Alkyl proline and oxytocin derivatives towards development of anti-diabetics and design of potential antifolates derived from guanine

by

Harsha Chilukuri
10CC17J26025

A thesis submitted to the
Academy of Scientific & Innovative Research
for the award of the degree of
DOCTOR OF PHILOSOPHY
in
SCIENCE

Under the supervision of
Dr. Moneesha Fernandes

National Chemical Laboratory

CSIR-National Chemical Laboratory, Pune

Academy of Scientific and Innovative Research
AcSIR Headquarters, CSIR-HRDC campus
Sector 19, Kamla Nehru Nagar,
Ghaziabad, U.P. – 201 002, India

May 2021

Certificate

This is to certify that the work incorporated in this Ph.D. thesis entitled, "*Alkyl proline and oxytocin derivatives towards development of anti-diabetics and design of potential antifolates derived from guanine",* submitted by *Harsha Chilukuri* to the Academy of Scientific and Innovative Research (AcSIR) in fulfillment of the requirements for the award of the Degree of *Doctor of Philosophy in Sciences*, embodies original research work carried-out by the student. We, further certify that this work has not been submitted to any other University or Institution in part or full for the award of any degree or diploma. Research material(s) obtained from other source(s) and used in this research work has/have been duly acknowledged in the thesis. Image(s), illustration(s), figure(s), table(s) etc., used in the thesis from other source(s), have also been duly cited and acknowledged.

(Signature of Student)   (Signature of Co-Supervisor)   (Signature of Supervisor)
                                            If-any

   Harsha Chilukuri   Dr. Moneesha Fernandes
       6.5.21   6.5.21

STATEMENTS OF ACADEMIC INTEGRITY

I Harsha Chilukuri, a Ph.D. student of the Academy of Scientific and Innovative Research (AcSIR) with Registration No. 10CC17J26025 hereby undertake that, the thesis entitled "Alkyl proline and oxytocin derivatives towards development of anti-diabetics and design of potential antifolates derived from guanine" has been prepared by me and that the document reports original work carried out by me and is free of any plagiarism in compliance with the UGC Regulations on *"Promotion of Academic Integrity and Prevention of Plagiarism in Higher Educational Institutions (2018)"* and the CSIR Guidelines for *"Ethics in Research and in Governance (2020)"*.

Signature of the Student
Date   : 6.5.21
Place   : Pune

It is hereby certified that the work done by the student, under my/our supervision, is plagiarism-free in accordance with the UGC Regulations on *"Promotion of Academic Integrity and Prevention of Plagiarism in Higher Educational Institutions (2018)"* and the CSIR Guidelines for *"Ethics in Research and in Governance (2020)"*.

Signature of the Co-supervisor (if any)
Name :
Date   :
Place :

Signature of the Supervisor
Name : Dr. Moneesha Fernandes
Date   : 6.5.21
Place : Pune

To

Family

Acknowledgements

If you wish to see the rainbow, you have to put up with the rain!

This quote aptly summarizes my thankfulness on completion of my doctoral thesis. And unquestionably my research supervisor Dr. Moneesha Fernandes is the first person to be acknowledged since she provided me with constant support, encouragement and patience to put up with the rain of scientific and unscientific problems I faced during the course of the study. I am also very thankful for the freedom I was given in my projects and publications which enabled me to acquire different scientific capabilities and perceive science from a broad perspective.

I would like to thank Dr. Vaijayanti Kumar for her constant guidance throughout my Ph.D. Her enthusiasm towards work is very inspirational. Her inputs and suggestions in the lab have made my thesis truly meaningful.

I would also like to thank Dr. Mahesh J Kulkarni who provided me with unconditional support right from the beginning of my joining NCL as a project assistant. I could learn biology and cell culture and can independently carry out experiments because of him and his students. Although I never actively worked in Dr Kulkarni's lab, he and his students always considered me as one of their own.

I would like to thank my Doctoral Advisory Committee members Dr. Narendra Kadoo, Dr. G. J. Sanjayan and Dr. Udaya Kiran Marelli for their timely evaluation of my progress and for providing me with useful suggestions during the course of my work.

I am very thankful for the support provided by Dr. Santhakumari and her students for mass spectrometric analysis. I am also thankful to Mr. Gati Nayak, whose timely and prompt evaluation of my thesis for plagiarism was done without any difficulty. Student Academic Office, specially Ms. Vaishali and Ms. Komal deserve a special appreciation for their work and helping me and many students in the tedious documentation procedure for successful completion of their degree. I would also like to thank the Director and Organic Chemistry Division office for providing me with constant support. My Ph.D. degree would not have been possible without the financial support of the DST-Women Scientist Scheme for which I am truly grateful.

Acknowledgements

I would like to thank Rajeshwari for patiently carrying out and teaching me all the cell culture experiments. I would also like to thank Linthoi and Dr. Karthikeyan for all the collaborative work they provided. My thesis would have been incomplete without the help of my collaborators.

I am very much thankful to my seniors and lab mates who guided, helped and patiently taught me how to work in a lab. My deepest gratitude to Anita Mam, Madhuri, Namrata, Seema, Venu, Kiran, Manoj, Anjan, Tanaya, Amit, Govind, Ragini, Manisha, Atish and Komal for keeping the lab atmosphere fun and cheerful all the time. I would also like to thank Mr. Bhumkar and Mr. Gurav for helping out in the lab.

I would specially like to thank Atish and Manisha for all the support and encouragement they provided and keeping the lab a pleasant place to work during the difficult times of Covid 19 lockdown.

I would like to express my deepest and sincerest gratitude to my parents for always being a source of encouragement and patiently bearing my tantrums. I would also like to thank my brother for the support he provided. My smooth journey of Ph.D. would not have been possible without their support.

I would specially like to thank my mother in-law and father in-law for understanding the importance of my work and thus having unconditional patience and support. I would also like to thank all the family members from my in-laws side for their constant support and encouragement.

This thesis would be incomplete without mentioning the support and constant encouragement provided by my husband, Venu. Not only has he inspired me to work harder, but he has also taught me the importance of perseverance. His patience and support has pushed me to complete my thesis and one day I would like to become a researcher like him.

Lastly, I would like to thank my Stars for always looking out for me and making my life more meaningful.

-Harsha

<u>Contents</u>

Chapter 1

Introduction to Antidiabetics, Antifolates and Peptides

SECTION A

SECTION B

SECTION C

Chapter 2

Oxytocin and its analogues mediated glucose uptake in Chinese Hamster Ovary cells

Chapter 3

N-(3-Aminoalkyl)proline derivatives with potent antigycation activity

Chapter 4

Design, synthesis and evaluation of antifolates derived from guanine

Appendix

Units

°C	Degree Celsius
mg	Milligram
g	Gram
h	Hour
Hz	Hertz
µg	Microgram
mL	Millilitre
min	Minutes
MHz	Megahertz
mmol	Millimole
ppm	Parts per million
M	Molar
µM	Micromolar

Chemical Notations

Boc	Di-tert butyl carbonate
Fmoc	Fluorenylmethyloxycarbonyl
DCM	Dichloromethane
DMF	*N, N'*- Dimethylformamide
DMAP	*N,N'*- Dimethylaminopyridine
DIPEA	*N, N*- Diisopropylethylamine
MeOH	Methanol
HOBt	Hydroxybenzotriazole
MBHA	4-Methylbenzhydrylamine hydrochloride

PyBOP	Benzotriazol-1-yl oxytripyrrolidinophosphonium hexafluorophosphate
HATU	Hexafluorophosphate Azabenzotriazole Tetramethyl Uronium

Single and three letter amino acid codes

G/ Gly	Glycine
L/Leu	Leucine
P/ Pro	Proline
K/Lys	Lysine
C/Cys	Cysteine
N/Asn	Asparagine
Q/Gln	Glutamine
Y/Tyr	Tyrosine
E/Glu	Glutamic acid
I/Ile	Isoleucine

Other Notations

δ	Chemical shift
J	Coupling constant
DEPT	Distortionless Enhancement by Polarization Transfer
IR	Infra Red
m/z	Mass-to-charge ratio
NMR	Nuclear Magnetic Resonance
rt	Room temperature
TLC	Thin Layer Chromatography
CHO	Chinese Hamster Ovary

- All the reagents were purchased from Sigma-Aldrich and used without further purification.

