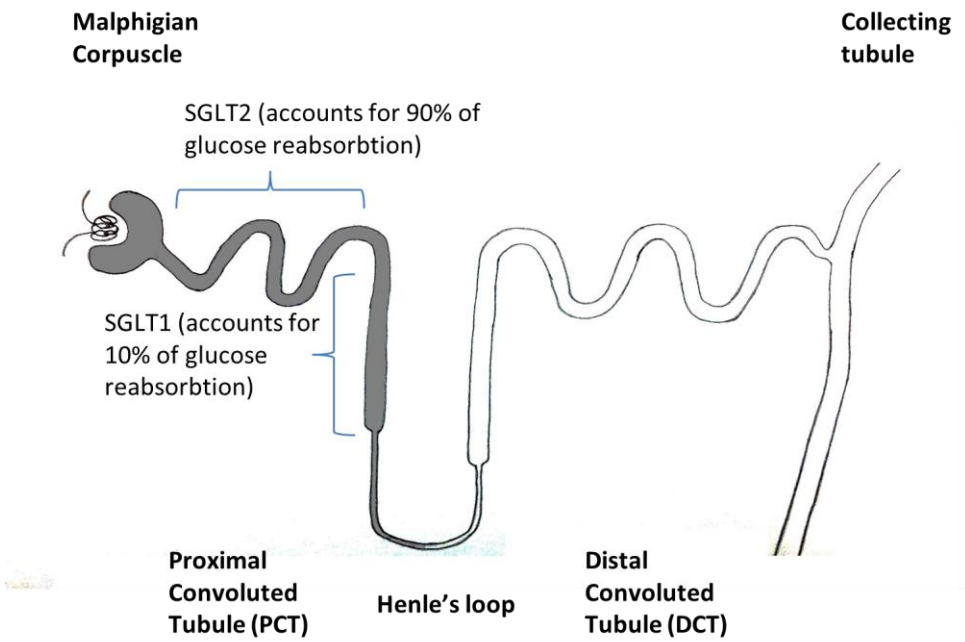


Figure 2. The figure shows the structure of a nephron and the location of the transporter protein involved in glucose reabsorbtion (SGLT1 and SGLT2). Structurally a nephron has three major regions ranging from the Malphigian corpuscle to the Collecting duct. Among those three regions both the SGLT proteins are found in Proximal Convoluted Tubule (PCT). SGLT2 is expressed in the earlier segments of PCT (S1 segments) whereas SGLT1 remains expressed in the later segments of PCT (S2/S3 segments).

SGLT2 is encoded by the SLC5A2 gene (Solute carrier family 5, member 2) located on p11.2 region of human chromosome 16. The gene contains 14 separate exons (~8kb) encoding a protein with 672 amino acids and remains associated with several rare mutations (Magen *et al.*, 2005). Missense, Non-sense and Frameshift mutations often cause functional impairments in the SGLT2 protein and consequently lead to familial renal glycosuria (FRG). FRG is an autosomal recessive disorder and is restricted to certain families or populations. Most of the FRG causing mutations are private in nature (Zimdahl *et al.*, 2017). However, these mutations hold beneficial effects for certain other disorders. Loss of excess calories through urinary excretion reduces the risk of diabetes and overweight. Therefore, these mutations are likely to be infrequent among the diabetic patients. Mutation in the SGLT1 encoding SLC5A1

gene (22q12.3) is mostly associated with glucose–glucose malabsorption syndromes. Furthermore, several SNPs have been discovered in the non coding regions of SLC5A2 gene with no influence on the functional integrity of the encoded protein. Henceforth, no significant established impact of the SGLT2 polymorphic markers has been observed upon type 2 diabetes mellitus. However, nominal association of rs9934336 and rs3116150 has been observed with plasma glucose concentration. Studies have also shown elevated expression of SGLT2 and GLUT2 in the isolated PCT cells of T2DM patients (Bautista *et al.*, 2004; de Leeuw and de Boer, 2016; Rahmoune *et al.*, 2005). Renal cells of T2DM patients showed significantly higher expression of SGLT2 compared to the non–diabetic individuals in a hyperglycemic medium (López *et al.*, 2010). Therefore, pharmacological blockade of SGLT2 has recently evolved as an apparent therapeutic means for treating T2DM patients (Zimdahl *et al.*, 2017).

Table 5. A comparison between SGLT1 and SGLT2 based on certain discriminatory characteristics.

	SGLT1 (*SLC5A1*)	**SGLT2** (*SLC5A2*)
Sites of expression	Usually in small intestine, but also found in kidney and heart.	Found in kidney with almost exclusivity.
Location in renal tubule	Later segment of Proximal Convoluted Tubule (S2/S3 segment)	Early segment of Proximal Convoluted Tubule (S1 segment)
Affinity towards glucose	High affinity with K_m = 0.4 mM	Low affinity with K_m = 2 mM
Glucose transportation potential	Low capacity	High capacity
Renal glucose reabsorbtion percentage	About 10%	About 90%

Chromosomal location	22q12.3	16p11.2
Gene size	69,998 bp	8,326 bp
Protein size	664 amino acids	672 amino acids
Molecular weight of protein	73498 Da	72897 Da

2.3 SGLT2 inhibitors

2.3.1 Phlorizin

For the very first time Phlorizin was isolated in 1835 from the root bark of apple tree. It had proven to be a potent but non selective inhibitor of both SGLT1 and SGLT2. Phlorizin was the first drug used in multiple studies to establish the relation between hyperglycemia and T2DM. Use of Phlorizin in controlled and diabetic (pancreactomised) rats demonstrated the affect of glucotoxicity on insulin sensitivity. However, the drug failed to be approved as a potent inhibitor due to low bio-availability and poor intestinal absorption (Chao and Henry, 2010). Structurally phlorizin has O-glucosidic linkage that makes it vulnerable to β-glucosidase mediated degradation, thus, reducing the bio-availability. Moreover, the drug gets metabolized into phloretin that inhibits GLUT1 and reduces intestinal absorption of glucose. Phlorizin also non - selectively inhibits intestinal SGLT1 and causes gastro–intestinal side effects (Chao and Henry, 2010). Therefore, significant structural modifications were done on both glyconic and aglyconic moieties to design selective inhibitors for SGLT2. T-1095 was the first selective SGLT2 inhibitor designed after exploring the structure–activity relationship (SAR) of phlorizin (Cai *et al.*, 2015). Substitution of phenol rings with other compounds enhanced the selectivity of the inhibitors along with their half life (Abdul-Ghani *et al.*, 2015).

2.3.2 Remogliflozin

Remogliflozin is a glucoside-based benzylpyrazole compound with selective SGLT2 inhibitory nature. It is proposed in the form of a prodrug named remogliflozin etabonate which metabolizes into active remogliflozin inside the body. The drug had shown significant reduction of fasting blood glucose level and glycated hemoglobin (HbA1c) in certain studies. However, the bio-availability of remogliflozin is very low, probably due to the presence of O–glucosidic linkage, which has rendered the drug currently under development (Madaan *et al.*, 2016).

2.3.3 Sergliflozin

Sergliflozin is a glucoside–based benzyl phenol compound with selective SGLT2 inhibitory nature (Katsuno *et al.*, 2007). However, it remained as an investigational drug with no further development after phase II clinical trial. In most of the studies, sergliflozin showed sustained suppression of glucose reabsorbtion in nephron and also displayed reduced body weight during the study period. It is also proposed with an attached methyl carbonate group (etabonate) which later dissociates to release the active compound inside the body (Katsuno *et al.*, 2007).

Although the above mentioned drugs were discontinued due to their unfavorable efficacy and nature of side effects yet many efficient drugs have recently been introduced into the market. Drugs such as Dapagliflozin, Canagliflozin, Ertugliflozin and Empagliflozin are agents currently approved by the FDA. Other SGLT2 inhibitors in the drug classes Ipragliflozin, Tofogliflozin and Luseogliflozin are approved in Japan.

2.3.4 Dapagliflozin

Dapagliflozin was the first SGLT2 inhibitor approved by the European Union on November 12, 2012 in the world. However, concerns over the associated side effects such as liver damage, urinary bladder cancer lead to its short term withdrawal. The resolved drug was approved by the US FDA and re-introduced into the market under the brand name ‘Farxiga’ on January 8, 2014 (Madaan *et al.*, 2016). AstraZeneca

Pharma India Ltd and Sun Pharmaceutical Industries Limited have recently joined hands with the agreement to distribute Dapagliflozin drugs into the Indian market at affordable price (Bureau, 2016). The drug is usually recommended with a primary dose of 5 mg which can further be increased to 10 mg. Dapagliflozin has also been approved by the FDA to be used in combination with biguanides as anti-diabetic agent. Dapagliflozin is metabolized by UGT1A9 encoded protein into dapagliflozin 3-O-glucuronide in the liver and kidneys. Cytochrome P450 (CYP) mediated metabolism also contributes to a minor extent in humans. Dapagliflozin has shown increased urinary glucose excretion, weight loss among diabetic patients. Clinical trials are currently under progress to check the efficacy of the drug in cardiovascular diseases. However, Dapagliflozin tends to cause adverse affects by decreasing the glomerular filtration rate (GFR) among patients suffering from renal impairment.

2.3.5 Canagliflozin

Canagliflozin received its approval from US FDA on March 29, 2013. It was the first SGLT2 inhibitor introduced in USA and Canada. Janssen (a division of Johnson and Johnson) introduced the drug under the brand name 'Invokana'. The drug is metabolized into inactive O-glucoronides in the liver mainly by the enzymes UGT1A9 and UGT2B4 (Madaan *et al.*, 2016). However, CYP3A4 mediated metabolism also contributse to a certain extent; about 7% in humans. The drug is available in two different doses of 100 mg as well as 300 mg and needed to be taken orally once daily by the patients. Among these 100 mg is the recommended dose to begin with which can further be increased up to 300 mg. Administration of canagliflozin have shown successful results in the T2DM patients (Madaan *et al.*, 2016). The outcomes were also supportive for the patients with moderate hepatic and renal impairment. Current studies are yet in progress to assess the efficacy of Canagliflozin with cardiovascular diseases in T2DM patients. However, contraindications were observed for the patients with severe renal dysfunction.