- DMF was dried over P_2O_5 and stored by adding 4 Å molecular sieves.

- TEA was dried over KOH and stored on KOH.

- Reactions were monitored by TLC. TLCs were run in either Petroleum ether with appropriate quantity of EtOAc or DCM with an appropriate quantity of MeOH

- TLC plates were visualized with UV light and/or by Ninhydrin solution and heating.

- Usual reaction work up involved sequential washing of the organic extract with water and brine followed by drying over anhydrous Na_2SO_4 and evaporation of the solvent under vacuum.

- Column chromatographic separations were performed using silica gel 60-120 mesh (Merck) or 200- 400 mesh (Merck) and using the solvent systems EtOAc/Petroleum ether or MeOH/DCM.

- TLC was run using pre-coated silica gel GF254 sheets (Merck 5554).

- ^{1}H and ^{13}C NMR spectra were obtained using Bruker AC-200, AC-400 and AC-500 NMR spectrometers.

- ^{1}H NMR data are reported in the order of chemical shift, multiplicity (s, singlet; d, doublet; t, triplet; br, broad; br s, broad singlet; m, multiplet and/ or multiple resonance), number of protons.

- Mass spectra were recorded on a Q Exactive Hybrid Quadrupole Orbitrap Mass spectrometer (Thermo Fisher Scientific

- RP-HPLC on C-18 column using Waters system

- MALDI-TOF spectra were recorded on a SCIEX TOF/TOF 5800 system, and the matrix used for analysis was

- CD spectra were recorded on a Jasco J-815 spectropolarimeter equipped with a Peltier-controlled cell holder.

Chapter 1

Introduction to

Antidiabetics, Antifolates

and

Peptides

Section A
Antidiabetics

1.0 Diabetes and Antidiabetics

A global chronic metabolic disease, Diabetes Mellitus is characterised by constant high glucose levels in the blood. [1] Rapidly growing in both the developed and developing world leading to life-threatening complications,[2] this disease is classified as Type 1 and Type 2 diabetes, pancreatic β-cell dysfunction being a common and important characteristic in both the types.[3] Immune mediated destruction of pancreatic β-cells leading to insufficient insulin is seen in Type 1 diabetes, usually diagnosed in children,[4] while Type 2 diabetes, the most common type, seen in adults is due to inadequate insulin secretion and/or insulin resistance.[5] The risk of developing type 1 diabetes is based on many factors of genetic and environmental origin such as increased weight, exposure to some viral infections and increased maternal age at delivery. Whereas factors such as obesity, changes in diet and physical activity, age, insulin resistance, a family history of diabetes and ethnicity contribute to the occurrence of type 2 diabetes.[6,7] Although the occurrence of type 2 diabetes accounts for 90-95% [8] both the classes are chronic conditions that cannot be cured but only managed. [9] A significant burden on healthcare systems globally is seen due to prevalence of diabetes reaching epidemic proportions.[10] Other severe conditions caused by diabetes mellitus over time are nerve damage, blindness, amputation of lower limbs, kidney damage and cardio vascular disease.[11]

Several different mechanistic classes distinguish currently deployed antidiabetic drugs such as insulin secretagogues, insulin sensitizers, insulin mimetics, α-glucosidase inhibitors, and DPP IV inhibitors which can be used either in monotherapy or in combination. In the last years, thousands of compounds have emerged and there is extensive literature on them.[12]

However some unwanted side effects have also been witnessed with the use of these drugs, hence the search for new antidiabetic drugs is of considerable interest to researchers worldwide.[13] Although progresses have been made, the pursuit for the "perfect" antidiabetic drug still continues.[12] Lately, people have also turned to the use of herbal medicine and natural products because of their availability, low cost, and low side effects.[14]

1.1 Classification of Diabetes

WHO classifies diabetes into two major types: Type 1 diabetes (T1DM) and type 2 diabetes (T2DM). Age, onset and degree of loss of β cell function play a major role in this classification. But the underlying common characteristic is the dysfunction or destruction of pancreatic β-cells in both diabetes classes.

Type 1 diabetes is an autoimmune disorder caused due to β-cell destruction resulting in absolute insulin deficiency. Onset of this type of diabetes is childhood and early adulthood.

Type 2 diabetes is the most common type and accounts for between 90% and 95% of diabetes. Varying degrees of β-cell dysfunction and insulin resistance frequently leads to obesity. Insulin treatment is generally not required for survival, but may be used to decrease the blood sugar level to avert chronic complications.[15]

1.2 Insulin, its structure and its role in diabetes

Insulin is synthesised in the β cells of the pancreas [16] and still continues to be the cornerstone of therapy for diabetes mellitus.[17] Maintaining glucose level in the blood[18] and energy storage[19] are the main functions of this peptide, leading to its high demand owing to dramatic increase in the number of diabetic patients worldwide. The year 1922 witnessed a major breakthrough in medicine and therapy for diabetes when insulin was discovered[17] which led to the Noble Prize in Physiology or Medicine being awarded to Banting and Macleod in 1923.[20] Following the discovery, another breakthrough in 1955 was witnessed where the primary structure of insulin, the two chains and the disulfide bonds was successfully determined by Sanger and co-workers leading to a Nobel Prize in 1958.[21]
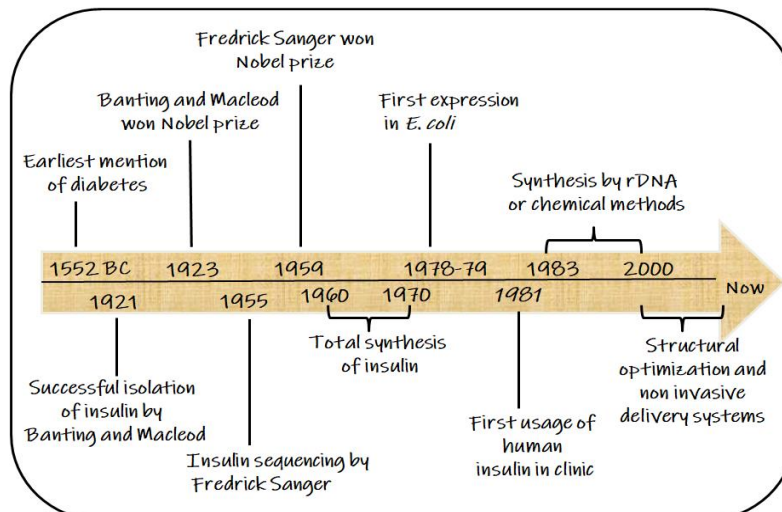
Figure 1: Timeline of insulin and its development

Since the discovery of this peptide hormone, its structure and stability has been a subject of numerous studies.[22]Structural elucidation revealed Insulin consists of 2 chains, chain A with 21 amino acids and a chain B with 30 amino acids linked by two disulfide bonds. In addition, chain A contains an intra-chain disulfide bond [22] and a terminal helix and is linked to chain B *via* two disulfide bonds[16] as depicted in figure 2. This structure is highly conserved among vertebrates, which is evidenced by its high degree of homology. Representative structural

features of large proteins such as a well-ordered hydrophobic core and canonical elements of secondary structure are exhibited by insulin despite its small size[19]
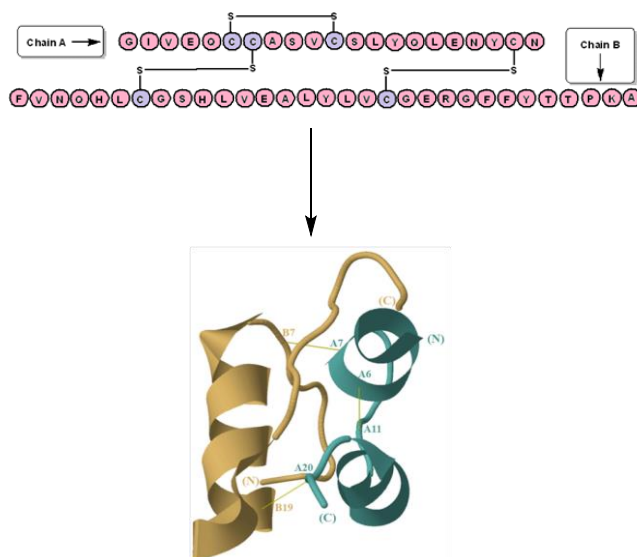
Figure 2: Two and three dimensional structures of insulin[21]

1.2.1 Mechanism of Insulin action

Secretion of insulin from the β cells results due to increased levels of glucose.[16] Insulin binds to insulin receptor present on the surface of the cell[21] giving rise to a biological cascade of events at molecular level. [23] Since glucose is a hydrophilic molecule, transporters such a glucose transporter proteins (GLUT) are required for glucose to cross the cell membrane. [24]The movement of these transporter proteins is stimulated by insulin causing the transport from intracellular pool to the plasma membrane occurring within seconds after the stimulus.[25]
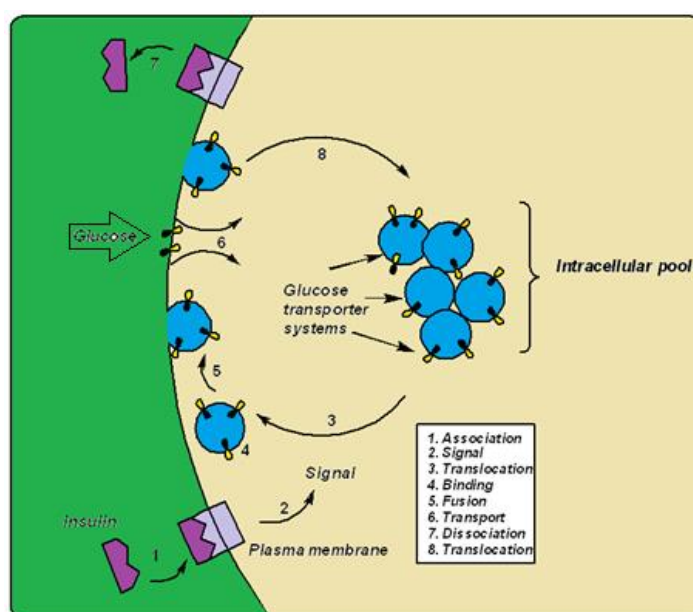
Figure 3: Molecular mechanism of insulin action[25]

Despite in depth study conducted on the relationship between the structure of insulin and its function, synthesis of new analogues of this peptide remains challenging.[26] These new analogues are designed to mimic insulin with improved properties such as stability, rate of release and duration of effect[27] as shown in figure 4.
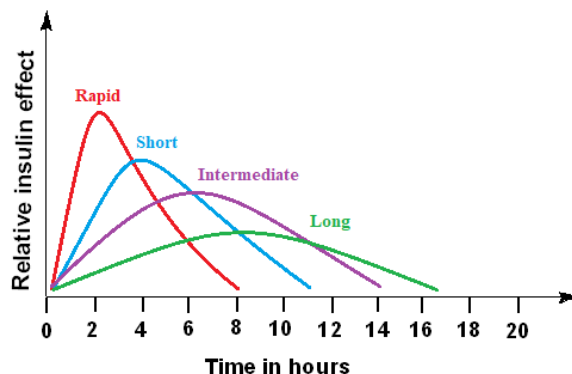
Figure 4: Release rates of different insulin preparations

1.3 Current small molecules and peptides as antidiabetic drugs

Continuous demand for improved diabetic therapies that provide precise glycemic control and maintainace and/ or improvement of pancreatic endocrine functions is evident although several tradiotional hypoglycaemic agents have been widely used for glycemic control.[28] The majorly used oral antidiabetic medications include dipeptidyl peptidase 4 (DPP-4) inhibitors, thiazolidinediones (TZD), sulfonylureas, biguanides, α-glucosidase inhibitors, sodium-glucose cotransporter (SGLT2) inhibitors and incretin mimetics.[29,30] A list of important marketed small molecule antidiabetic agents such as metformin, rosiglitazone, glimepiride, acarbose, sitagliptin and canagliflozin are summarised in figure 5.
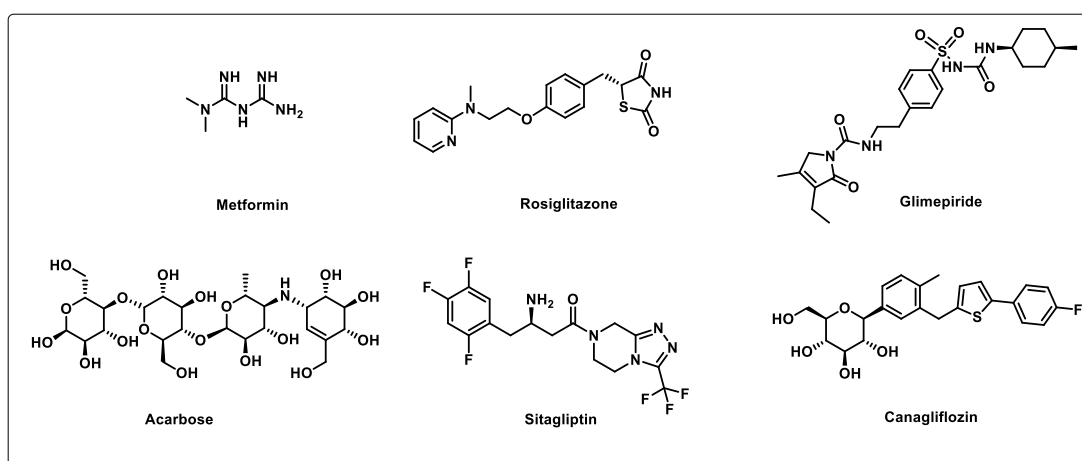
Figure 5: Some of the important marketed antidiabetic drugs.