2.3.6 Empagliflozin

Empagliflozin was developed by the Boehringer Ingelheim and Eli Lily & Company. After being approved by the US FDA in August, 2014 the drug is marketed under the brand name ‘Jardiance’. Recently, Lupin Limited and Boehringer Ingelheim have joined hands to co-market Empagliflozin so as to tackle the increased prevalence of T2DM. In vitro studies have indicated that UGT2B7, UGT1A3, UGT1A8, and UGT1A9 play chief role in glucuronidation of the drug (Madaan *et al.*, 2016). Empagliflozin has a starting recommended dose of 10 mg which can further be increased up to 25 mg. Similar to Canagliflozin, Empagliflozin has also shown positive indications in T2DM patients. Although contraindicated to T2DM patients in conjunction with severe renal dysfunction, this drug have shown indications to reduce cardiovascular events.

Ertugliflozin is a recently approved SGLT2i which is available in the market. Structurally it contains C-glucoside linkage with an additional cyclic ether ring and aglycone similar to that of dapagliflozin.

2.3.7 Ipragliflozin

Ipragliflozin was the first SGLT2 inhibitor to be approved in Japan by 2014. However, clinical trials are yet in process to check the safety of the drug before approving globally. It is available under the brand name ‘Suglat’ in Japan (Madaan *et al.*, 2016). With recommended dose of 50 mg and 100 mg, Ipragliflozin had shown improved glycemic control in T2DM patients with mild renal impairment. However, contraindications were observed for those with severe renal dysfunction. The drug metabolism occurs through glucuronidation as in the other drugs of the class. Studies have shown indications for the drug to alleviate glycemic control in T1DM patients (Madaan *et al.*, 2016).

2.3.8 Tofogliflozin

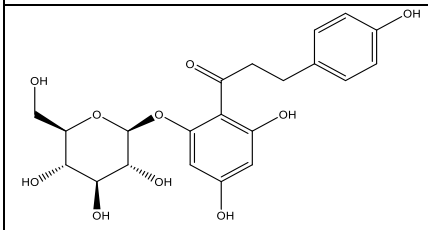
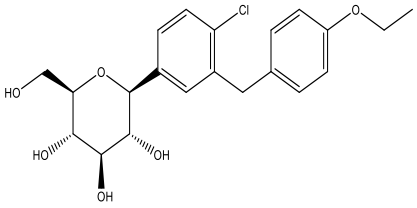
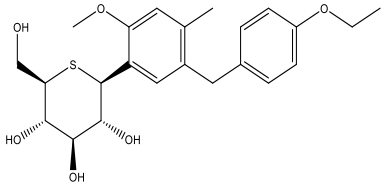
Tofogliflozin is another SGLT2 inhibitor approved globally by Japan FDA in March 2014. It is marketed under the brand name ‘Deberza’ in the markets of Japan, US and European Union with 20 mg oral drug dose. Although a brief study on its efficacy is

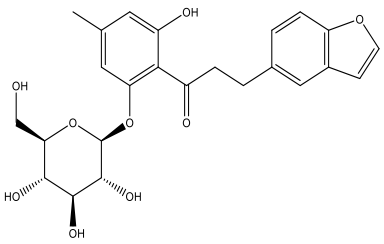
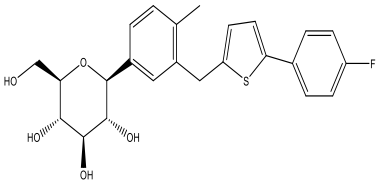
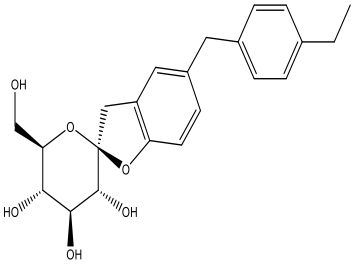
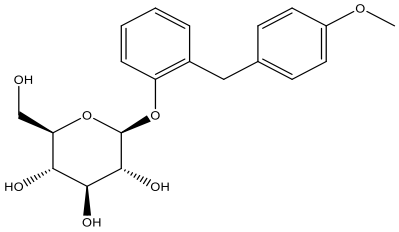
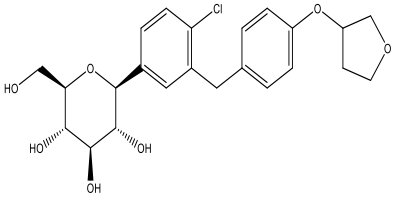
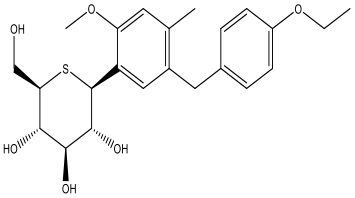
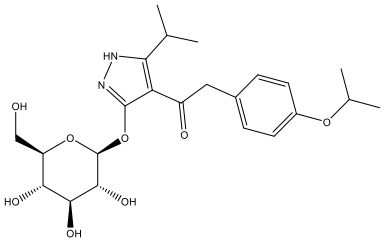
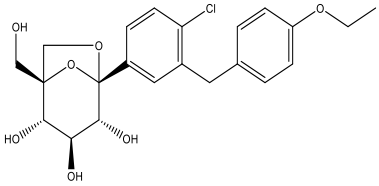
yet in progress, the treatment results are expected to deplete with increase in renal impairment (Madaan *et al.*, 2016). However, a study on diabetic mice has demonstrated alleviation in renal and pancreatic function (Nagata *et al.*, 2013).

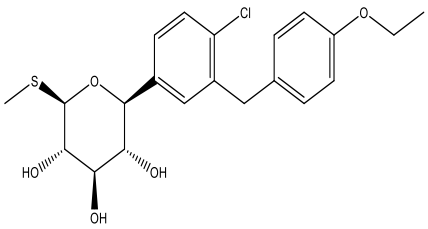
2.3.9 Luseogliflozin

Luseogliflozin was approved by Japan FDA in March 2014. It is marketed by Taisho pharmaceuticals and Novartis under the brand name 'Lusefi' with recommended oral doses of 2.5 mg and 5.0 mg. Although the efficacy of the drug falls with increasing renal impairment yet studies have shown the effect of renal dysfunction, over drug action, only on post prandial glucose level. However, the fasting glucose level in the same population remained unaffected (Madaan *et al.*, 2016).

Table 6. A structural comparison between SGLT2 inhibitors with O-linkage and C-linkage.

SGLT2 inhibitors with O–aryl glucosidic linkage (discontinued drugs or currently under development)	**SGLT2 inhibitors with C–aryl glucosidic linkage (available in the market)**	**SGLT2 inhibitors with C–C–aryl glucosidic linkage (approved mostly in Japan)**
 phlorizin (discontinued)	 dapagliflozin	 ipragliflozin

 T – 1095 (discontinued)	 canagliflozin	 tofogliflozin
 sergliflozin (discontinued)	 empagliflozin	 luseogliflozin
 remogliflozin (under development)	 ertugliflozin	

	 sotagliflozin (under development)	

Chapter 04

MATERIALS AND METHODS

3.1 Procurement of 3D structure of protein

The 3D structure of SGLT2 was obtained from Protein Data Bank (www.rcsb.org) with PDB code 3DH4. Although protein structures can easily be obtained from PDB yet the criteria to select a reliable 3D protein structure depends upon the resolution of protein and protein-ligand interactions.

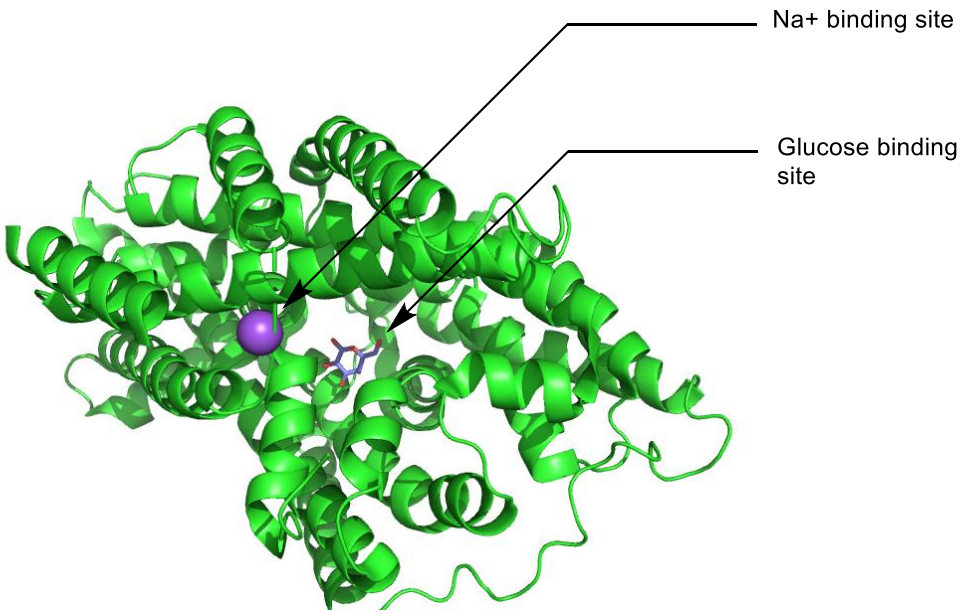


Figure 3. The 3D structure of a single subunit of SGLT2 (obtained from RCSB PDB with PDB ID – 3DH4) with the co crystallized glucose.

The procured SGLT2 protein (PDB code – 3DH4) showed a tetrameric structure with three ligand molecules; Erbium ion, sodium ion and β–D–glucose along with a unique L–peptide linking. The protein was reported to be isolated from *Vibrio parahaemolyticus* with a resolution of 2.7 Å and each ligand molecules located far apart preventing any untoward interaction. Structurally the protein is a transmembrane protein with 14 membrane–spanning helices with both the amino and carboxyl termini facing the periplasm. Among the 14 helices, 10 helices are known to form the structural core of the protein and show inverted topology in accordance with the plane of the membrane. The glucose molecule, however, remains bound to the centre of protein core establishing unique interactions with certain hydrophobic residues (Faham *et al.*, 2008) (Figure 4). Moreover, the protein structure with the

above mentioned PDB ID is the only crystallized structure available for SGLT2 protein.