Table 1: Current oral antidiabetic medications and targeted organs[29,30]

Therapeutic Class	Drug	Targeted organ(s)	Mechanism of action
Biguanides	Metformin	liver, intestine, pancreas	Insulin sensitizer Numerous effects on inhibition of hepatic glucose production
Thiazolidinediones	Rosiglitazone	liver, adipose tissue, muscle	Insulin sensitizer
Sulfonylureas	Glimepiride	pancreas	Insulin secretion
α-glucosidase inhibitors	Acarbose	pancreas, small intestine	act as competitive inhibitors of enzymes needed to digest carbohydrates
DPP-IV inhibitors	Sitagliptine	intestines	Inhibition of degradation of GLP
SGLT2 inhibitors	Canagliflozin	kidneys	Glucosuria due to blocking (90%) of glucose reabsorption in renal PCT; insulinindependent mechanism of action
Incretin mimetics	Liraglutide, Exenatide	pancreas	Activate GLP1 receptor Increased insulin secretion, decreased glucagon, delayed gastric emptying, increased satiety

β cell preservation, improvement of glycemia ultimately leading to reduction or halting the progression of diabetes with excellent safety profile and minimum side effects would be an ideal therapeutic intervention for diabetes. No such therapy existed until a new therapeutic

class known as incretin mimetics or incretin enhancers was first identified. Glucose dependent insulinotropic polypeptide (GIP) and glucagon like peptide (GLP-1) are the peptides responsible for the incretin effect. Released from the gastrointestinal tract in response to food intake, incretin hormones enhance glucose dependent insulin secretion from pancreas. These peptides govern glucose homeostasis by several mechanisms like increasing glucose stimulated secretion of insulin, reduce glucose production in the liver and decrease secretion of glucagon, delay gastric emptying leading to reduction of body weight.[31] About 70% of insulin release is caused due to incretin effects and any disturbances or defect in this system is thought to be the main contributor for type 2 diabetes mellitus.[32] Thus this system has become an important and attractive target for T2DM. GLP-1 receptor agonists and DPP-4 inhibitors are the two drug classes that cause the above effect.[29,30] Many incretin mimetics are currently approved for the treatment of diabetes targeting the GLP-1 such as albiglutide, dulaglutide, exenatide, liraglutide and lixisenatide[29] structures as shown in figure 6.
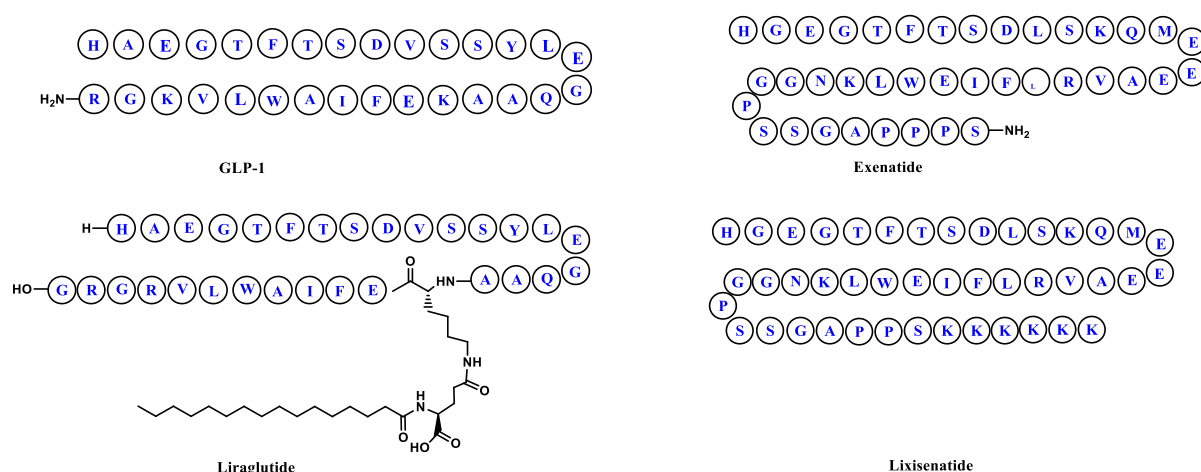
Figure 6: Sequences of GLP-1, Exenatide, Liraglutide and Lixisenatide

1.4 Hyperglycemia and Advanced Glycation End Products(AGEs)

A common outcome in type 1 and type 2 diabetes is the high level of glucose in the blood called as hyperglycemia[33] which is caused due to absolute or relative deficiency of insulin.[5] Prolonged hyperglycaemic conditions give rise to several other complications apart from diabetes and impaired insulin secretion such as heart diseases, nephropathy, neuropathy etc.[34] Such hyperglycaemic conditions also favour the formation of covalent adducts between glucose and plasma proteins[35]/ lipids/ nucleic acids[36] through a non enzymatic process known as glycation reaction further forming undesired Advanced Glycation End products (AGEs).[37] These AGEs that form *via* the Maillard reaction[38] represent one of the most promising areas of

research today. Important work and emphasis on AGEs since the last 20 years has been accomplished although the initial chemistry behind their formation was known since the early 1900's.[39] AGEs are formed as a result of complex and multistep processes wherein the electrophilic carbonyl groups of reactive sugars (glucose) react with the side chains of nucleophilic amino groups of amino acids (lysine or arginine) and form a non-stable Schiff base which further rearranges to Ketoamine (Amadori products). These products can further irreversibly react with amino acid residues of peptides or proteins to form complex protein adducts or cross links. In continuation, numerous AGEs result from dehydration, oxidation, oxidative breakdown and polymerization reactions.[40]
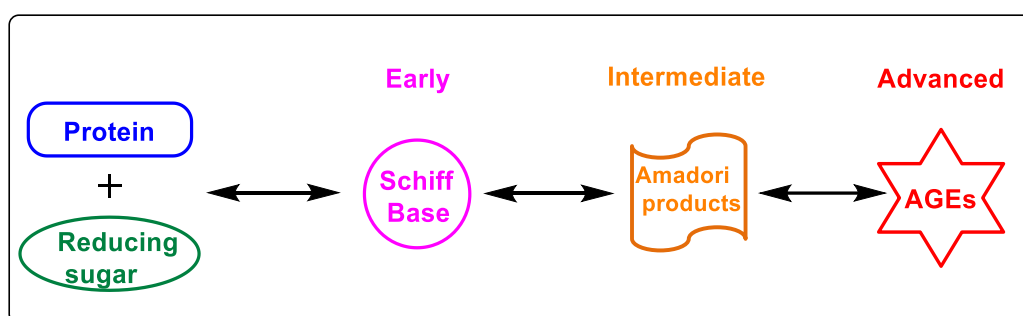
Figure 7: Glycation reaction between reducing sugar and protein/DNA leading to formation of AGEs

1.4.1 Antiglycation

Natural and synthetic chemical compounds have been previously employed for the intervention of Maillard reaction towards removing one of the reactants; either the amino group or the sugar.[41] Early stage inhibition of the glycation reaction is seen with a few synthetic AGE inhibitors which occur due to initial attachment of the sugars to the amino groups thus inhibiting further reaction. Late stage glycation inhibition is seen in most synthetic inhibitors mainly because reactive carbonyls are scavenged and thus halting the formation of intermediate Amadori products.[42] Figure 8 depicts the early and late stage glycation inhibitors.

Thus, an effective strategy to prevent AGEs formation and to manage diabetic complications, antiglycation therapy would prove effective.[43]
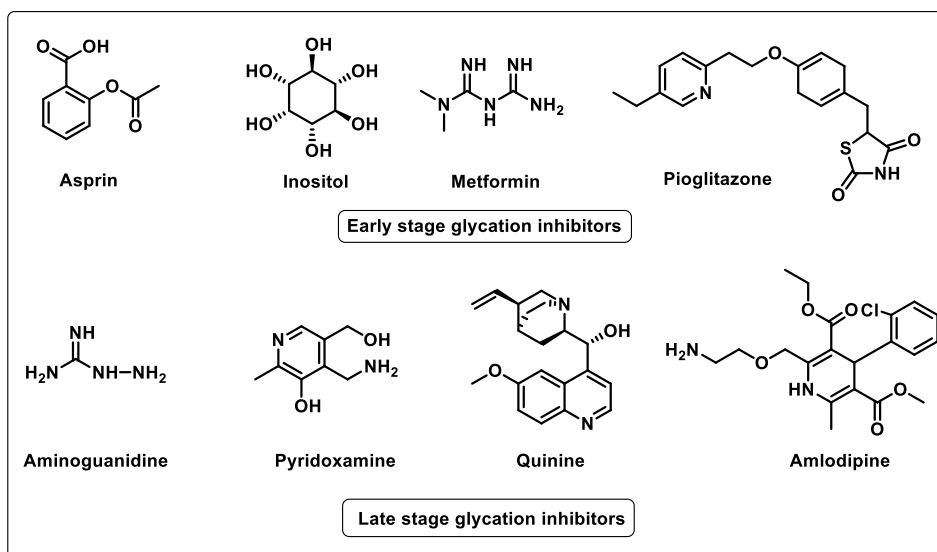
Figure 8: Structures of AGE inhibitors that inhibit glycation at early stage and late stage.[42]

1.4.2 Glucose Uptake

All the cells require glucose as the main source of energy thus playing a central role in various functions metabolic and homeostatic functions.[44] The eukaryotic plasma membrane is a lipid bilayer hence glucose is transported across the membrane by carrier proteins and glucose transporters.[45] GLUTs is a family of membrane proteins which is mainly responsible for the transport of glucose into tissues or cells[46] failure of which causes impaired glucose homeostasis leading to hyperglycemia[47,48] and in turn type 2 diabetes and obesity.[47] These proteins are structurally related membrane proteins designated as GLUT1-GLUT-5. Of these, GLUT1, GLUT2, GLUT3, and GLUT5 appear to be localized primarily to the plasma membrane whereas GLUT4 is located near the nucleus and moves to the membrane under the stimulus of insulin[49] in order to decrease the concentration of glucose in the blood.[50] Uptake of glucose by this transporter is mainly seen in tissues such as skeletal, cardiac muscles and adipose tissue.[51] Thus developing safer antidiabetic agents while keeping in mind molecular mechanisms to regulate blood glucose level have become a major research focus.[47]

Section B
Antifolates

1.5 Folates and Antifolates

Understanding the role of folates and its derivatives in humans led to the development and identification of antifolates as therapeutics. Therefore before discussing antifolates, work that led to the discovery of folates should be given due attention. A group of B vitamins also known as "folates" share the same vitamin activity based on the parent structure of folic acid. Fully oxidised and most stable form of the vitamin is generally referred as 'folic acid' which is generally used in supplements and fortified foods.[52] Since humans cannot synthesise folates, high dependence on dietary sources makes up for the daily folate requirement[53] Biosynthesis of critical cellular components such as nucleobases, proteins and co factors are brought about by various folate species.[54] Chemically, pteridine ring, *para*-amino benzoic acid and glutamate moieties comprise folic acid.[55] Specialised membrane transport systems allow folates which are highly hydrophilic anionic molecules to cross the cell membrane.[56] Majorly two different mechanisms govern the entry of natural and synthetic folates into cells of which one involves reduced folate carrier (RFC) having an affinity of micromolar range. The other transport system utilises a high affinity, membrane binding protein/ receptor called the folate receptor having nanomolar range affinity towards folates.[57]
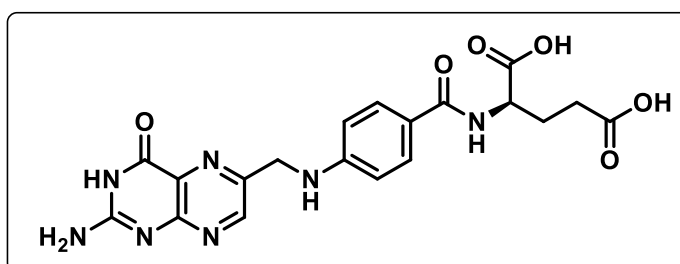
Figure 9: Chemical structure of folic acid

1.6 Genesis of the use of antifolates against human disease

In humans, ingested folic acid is actively transported into cells which is then reduced by the enzyme dihydrofolate reductase (DHFR) first to dihydrofolate and subsequently to tetrahydrofolate thus becoming biologically active.[53] Antifolates were the first class of antimetabolites to enter the clinics, act by the disruption of the metabolic pathways that require one carbon moieties supplied by the B9 folate vitamins due to their structural similarity.[58] The antifolate therapy comprises of three major clinically validated folate targets: dihydrofolate reductase (DHFR) and thymidylate synthase (TS) and serine hydroxymethytransferase (SHMT).[59] Incidentally unlike humans, bacteria, fungi and protozoa

possess an endogenous folate biosynthetic pathway wherein dependency on external folates is not seen.[60]
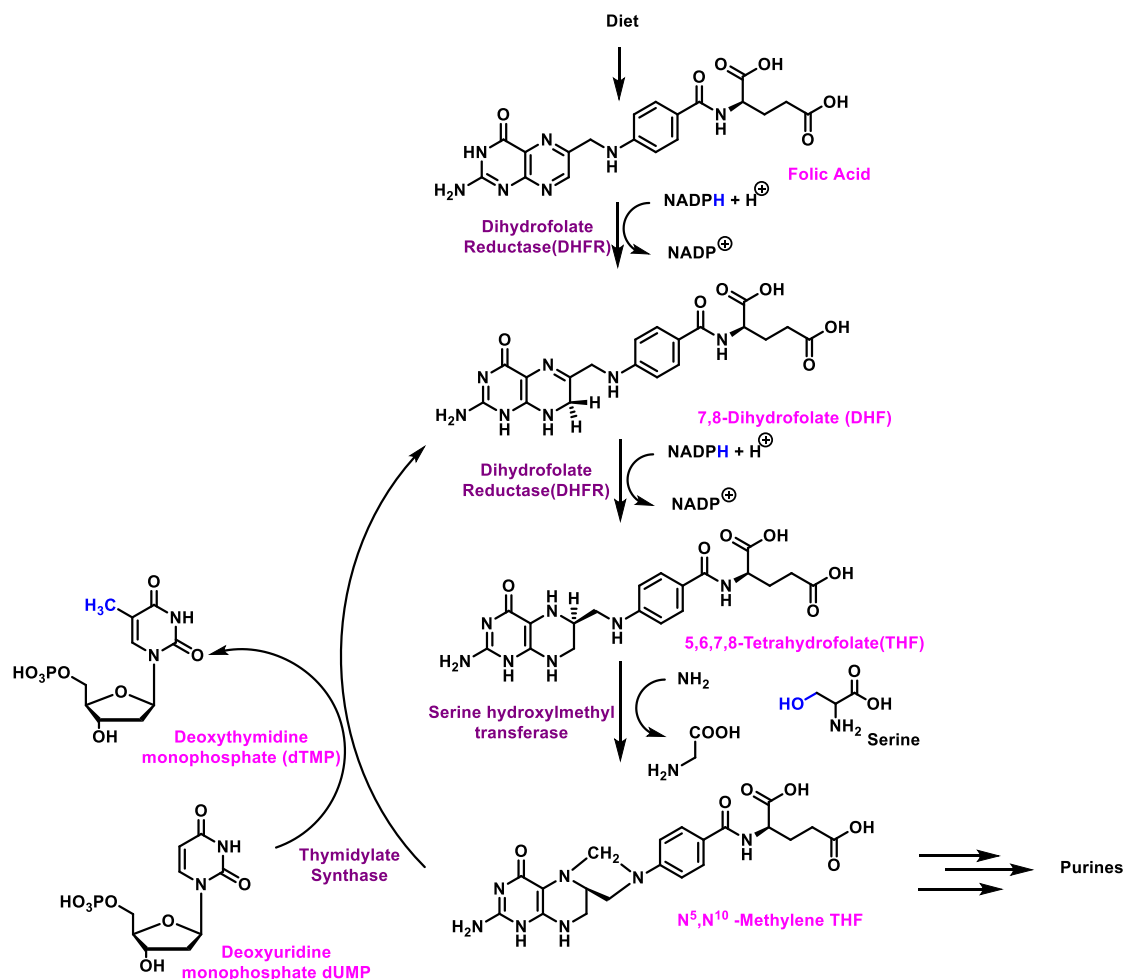
Figure 10: Folate biosynthetic pathway.[61]