Binding sites were identified on the basis of the associated co-crystallized ligand (glucose) using Schrodinger molecular modeling environment (Maestro).

3.2 Structure Drawing and database generation of ligand molecules

The structures of all ligand molecules were created using Maestro interactive enumeration. Maestro v11.3 (2017-4) was used to create an entire database of ligands taking empagliflozin as the standard reference molecule. A fragment library with a total of 45 fragments was created prior to ligand preparation on the basis of available literature. Ligand preparation wizard was employed to prepare the standard empagliflozin molecule for enumeration. The attachment bonds for enumeration were selected (Figure 4) and then enumerated with all the fragments (table 7). The conformations of the ligand molecules were subsequently optimized by OPLS3 (Optimized Potential for Liquid Simulations version 3) force field.

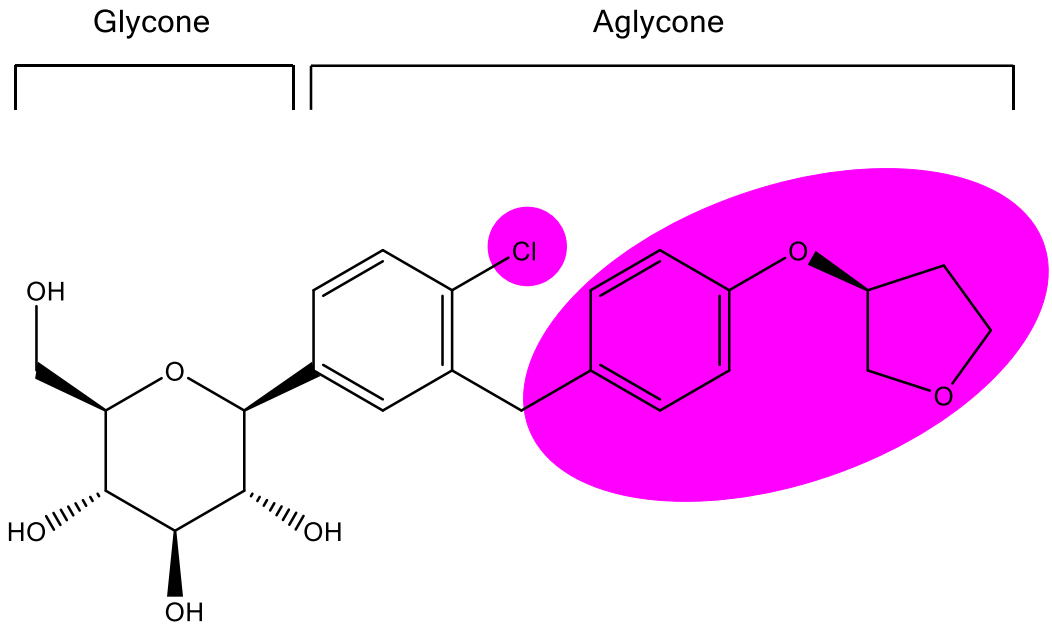
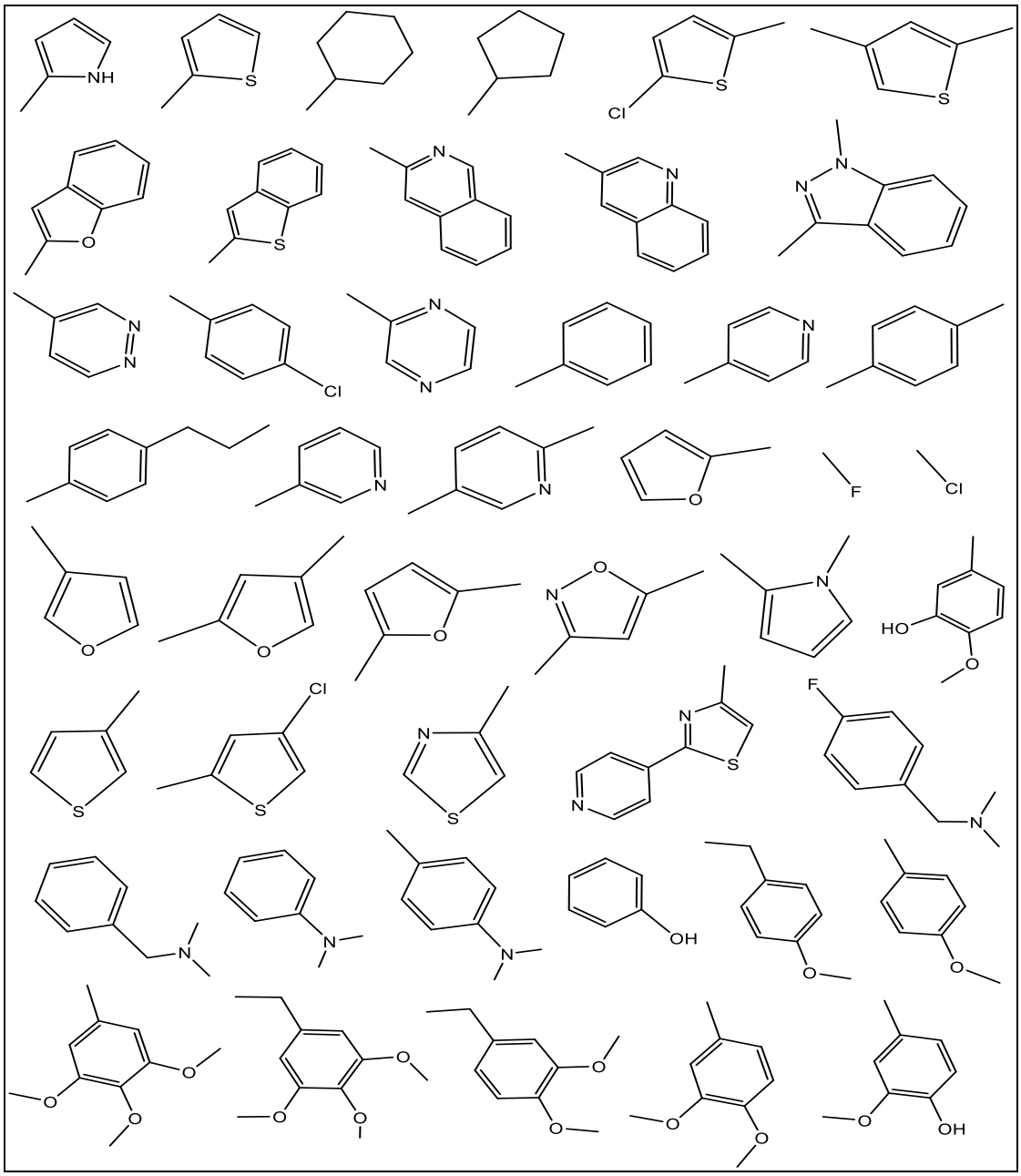


Figure 4. Structure of empagliflozin (the standard molecule). The coloured (highlighted) regions indicate the positions where the fragments are enumerated. The process of enumeration appends the fragments randomly one

after another to the pre-defined positions and generates the ligands. Moreover, the stable conformations of the enumerated ligand molecules are obtained using OPLS3 force field.

As a result of the interactive enumeration a total of 3250 ligand molecules were produced. The ligands were then docked on the prepared protein using Maestro v11.3 (2017-4).

Table 7. A tabular representation of the fragments constructed for ligand database generation using Maestro v11.3 (2017-4).



3.3 Protein preparation

Protein structures obtained from PDB often bear heavy atoms, water molecules, co-factors, metal ions and multimeric structures. Moreover, they do not show any

information on bond orders, formal atomic charges or topologies. X–ray structure analysis often fails to distinguish between O and NH_2 conferring structures with misaligned terminal amide groups. Furthermore, ionization and tautomeric states also remain unassigned in general. Therefore, the protein structures are converted to certain suitable forms prior to docking using force fields. The Glide (Grid based ligand docking with energetics) module of Maestro v11.3 requires well assigned bond orders and ionization states and performs better in the presence of re-oriented side chains and reduced steric clashes. Maestro v11.3 (2017-4) is installed with *protein preparation wizard* to get rid of water molecules, peptide substrates and re-establish the bond orders and other topologies. The protein preparation wizard has three essential components; *Import and process*, *Review and modify* and *Refine*. OPLS3 force field was employed for energy minimization.

3.4 Receptor Grid Generation

Receptor Grid Generation is an essential method of Glide that requires a prepared protein structure with appropriate bond orders and formal charges. Ligand-receptor interactions normally take place at certain specific binding sites. Presence of any co-crystallized ligand aids in determining the position and size of the active sites. Grid generation represents the shape and properties of the receptor and subsequently defines the active site for the ligand-receptor interaction.

3.5 Docking

The process of molecular docking for all the ligand molecules generated was done using Maestro v11.3 (2017-4). The co-crystallized ligand glucose was split from the procured protein and docked for analyzing the interactions. The receptor grid file generated defines the active site for ligand docking.