1.6.1 Dihydrofolate reductase (DHFR) inhibitors

DHFR is a small protein with molecular weight of 20 kilo Daltons[62] having a role in thymidine (one of the four nucleosides required for DNA replication) biosynthesis. This enzyme causes NADPH catalysed reduction of dihydrofolate to tetrahydrofolate, which is essential for protein and nucleic acid biosynthesis.[63] In the folate metabolism, tetrahydrofolate acts as a critical one-carbon-unit donor. [64] This is by far the most intensively studied enzyme in the folate pathway. [65]Malaria and other protozoal infections, infections caused due to fungus, bacteria and mycobacteria are commonly treated with DHFR inhibitors.[60,66] Over the years, several compounds have been discovered and different drugs have entered the market. Among them, we have to mention pyrimethamine and proguanil as antimalarial drugs trimethoprim, an antibacterial drug commonly used in association with sulfonamides, like sulfamethoxazole and

methotrexate[67] (figure 11) Designing new agents that target crucial enzyme Dihydrofolate reductase (DHFR) is important[68] because it plays a critical role in the folate pathway.[68] Unfortunately, these negatively charged antifolates (at physiological pH) have insufficient intracellular concentrations to show required activity due to limited passive diffusion.[69]

(a) Methotrexate **(b) Pyremethamine** **(c) Proguanil**
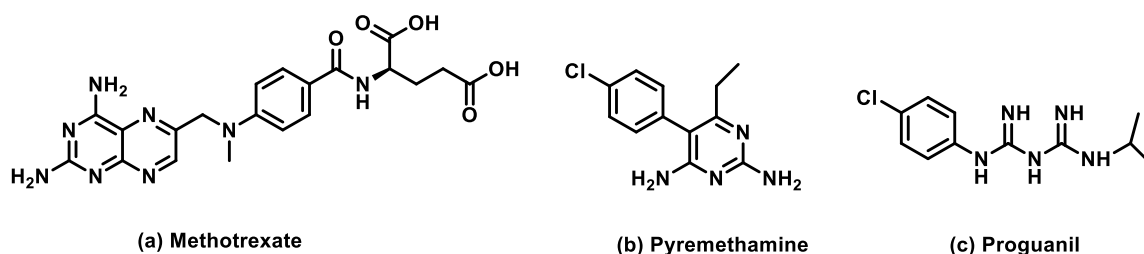
Figure 11: Chemical structures of commonly used approved drugs acting on DHFR

1.6.2 Thymidylate synthase (TS) inhibitors

Thymidylate-mediated pathway was the only known mechanism by which *de novo* thymidine nucleotide biosynthesis occurred.[70] Thymidylate synthase catalyses the reductive methylation of 2′-deoxyuridine-5′-monophosphate (dUMP) to 2′-Deoxythymidine-5′-monophosphate (TMP) [71]Replication process is affected due to a complex mechanism called as thymineless death is caused due to TS inhibition, hence regarded as a target for cancer chemotherapy. This enzyme is also involved in protein regulation pathway and apoptotic processes. Along with being an attractive target for cancer chemotherapy, this protein also acts as a target for antibacterial and antifungal chemotherapy. The structural elucidation by X-ray in combination with computational studies and site-directed mutagenesis revealed multiple binding modes, providing a direction towards reliable drug design. The primary focus in designing drugs against TS has been the small molecules that structurally resemble the substrate, dUMP, or its cofactor.[72,73] The enzyme's activity is a two-stage process. Firstly, a configurational change is introduced upon binding of dUMP to the receptor creating a neighbouring binding site for *N*-5,10-methylene-tetrahydrofolate (CH$_2$THF). Then the folate's one carbon group is delivered to the uridine ring, furnishing TMP and dihydrofolate. TMP is subsequently phosphorylated by a kinase to corresponding diphosphate and triphosphate derivatives of thymidine, TDP and TTP, respectively. [74]

(a) Raltitrexed **(b) 5-Fluorouracil** **(c) Pemetrexed**
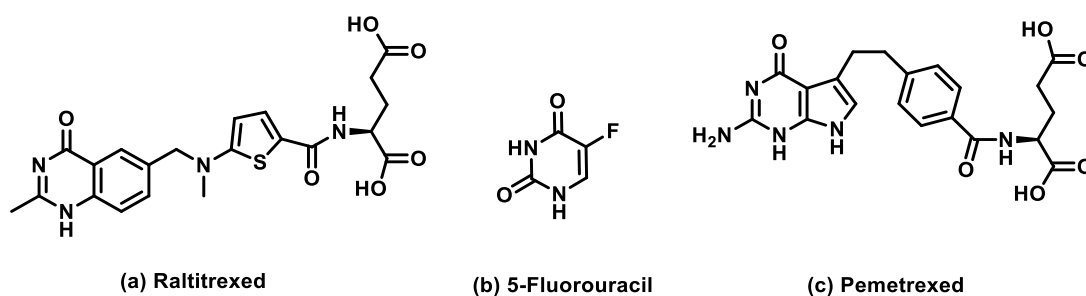
Figure 12: Chemical structures of approved drugs acting on TS[75]

1.6.3 Serine hydroxymethyltransferse (SHMT) inhibitors

The metabolic role of *L*-serine is not only being a building block for conventional protein synthesis [76] but also a principal one-carbon source for DNA biosynthesis.[77] The enzyme serine hydroxymethyltransferase (SHMT) fragments serine into glycine and a tetrahydrofolate-bound one-carbon unit.[78] Two isoforms of this enzyme are present in humans *viz* cytosolic SHMT and mitochondrial SHMT that share a ~66% amino acid sequence identity and possess very similar catalytic properties.[79] It has been proposed that the cytosolic SHMT catalyses the reaction of 5,10-methylene-THF and glycine to THF and serine, whereas mitochondrial SHMT might catalyze the reverse reaction.[80] It being a ubiquitous enzyme, the main function is to be a major source of one carbon groups required for purine, methionine and thymidylate biosynthesis.[81] Although SHMT has biological and medical importance, drugs that target this enzyme have not been developed yet. This is because this enzyme catalyses serine to glycine, having limited space thus the substrates that can be accepted into the active site of SHMT are extremely limited[82] although several studies targeting the protein have been reported throughout time.[83-86]
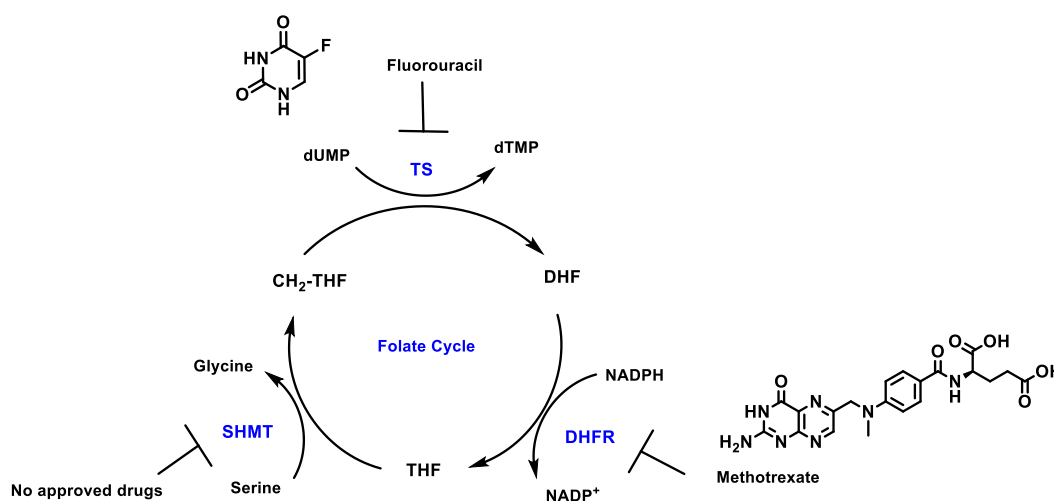
Figure 13: Folate cycle and approved drugs acting on the key enzymes of the cycle[82]

Section C

Peptides

1.7 Introduction

From humble beginnings as substances isolated from livestock glands, peptides have proved to be very useful in the pharmaceutical scene.[87] Peptides have remarkably influenced therapeutic approach and have come a long way since the first isolation of Insulin. Usually composed of 2-50 amino acids that are linked together *via* peptide bonds/ amide bonds, (–CONH–)[88] peptides represent exclusive properties that are not seen in either small molecule drugs or in proteins and are also biochemically and therapeutically distinct from both.[89] Peptides attain different structures and conformations owing to flexibility around alpha carbon.[90] In comparison to other classes of therapeutic moieties, peptide based drugs display many advantages such as high selectivity, specificity, and efficacy for recognizing and binding to targets.[91] Broad range of molecules can be designed to give therapeutically important peptides in fields such as oncology, immunology, infectious disease and endocrinology.[92]
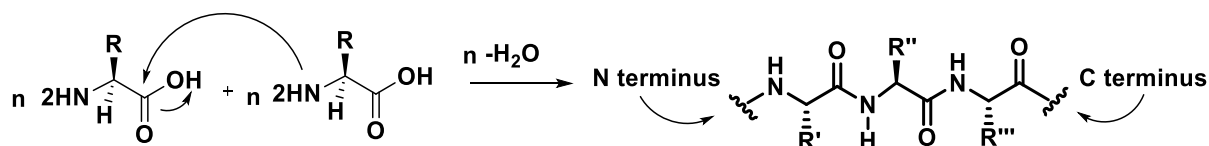
Figure 14: Formation of peptide bond and general structure of peptides

1.8 Chemical synthesis of peptides

Traditionally, synthesis of peptides can be accomplished by 2 major techniques; solid phase and liquid phase synthesis. Typically peptides are synthesized from C → N direction by the coupling of Nα protected amino acids, followed by removal of protecting group of the amino functionality and repetition of this 2-step cycle.[93] In the early days of peptide discovery, synthesis was carried out in solution which required careful manipulation of protecting groups, multiple challenging workups and purifications. Although this method may be suitable for synthesis of short peptides in large scale, synthesis of longer and more complex peptides proved problematic.[94] These shortcomings of solution phase synthesis were overcome by the use of solid support since it displayed obvious advantages.[95] Currently employed and most prevalent method of solid phase peptide synthesis technique which was initially developed by R. B Merrifield in 1963 has displayed numerous advantages over traditional peptide synthesis which resulted in the Nobel Prize in Chemistry in 1984.[96] Simplicity, efficiency and speed of the reactions carried out using a solid support were the main advantages of this method.[97] Shortcomings of solution phase synthesis such as time consuming laborious purification, usage of large quantities of other reagents and chemicals such as coupling reagents, solvents etc were overcome by this milestone.[98]

1.8.1 Solid Phase Peptide Synthesis

Advantages over conventional solution phase synthesis in terms of efficiency and the ease of purification are the main attractions of this method. Reagents that are used in excess to drive the reaction to completion can be easily removed by simple filtration and washing. Therefore, high boiling solvents such as dimethylsulfoxide (DMSO), N-methylpyrrolidone (NMP), N,N'-dimethylformamide (DMF) can be used without having to evaporate them. Additionally, all the above simple steps can be easily incorporated as an automated synthesis.[99] As shown in figure 15, attachment of first amino acid to the solid support usually *via* a linker (i), selective deprotection of the protecting group of the α -amino group (ii) further cycles of coupling and deprotection (iii) cleavage of peptide from the resin to give desired peptide (iv) are the main stages of solid phase synthesis. Newer strategies have been to increase the yield and efficiency of the synthesis.[100]
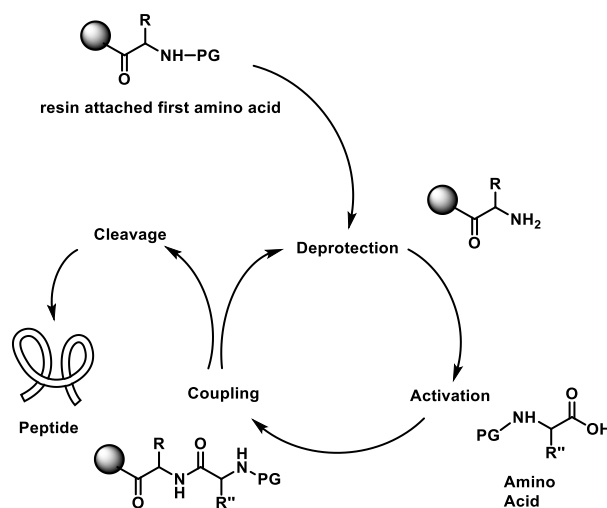
Figure 15: General steps for solid phase peptide synthesis

Table 2: Key techniques shaping synthesis of peptides
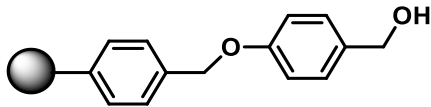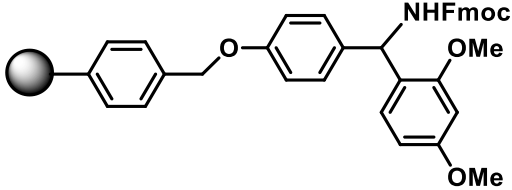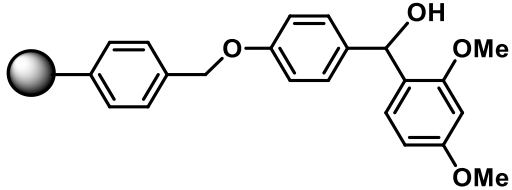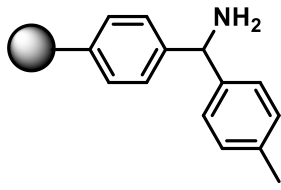
Sr. No.	Year	Discovery/ Invention
1	1881	Discovery
2	1901	First published synthesised dipeptide
3	1932	Discovery of Z group
4	1953	First synthesis of oxytocin
5	1957	Introduction of Boc protecting group
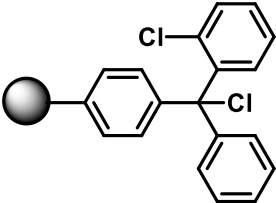
6	1963	Discovery of SPPS
7	1970	Introduction of Fmoc protecting group
8	1970-73	Introduction of other resins such as MBHA and Wang resin
9	1990's	Development of Native Chemical Ligation for peptide synthesis