The Glide module of Maestro follows a stringent hierarchical protocol to dock the various ligand poses and accordingly rank them based on their scores. The Glide docking hierarchy mainly consists of the four essential stages (Figure 5). In the first stage, the module searches and selects the “site points” essential for docking. The equally spaced 2 Å receptor grid identifies the active site and selects all the possible

site points. For a given site point, the grid measures the distance from the site point to the receptor surface and compares the same with the distance from the ligand center (midpoint of the ligand diameter) to the ligand surface. In case of a good match, the ligand center is positioned at the site point or the same is skipped if found not enough good matches. In the second stage, a “diameter test” is followed so as to select the possible orientation of the ligands. For this a ligand diameter is drawn between the widely spaced atoms within the ligand structure. Ligand diameter helps in identifying the orientations with less steric clashes. Subsequently, “the subset test” comes with a subset of ligand atoms capable of making H-bonds or ligand-metal interactions with the receptor. In case the subset gives a favorable score with the receptor, the module further begins to evaluate all interaction scores with the receptor. The next sub stage uses ChemScore empirical scoring function to recognize the favorable hydrophobic, H-bonding and ligand-metal interactions along with penalizing the steric clashes. At this stage the atoms, irrespective of their position and orientation, prefer to move along x, y or z axis for about ±1 Å in order to attain the best possible score. Therefore, this stage is also known as “greedy scoring”. Finally the last sub stage of stage 2 involves re-scoring the ligand poses with top greedy score using a procedure of “refinement”. However, only a small number of top best refined poses are allowed to enter into the next stage of energy minimization. Close interatomic contacts often results in large potential energy and gradient terms leading to lower stability of the molecule. Glide uses OPLS (Optimized Potential for Liquid Simulation) force field to reduce the energy and obtain stable molecule with minimum energy. Finally, the minimized poses are re-scored using Schrodinger’s proprietary GlideScore (GScore) scoring function.

GScore = (0.05 x vdW) + (0.15 x Coul) + Lipo + Hbond + Metal + Rewards + RotB + Site

Where,

vdW = Van der Waals energy

Coul = Coulomb energy

Lipo = Lipophilic term

Hbond = Hydrogen-bonding term

Metal = Metal-binding term

Rewards = Rewards and penalties for various features

RotB = rotatable bonds penalty

Site = Polar interactions

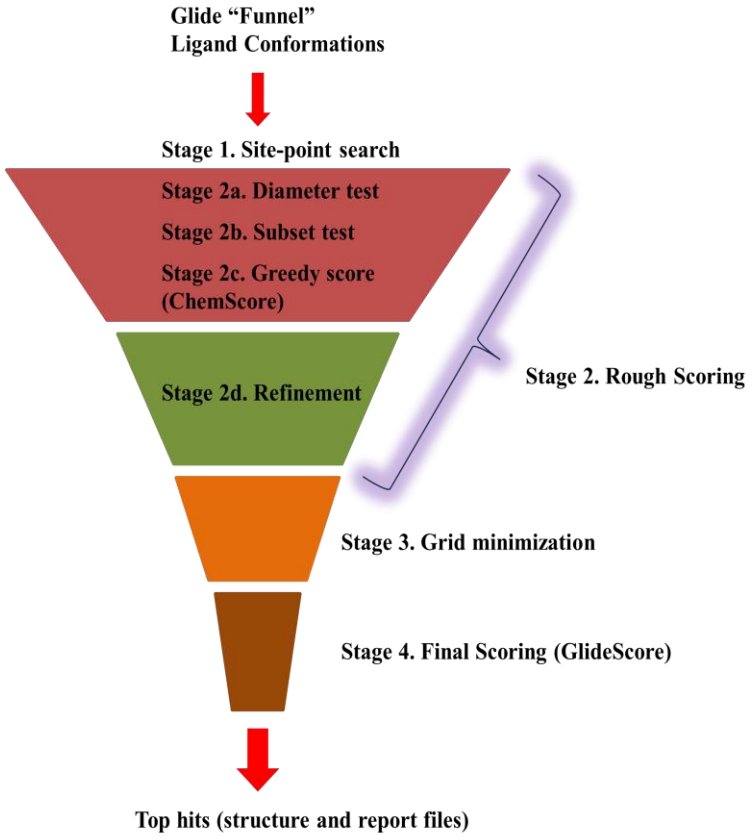


Figure 5. The Glide docking hierarchy

However, the precision of docking is determined by three other key components; HTVS, SP and XP. HTVS (High Throughput Virtual Screening) mode is used to rapidly screen a large number of ligands. Conformational sampling of HTVS is

restricted more than that in SP docking. SP (Standard Precision) mode is suitable for large number of ligands with unknown quality. SP is the method often recommended for molecular docking followed by XP docking of the ligand poses with high score. Extra-precision (XP) mode requires considerably more CPU time and hence often recommended to dock the top-scoring ligand poses after the completion of SP dock run.

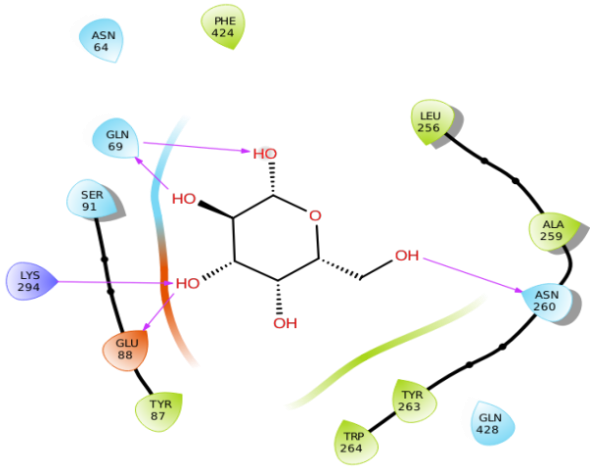
Chapter 05

RESULTS AND DISCUSSION

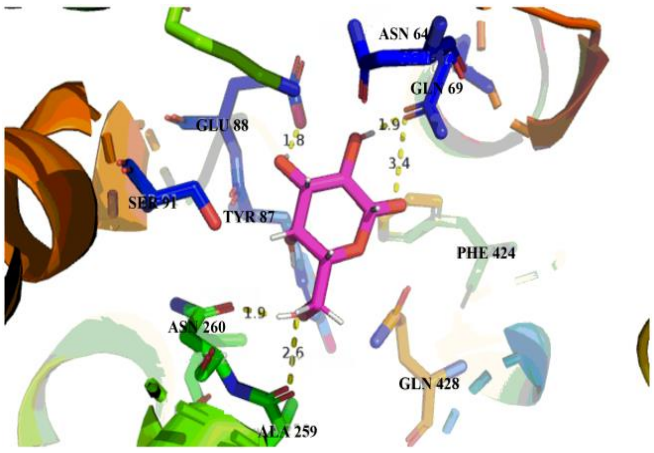
4.1 Binding site identification

The binding site of SGLT2 (PDB ID – 3DH4) was identified using the available data from RCSB PDB. The protein is made up of four subunits with a sequence length of 530 amino acids. The ligand glucose molecule is associated with all four subunits and acts as an agonist triggering protein function. The binding site of glucose, therefore, was identified using Schrodinger Maestro after procuring the 3D structure of protein from the database (refer to figure 3). The various amino acid residues which interact with ligand glucose are shown in Figure 6.

A)



B)



C)

Ligand	**Hydrogen bond Interactions**
Glucose	GLN428, GLN69, ASN64, GLU88, LYS294, TYR87, SER91, TRP264 and ASN260

Figure 6. (A) 2D interaction diagram of ligand glucose molecule (obtained using Schrodinger Maestro v11.3) representing hydrogen bond interactions with the amino acid residues of the binding site. (B) 3D interaction diagram of ligand glucose molecule (obtained using PyMOL) representing bond length (in Å) in the binding site of SGLT2. C) Tabular representation of the various amino acid residues reported to interact with the ligand glucose molecule through hydrogen bond.

4.2 Docking study

Molecular docking was performed using Schrodinger Maestro v11.3. The SGLT2 (PDB ID – 3DH4) was docked with ligand database of 3250 molecules. Empagliflozin, dapagliflozin, ertugliflozin and canagliflozin were also docked as the available standard molecules. Co-crystallized ligand glucose was also docked for the validation of the docking method. GlideScore was used as the scoring function. Ten binding conformations were generated for each ligand molecule. The conformation with highest GlideScore was selected as the optimum binding pose (best fit conformation) or hit molecule.

The ligand glucose molecule (docking score: -7.787) has shown major interactions with Gln 69, Lys 294, Glu 88 and Asn 260 (figure 6A). The interaction data obtained from the database validates the aforementioned residues to be among the major interacting residues of glucose molecule in the binding site of SGLT2 (figure 6C). Moreover, the ligand glucose molecule has shown an RMSD of 0.0293 (which is less than 2) validating the reliability of the docking method.

The interactions of empagliflozin, ertugliflozin and canagliflozin complexed with SGLT2 were studied and analyzed. Dapagliflozin, however, failed to interact with the binding site of SGLT2 in our study. Canagliflozin (docking score: -12.070) showed the highest docking score compared to the other standard molecules, that is, empagliflozin (docking score: -9.800) and ertugliflozin (docking score: -8.445). The amino acids interactions established by the glucose ring of empagliflozin are Glu 68 and Ala 63. The aglycones, on the other hand, demonstrated overt interactions with Tyr 263 and Lys 294 (Figure 8). Canagliflozin showed glyconic interactions with Ala 63, Asn 64 and aglyconic interaction with Tyr 263 in our study (Figure 7). Ertugliflozin showed almost similar interactions for the glyconic ring with Asn 64, Ala 63, Ser 66 and aglycone with Tyr 263 (Figure 9).

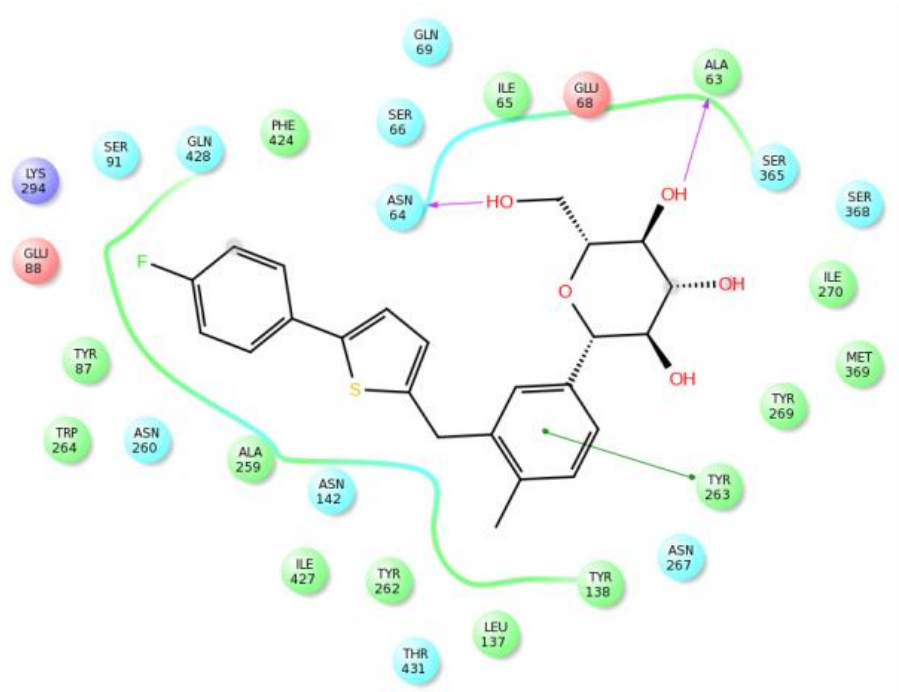


Figure 7. 2D interaction diagram of canagliflozin (docking score: -12.070)

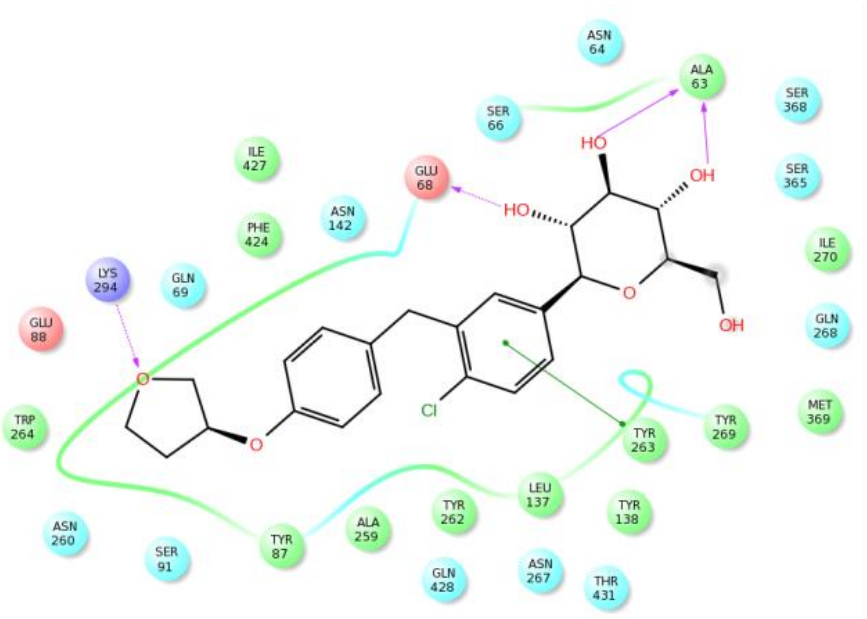


Figure 8. 2D interaction diagram of empagliflozin (docking score: -9.800)

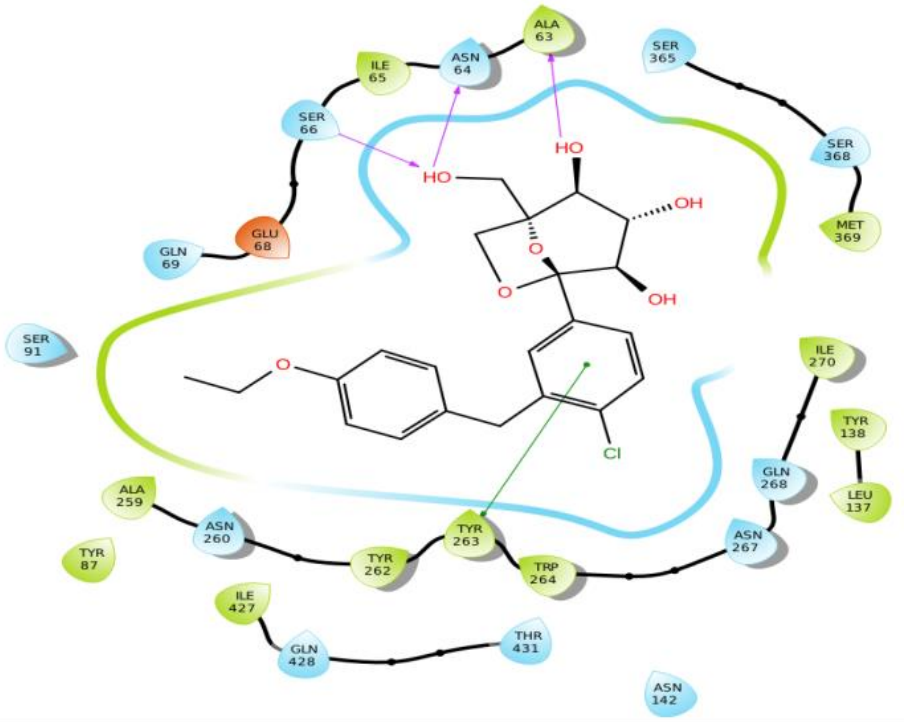
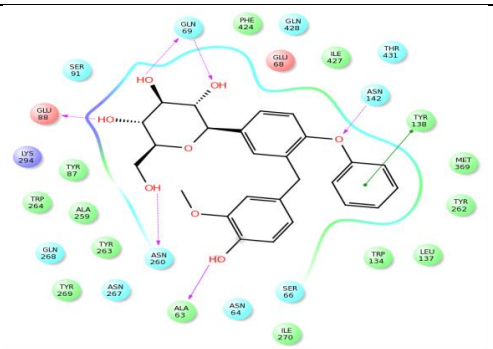
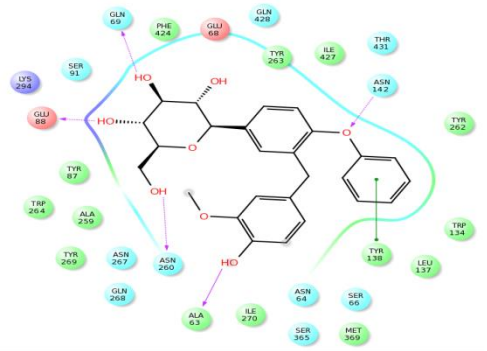
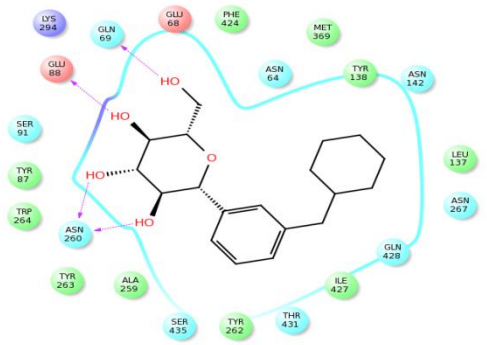
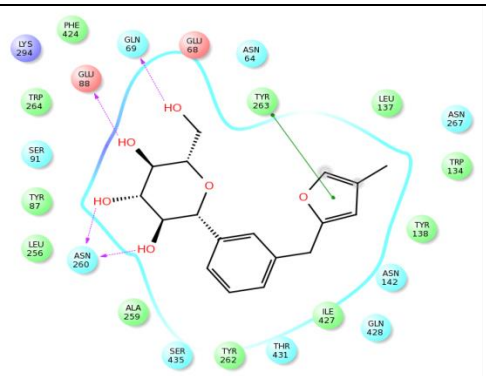
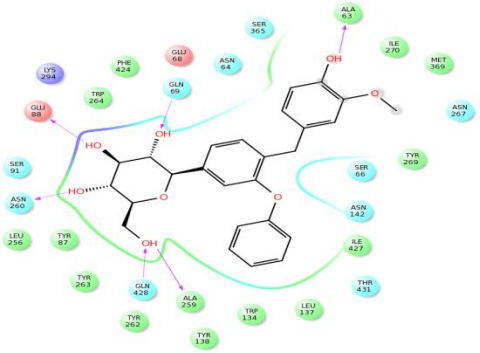
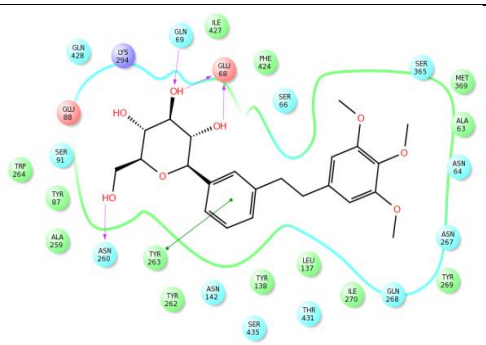
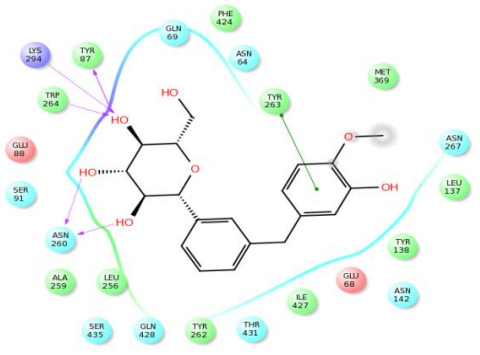


Figure 9. 2D interaction diagram of ertugliflozin (docking score: -8.445)

Virtual screening of the ligand database of 3250 molecules has shown fascinating outcomes. Based on the results all the ligand molecules have been classified into set A molecules and set B molecules as shown in Table 8 and 9 respectively. The molecules in set A have established interactions mostly with Glu 88, Gln 69 and Asn 260 which are, in particular, among the major binding site residues for ligand glucose. Therefore set A molecules have attained binding conformations in a way to negotiate with the preferred pose of ligand glucose and bind in its position. On the other hand, set B molecules demonstrated different binding conformations with Tyr 263, Ala 63 and Asn 64 as the major interacting residues. Moreover, set B molecules have shown higher docking score than the set A molecules.

Table 8. Docking scores and interactions of the set A molecules with best fit conformation.

Ligands	Docking Score	2D interaction diagram	Interacting amino acid residues
A1	-11.664		GLU88, GLN69, ASN260, ASN142, TYR138, ALA63
A2	-11.577		GLU88, GLN69, ASN260, ASN142, TYR138, ALA63
A3	-11.124		GLU88, GLN69, ASN260

A4	-10.808		GLU88, GLN69, ASN260, TYR263
A5	-10.768		GLU88, GLN69, ASN260, GLN428, ALA259, ALA63
A6	-10.421		GLU68, GLN69, ASN260, TYR263
A7	-10.238		TYR87, LYS294, TRP264, ASN260, TYR263

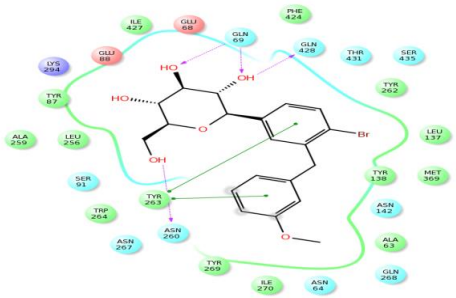
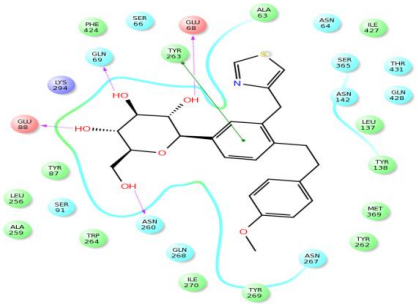
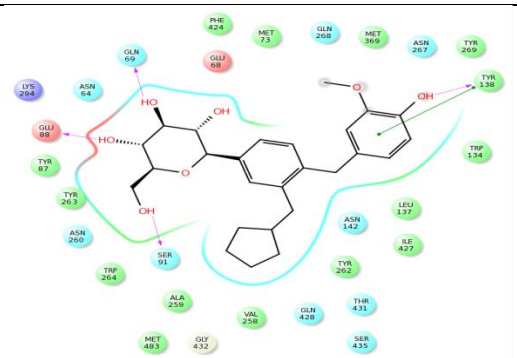
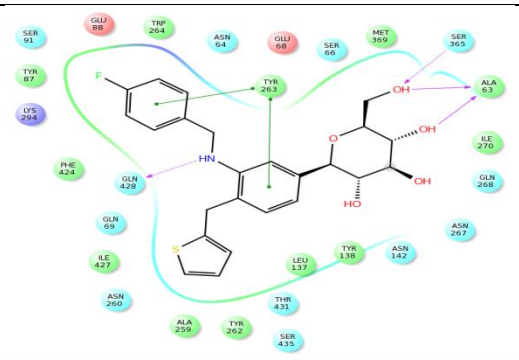
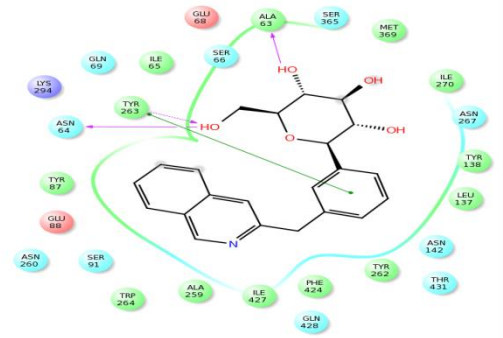
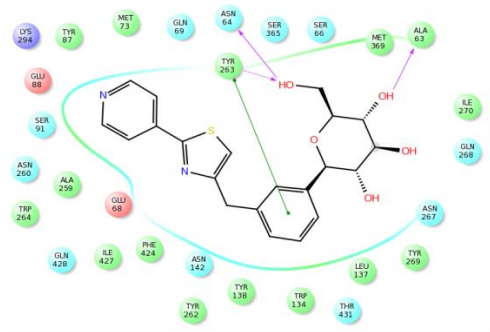
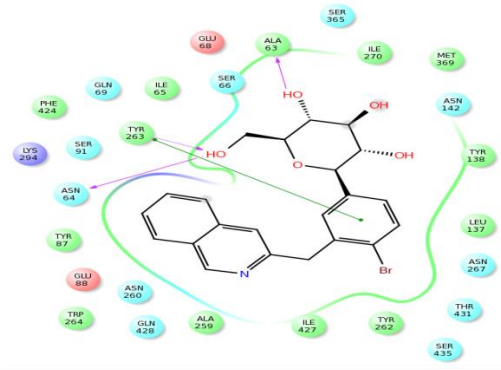
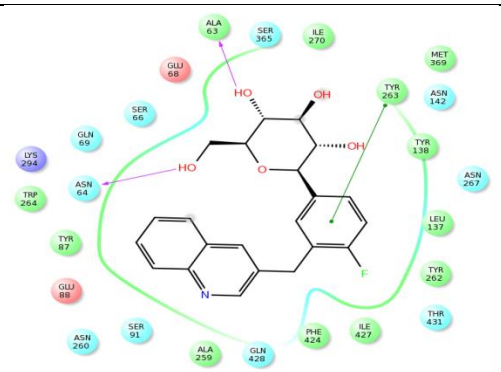
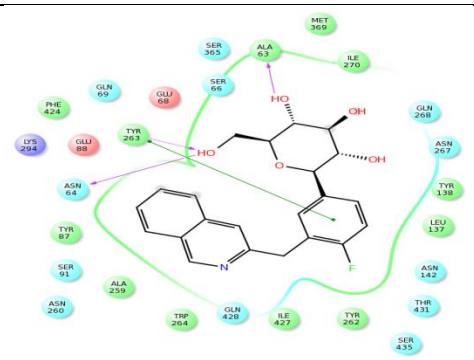
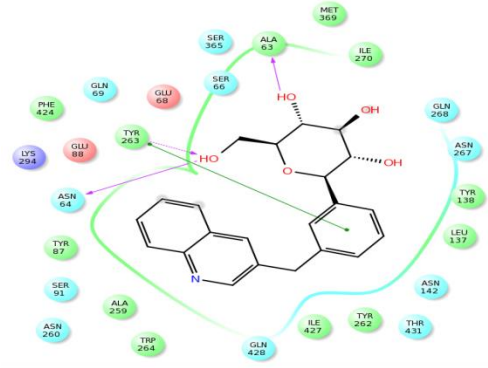
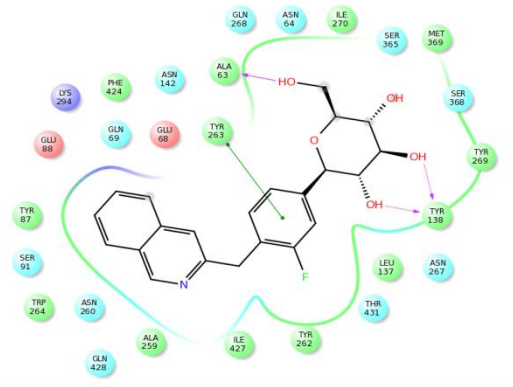
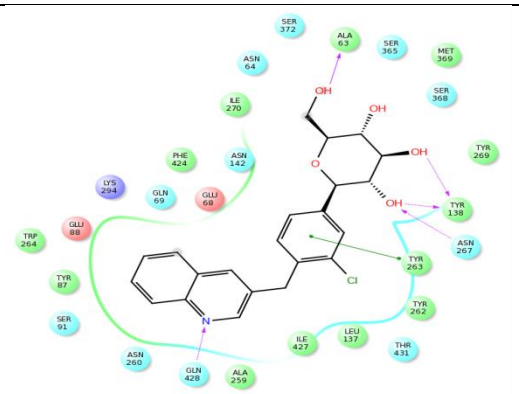
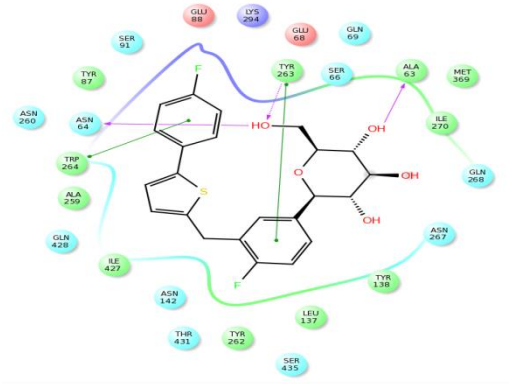
A8	-10.094		GLN69, GLN428, ASN260, TYR263
A9	-10.032		GLU88, GLN69, GLU68, ASN260, TYR263
A10	-10.028		GLU88, GLN69, SER91, TYR138

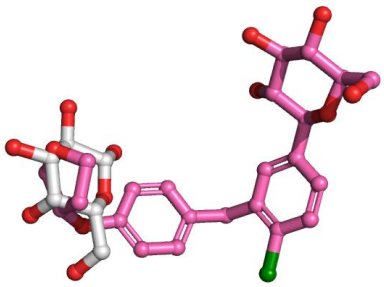
Table 9. Docking scores and interactions of the set B hit molecules with best fit conformation.

Ligands	Docking Score	2D interaction diagram	Interacting amino acid residues
B1	-14.060		GLN428, TYR263, ALA63, SER365
B2	-13.544		ALA63, TYR263, ASN64
B3	-13.514		ALA63, TYR263, ASN64

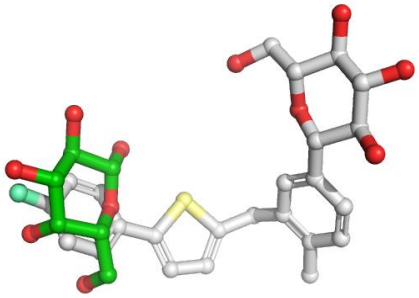
B4	-13.431		ALA63, TYR263, ASN64
B5	-13.412		ALA63, TYR263, ASN64
B6	-13.373		ALA63, TYR263, ASN64
B7	-13.318		ALA63, TYR263, ASN64

B8	-13.068		ALA63, TYR263, TYR138
B9	-13.063		ALA63, TYR263, TYR138, GLN428
B10	-12.870		ALA63, TYR263, ASN64, TRP264

A)



B)



C)

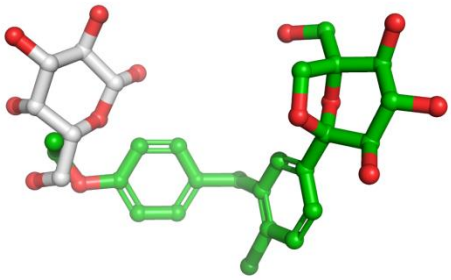


Figure 10. A) A comparison between the binding patterns of empagliflozin and glucose in the binding site of SGLT2. B) A comparison between the binding patterns of canagliflozin and glucose in the binding site of SGLT2. C) A comparison between the binding patterns of ertugliflozin and glucose in the binding site of SGLT2.

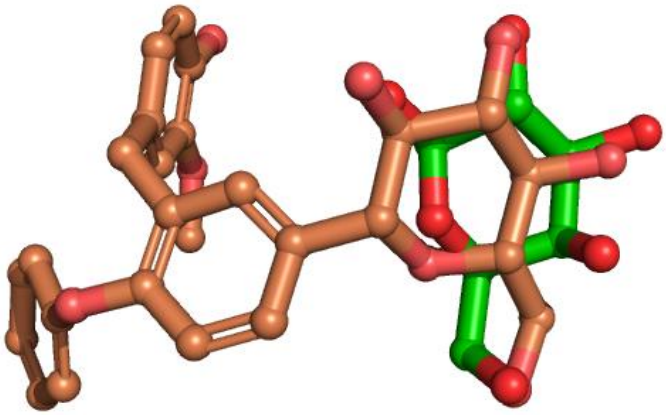


Figure 11. A1 molecule showing glyconic interaction into the glucose binding site of SGLT2.

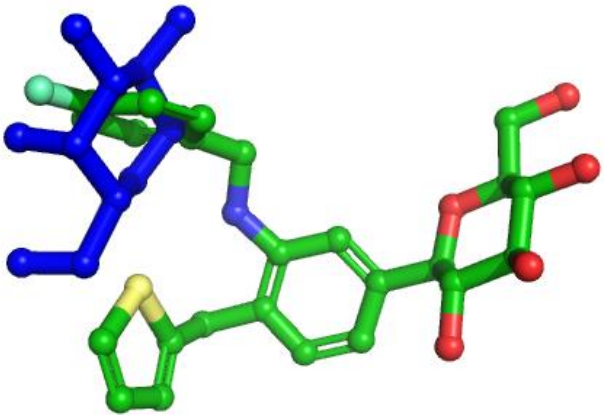


Figure 12. B1 molecule showing aglyconic interaction into the glucose binding site of SGLT2.

Structurally, all the available gliflozins contain a core glucose ring with unique side chains (aglycone). Following their mode of competitive inhibition, the core glucose ring was expected to interact with the binding site residues similar to the ligand glucose molecule. However, the docking results showed that some molecules interact

with the glucose binding site through their glycone ring and some through the aglyconic moieties. The interactions of set B molecules (with higher docking score) demonstrated the aglyconic mode of interaction into the glucose binding site similar to that of the standard gliflozins (Figure 10 and 12). The interactions of the aglyconic moieties in the glucose binding site may not be as stable as the glycones. Such binding conformations, therefore, might reduce the stability of inhibitor–protein complex and further promote renal dysfunction and urinogenital side effects in T2DM patients. However, set A molecules (with lower docking score) showed the expected glyconic mode of interaction (Figure 11). Such binding conformations, on the other hand, can be expected to enhance the complex stability and promote SGLT2 inhibition with further reduced adverse effects.

The structural design of the old SGLT2 inhibitors was mostly based on the structure of phlorizin. However, the reduced bio-availability of those O-glucosides led to their discontinuation. The second stage of SGLT2i began with the introduction of dapagliflozin in the market. Structurally, all the available SGLT2 inhibitors along with those under development are considered to be derivatives of dapagliflozin. Modifications in aglycone and glucose moiety of dapagliflozin are the structural basis for designing the SGLT2i (Cai *et al.*, 2015). Although dapagliflozin failed to show any docking result in our study, yet the aforementioned sets of molecules are also structural derivatives of the same drug.

Molecular docking has revealed crucial features relevant to drug interactions. However, a comparative molecular dynamic analysis of A1 and B1 molecules and their organic synthesis can probably justify the above proposed concepts of glyconic and aglyconic interaction.

Chapter 06

REFERENCES

Abdul-Ghani, M. A., Norton, L., and DeFronzo, R. A. (2015). Renal sodium-glucose cotransporter inhibition in the management of type 2 diabetes mellitus. *American Journal of Physiology - Renal Physiology* **309**(11): F889-F900.

Association, A. D. (2014). Diagnosis and Classification of Diabetes Mellitus. *Diabetes care* **37**(Supplement 1): S81-S90.

Back, S. H., and Kaufman, R. J. (2012). Endoplasmic Reticulum Stress and Type 2 Diabetes. *Annual review of biochemistry* **81**: 767-793.

Basit, A., Riaz, M., and Fawwad, A. (2012). Glimepiride: evidence-based facts, trends, and observations. *Vascular Health and Risk Management* **8**: 463-472.

Bautista, R., Manning, R., Martinez, F., del Carmen Avila-Casado, M., Soto, V., Medina, A., and Escalante, B. (2004). Angiotensin II-dependent increased expression of Na⁺-glucose cotransporter in hypertension. *American Journal of Physiology - Renal Physiology* **286**(1): F127-F133.

Bureau, B. B. B. (2016). Mar, 23. homepage. <www.business-standard.com/content/b2b-pharma/sun-pharma-partners-astrazeneca-to-sell-diabetes-drug-dapagliflozin-in-india-116032300628_1.html> Accessed 2017 Nov, 19.

Cai, W., Jiang, L., Xie, Y., Liu, Y., Liu, W., and Zhao, G. (2015). Design of SGLT2 inhibitors for the treatment of type 2 diabetes: a history driven by biology to chemistry. *Medicinal Chemistry* **11**(4): 317-328.

Cernea, S., and Dobreanu, M. (2013). Diabetes and beta cell function: from mechanisms to evaluation and clinical implications. *Biochemia medica: Biochemia medica* **23**(3): 266-280.

Chao, E. C., and Henry, R. R. (2010). SGLT2 inhibition — a novel strategy for diabetes treatment. *Nature Reviews Drug Discovery* **9**: 551.

Chawla, A., Chawla, R., and Jaggi, S. (2016). Microvasular and macrovascular complications in diabetes mellitus: Distinct or continuum? *Indian journal of endocrinology and metabolism* **20**(4): 546-551.

Collaboration, N. R. F. (2016). Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4· 4 million participants. *The Lancet* **387**(10027): 1513-1530.

de Leeuw, A. E., and de Boer, R. A. (2016). Sodium– glucose cotransporter 2 inhibition: cardioprotection by treating diabetes—a translational viewpoint explaining its potential salutary effects. *European Heart Journal - Cardiovascular Pharmacotherapy* **2**(4): 244-255.

Deacon, C. F., Johnsen, A. H., and Holst, J. J. (1995). Degradation of glucagon-like peptide-1 by human plasma in vitro yields an N-terminally truncated peptide that is a major endogenous metabolite in vivo. *The Journal of Clinical Endocrinology & Metabolism* **80**(3): 952-957.

DeFronzo, R. A., E, F., P, Z., and M, A. K. G. M. (2015). *International Textbook of Diabetes Mellitus*. Wiley Blackwell.

Deutschländer, M., Lall, N., and Van de Venter, M. (2010). *Isolation and identification of a novel anti-diabetic compound from Euclea undulata Thunb.* (76).

Drugs.com. (2018). homepage. <www.drugs.com/sfx/insulin-side-effects.html> Accessed 2018 Jan, 21.

Faham, S., Watanabe, A., Besserer, G. M., Cascio, D., Specht, A., Hirayama, B. A., Wright, E. M., and Abramson, J. (2008). The Crystal Structure of a Sodium Galactose Transporter Reveals Mechanistic Insights into Na⁺/Sugar Symport. *Science* **321**(5890): 810-814.

Fiorentino, V. T., Prioletta, A., Zuo, P., and Folli, F. (2013). Hyperglycemia-induced oxidative stress and its role in diabetes mellitus related cardiovascular diseases. *Current pharmaceutical design* **19**(32): 5695-5703.

Foretz, M., Guigas, B., Bertrand, L., Pollak, M., and Viollet, B. (2014). Metformin: From Mechanisms of Action to Therapies. *Cell Metabolism* **20**(6): 953-966.

Fowler, M. J. (2011). Microvascular and macrovascular complications of diabetes. *Clinical Diabetes* **29**(3): 116-122.

Gallo, L. A., Wright, E. M., and Vallon, V. (2015). Probing SGLT2 as a therapeutic target for diabetes: basic physiology and consequences. *Diabetes and Vascular Disease Research* **12**(2): 78-89.

Gambineri, A., Patton, L., Altieri, P., Pagotto, U., Pizzi, C., Manzoli, L., and Pasquali, R. (2012). Polycystic Ovary Syndrome Is a Risk Factor for Type 2 Diabetes: Results From a Long-Term Prospective Study. *Diabetes* **61**(9): 2369-2374.

Ghosh, S., and Collier, A. (2012). *Churchill's Pocketbook of Diabetes E-Book*. Elsevier Health Sciences.

Gunasekaran, U., and Gannon, M. (2011). Type 2 Diabetes and the Aging Pancreatic Beta Cell. *Aging (Albany NY)* **3**(6): 565-575.

Hansen, K. B., Vilsbøll, T., and Knop, F. K. (2010). Incretin mimetics: a novel therapeutic option for patients with type 2 diabetes – a review. *Diabetes, metabolic syndrome and obesity : targets and therapy* **3**: 155-163.

Harada, N., and Inagaki, N. (2012). Role of sodium-glucose transporters in glucose uptake of the intestine and kidney. *Journal of Diabetes Investigation* **3**(4): 352-353.

Hediger, M. A., Romero, M. F., Peng, J.-B., Rolfs, A., Takanaga, H., and Bruford, E. A. (2004). The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteins. *Pflügers Archiv* **447**(5): 465-468.

Holst, J. J. (2007). The Physiology of Glucagon-like Peptide 1. *Physiological Reviews* **87**(4): 1409-1439.

Jacobson, R., Mignemi, N., Rose, K., O ˛ Rear, L., Sarilla, S., Hamm, H. E., Barnett, J. V., Verhamme, I. M., and Schoenecker, J. (2014). The hyperglycemic byproduct methylglyoxal impairs anticoagulant activity through covalent adduction of antithrombin III. *Thrombosis research* **134**(6): 1350-1357.

Katsuno, K., Fujimori, Y., Takemura, Y., Hiratochi, M., Itoh, F., Komatsu, Y., Fujikura, H., and Isaji, M. (2007). Sergliflozin, a novel selective inhibitor of low-affinity sodium glucose cotransporter (SGLT2), validates the critical role of SGLT2 in renal glucose reabsorption and modulates plasma glucose level. *Journal of Pharmacology and Experimental Therapeutics* **320**(1): 323-330.

Kaveeshwar, S. A., and Cornwall, J. (2014). The current state of diabetes mellitus in India. *The Australasian Medical Journal* **7**(1): 45-48.

Kohlroser, J., Mathai, J., Reichheld, J., Banner, B. F., and Bonkovsky, H. L. (2000). Hepatotoxicity due to troglitazone: report of two cases and review of adverse events reported to the United States Food and Drug Administration. *The American Journal of Gastroenterology* **95**(1): 272-276.

Laboratories, I. (2016). homepage. <http://pre-diabetes.insulitelabs.com/gestational_diabetes.php> Accessed 2017 Nov, 19.

López, G. P., Albarrán, O. G., and Megías, M. C. (2010). Type 2 sodium-glucose cotransporter (SGLT2) inhibitors: from familial renal glucosuria to the treatment of type 2 diabetes mellitus. *Nefrologia* **30**(6).

Luna, B., and Feinglos, M. N. (2001). Oral agents in the management of type 2 diabetes mellitus. *American family physician* **63**(9): 1747-1756.

Madaan, T., Akhtar, M., and Najmi, A. K. (2016). Sodium glucose CoTransporter 2 (SGLT2) inhibitors: Current status and future perspective. *European Journal of Pharmaceutical Sciences* **93**: 244-252.

Magen, D., Sprecher, E., Zelikovic, I., and Skorecki, K. (2005). A novel missense mutation in SLC5A2 encoding SGLT2 underlies autosomal-recessive renal glucosuria and aminoaciduria. *Kidney international* **67**(1): 34-41.

Mathers, C. D., and Loncar, D. (2006). Projections of global mortality and burden of disease from 2002 to 2030. *PLoS medicine* **3**(11): e442.

Nagata, T., Fukuzawa, T., Takeda, M., Fukazawa, M., Mori, T., Nihei, T., Honda, K., Suzuki, Y., and Kawabe, Y. (2013). Tofogliflozin, a novel sodium– glucose co-transporter 2 inhibitor, improves renal and pancreatic function in db/db mice. *British journal of pharmacology* **170**(3): 519-531.

Organisation, W. H. (2018). homepage. <http://www.who.int/diabetes/action_online/basics/en/index3.html> Accessed 2017 Nov, 19.

Organization, W. H. (2016). *Global report on diabetes*. World Health Organization.

Otero, Y. F., Stafford, J. M., and McGuinness, O. P. (2014). Pathway-selective Insulin Resistance and Metabolic Disease: The Importance of Nutrient Flux. *The Journal of Biological Chemistry* **289**(30): 20462-20469.

Pathak, R., and Bridgeman, M. B. (2010). Dipeptidyl Peptidase-4 (DPP-4) Inhibitors In the Management of Diabetes. *Pharmacy and Therapeutics* **35**(9): 509-513.

Patnaik, P. K., Jain, K. K., Chandra, P., Pathak, J., Raman, K., and Shah, A. (2016). Diabetes in India: Measuring the dynamics of a public health catastrophe. *Journal of Social Health and Diabetes* **4**(2): 77.

Perland, E., and Fredriksson, R. (2017). Classification systems of secondary active transporters. *Trends in pharmacological sciences* **38**(3): 305-315.

Prabhakaran, D., Jeemon, P., and Roy, A. (2016). Cardiovascular diseases in India. *Circulation* **133**(16): 1605-1620.

Rahmoune, H., Thompson, P. W., Ward, J. M., Smith, C. D., Hong, G., and Brown, J. (2005). Glucose Transporters in Human Renal Proximal Tubular Cells Isolated From the Urine of Patients With Non– Insulin-Dependent Diabetes. *Diabetes* **54**(12): 3427-3434.

Rizos, C. V., Elisaf, M., Mikhailidis, D. P., and Liberopoulos, E. N. (2009). How safe is the use of thiazolidinediones in clinical practice? *Expert Opinion on Drug Safety* **8**(1): 15-32.

Ryder, R., Thong, K., Cull, M., Mills, A., Walton, C., and Winocour, P. (2010). The Association of British Clinical Diabetologists (ABCD) nationwide exenatide audit. *Practical Diabetes* **27**(8): 352.

Sandu, M.-M., Protasiewicz, D. C., Firănescu, A. G., Lăcătuşu, E. C., Bîcu, M. L., and Moţa, M. (2016). Data regarding the prevalence and incidence of diabetes mellitus and prediabetes. *Romanian Journal of Diabetes Nutrition and Metabolic Diseases* **23**(1): 95-103.

Sena, C. M., Pereira, A. M., and Seiça, R. (2013). Endothelial dysfunction — A major mediator of diabetic vascular disease. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **1832**(12): 2216-2231.

Soccio, Raymond E., Chen, Eric R., and Lazar, Mitchell A. (2014). Thiazolidinediones and the Promise of Insulin Sensitization in Type 2 Diabetes. *Cell Metabolism* **20**(4): 573-591.

Sola, D., Rossi, L., Schianca, G. P. C., Maffioli, P., Bigliocca, M., Mella, R., Corlianò, F., Fra, G. P., Bartoli, E., and Derosa, G. (2015). Sulfonylureas and their use in clinical practice. *Archives of Medical Science : AMS* **11**(4): 840-848.

Spanakis, E. K., and Golden, S. H. (2013). Race/Ethnic Difference in Diabetes and Diabetic Complications. *Current diabetes reports* **13**(6): 10.1007/s11892-11013-10421-11899.

Thulé, P. M., and Umpierrez, G. (2014). Sulfonylureas: A New Look at Old Therapy. *Current diabetes reports* **14**(4): 473.

Trujillo, J. M., Nuffer, W., and Ellis, S. L. (2015). GLP-1 receptor agonists: a review of head-to-head clinical studies. *Therapeutic Advances in Endocrinology and Metabolism* **6**(1): 19-28.

Vergès, B. (2015). Pathophysiology of diabetic dyslipidaemia: where are we? *Diabetologia* **58**(5): 886-899.

Wright, E. M., Loo, D. D. F., and Hirayama, B. A. (2011). Biology of Human Sodium Glucose Transporters. *Physiological Reviews* **91**(2): 733-794.

Zhang, H., Ackermann, A. M., Gusarova, G. A., Lowe, D., Feng, X., Kopsombut, U. G., Costa, R. H., and Gannon, M. (2006). The FoxM1 Transcription Factor Is Required to Maintain Pancreatic β-Cell Mass. *Molecular Endocrinology* **20**(8): 1853-1866.

Zimdahl, H., Haupt, A., Brendel, M., Bour, L., Machicao, F., Salsali, A., Broedl, U. C., Woerle, H.-J., Häring, H.-U., and Staiger, H. (2017). Influence of common polymorphisms in the SLC5A2 gene on metabolic traits in subjects at increased risk of diabetes and on response to empagliflozin treatment in patients with diabetes. *Pharmacogenetics and Genomics* **27**(4): 135-142.

Zou, H., Zhou, B., and Xu, G. (2017). SGLT2 inhibitors: a novel choice for the combination therapy in diabetic kidney disease. *Cardiovascular diabetology* **16**(1): 65.

Student approval form

Name of the author	Kamaljyoti Chakravorty
Department	Department of Human Genetics and Molecular Medicine
Degree	M.Sc. life sciences with specialization in Human Genetics
University	Central University of Punjab, Bathinda
Guide	Dr. Anjana Munshi
Project title	*In silico* design of compounds as SGLT2 inhibitors
Year of award	2018

Agreement

1. I hereby certify that, if appropriate, I have obtained and attached here to a written permission / statement from the owner (s) of each third party copyrighted matter to be included in my dissertation/thesis, allowing distribution as specified below.

2. I hereby grant to the university and its agents the non-exclusive license to archive and make accessible, under the conditions specified below, my dissertation/thesis, in whole or in part in all forms of media, now or hereafter known. I retain all other ownership rights to the copyright of the dissertation/thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis, dissertation, or project report.

Conditions:

1. Release the entire work for access worldwide	Not applicable
2. Release the entire work for Central University of Punjab only for: 1 year 2 years 3 years and after this time release the work for access	After 03 years

worldwide.	
3. Release the entire work for Central University of Punjab only, while at the same time releasing the following parts of the work (e.g. because other parts relate to publications) for worldwide access: a) Bibliographic details and synopsis only b) Bibliographic details, synopsis and the following chapters only c) Preview / Table of contents / page only	Not applicable
4. View only (No downloads) (world-wide)	Not applicable

Signature of the candidate **Signature and seal of the supervisor**

Place: Bathinda

Date: May,2018

URKUND

Urkund Analysis Result

Analysed Document:	Kamal's thesis copy (1).docx (D38770054)
Submitted:	5/16/2018 11:45:00 AM
Submitted By:	nonakalra@gmail.com
Significance:	1 %

Sources included in the report:

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3480706/

http://shodhganga.inflibnet.ac.in/bitstream/10603/42806/11/11_chapter%206.pdf

https://isp.ncifcrf.gov/files/isp/uploads/2010/07/gli55_user_manual.pdf

Instances where selected sources appear:

3