1.8.2 The solid support

The most common solid support or matrix made of Polystyrene (PS) is used for peptide synthesis.[97,101] Appropriate physical and chemical properties such as uniform shape, size, inertness, good mechanical stability, should be insoluble in all solvents are the important features for a good solid support. Additionally, the solid support must also contain appropriate functional groups for linker attachment.[101]

Table 3: Commonly employed solid supports for solid phase peptide synthesis[97]

Resin name	Resin structure	Cleavage conditions	Peptide product
Wang resin		90-95% TFA in DCM, 1-2 h	Acid
Rink amide resin		50% TFA in DCM 1 h	Amide
Rink acid resin		1–5% TFA in DCM 5–15 min or 10% AcOH in DCM, 2h	Acid
MBHA resin		HF 0º C, 1 h	Amide

2 Chlorotrityl resin		1–5% TFA in DCM 1 min	Acid

1.8.3 Protecting groups

Protecting groups are essential during coupling reactions to avoid by product formation. They should be easily introduced and should be easily removed during synthesis and should be stable to a broad range of reaction conditions. Selectively removing protecting groups under different conditions either during the synthesis or during cleavage of the peptide from the resin is followed during the synthesis. Groups such as Fmoc/Boc/tBu are most commonly used that serve as protecting groups for N α -amino of amino acids and the functional groups of the amino acid side-chains.[102]
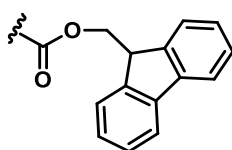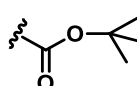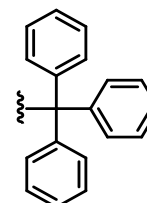
Figure 16: Most commonly employed orthogonal protecting groups for amines

1.8.4 Coupling reagents

Activation of the carboxylic acid is a pre-requisite for the amide bond formation since this bond formation does not take place spontaneously. Activation is commonly achieved using benzotriazole coupling reagents such as HCTU, PyBOP, HATU, or HBTU and carbodiimides, such as DIPCDI and DCC. Recently, a new and safer class of coupling reagents with high efficacy such as COMU and PyOxim, have been incorporated during the synthesis.
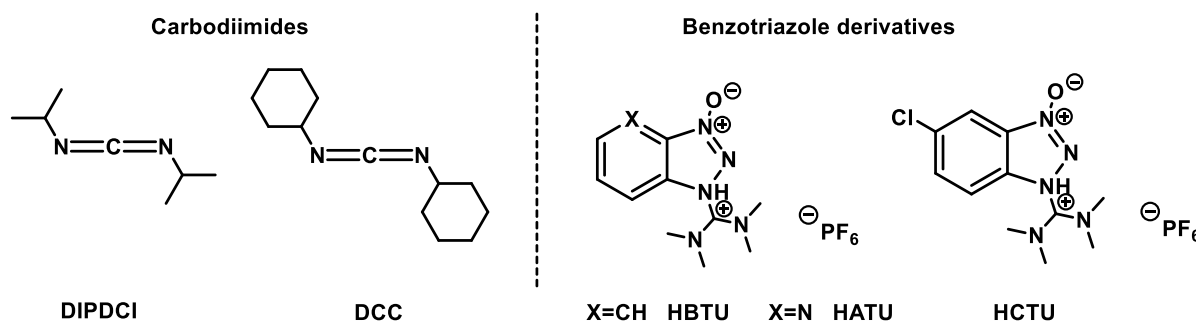
Figure 17: Representative coupling reagents

PRESENT WORK

Chapter 1 introduces the general concepts of antidiabetics, antifolates and peptides which form the basis of this thesis. Titled "**Introduction to antidiabetics, antifolates and peptides**" it briefly describes the importance of antidiatetics and antifolates as a class of drugs used to treat various diseases. Although these two classes of drugs can treat specific conditions in the body, peptides on the other hand fall under a wider category of drugs molecules that can be used to treat a wide range of diseases. The applications of peptides to multiple targets have emerged as powerful tools in research, diagnostics, vaccine development, and therapeutics. In addition, their high specificity and low toxicity offer a viable alternative to the small molecule drugs. in recent years, while the number of new chemical entities attributable to small molecules discovered each year has remained unchanged, those attributable to peptides and proteins has seen a steady increase. Antidiabetics are a class of drugs aimed towards reducing high blood glucose levels that prevail in both type of Diabetes Mellitus (type 1 in children and type 2 diabetes in adults). The high glucose level in the blood or hyperglycemia is due to loss of insulin secretion, action, or both. Although progresses have been made, the pursuit for the "perfect" antidiabetic drug still continues. Antifolates are the first class of antimetabolites to enter clinics more than 50 years ago. They are similar in structure to folates which makes them folate antagonists. Since folates are essential for various metabolic processes of cells, folate antagonism can be successfully employed for multiple diseases such as cancers, malaria and bacterial infections. Additionally, the absence of *de novo* folate biosynthetic pathway in humans makes it an attractive and potential target for drug development.

Chapter 2 titled "**Oxytocin and its analogues mediated glucose uptake in Chinese hamster ovary cells**" describes the solid phase peptide synthesis of native oxytocin and its analogues and their ability to cause glucose uptake in Chinese hamster ovary cells. Oxytocin is a bioactive peptide with 9 amino acids and a disulfide bond. It is a well known peptide and its functions were thought to be associated with inducing lactation and labour. But recent studies have revealed many other functions of this peptide which include glucose uptake in adipocytes. Hence a study was carried out to assess this peptide's ability to cause a glucose uptake in ovary cells. Along with native oxytocin, its disulfide depleted linear and amide cyclised analogues were also synthesised and studied for their glucose uptake ability. Since disulfide bond in peptides plays a role in maintaining structural integrity and play no direct role in the function

of the peptide, the disulfide bond of the native oxytocin was replaced by a more robust amide bond. The synthesised analogues were further purified by RP-HPLC and characterised by MALDI-TOF. Any perturbations in the secondary structure caused due to the replacement of the disulfide bond were studied by circular dichroism. Glucose uptake study in the ovary cells revealed a glucose uptake of up to 47% by the analogues.

Native oxytocin
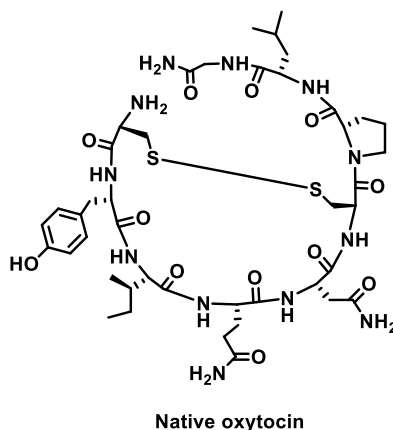
Chapter 3 brings to attention the role and importance of amino acids in many disease conditions. The importance of amino acids in the therapy of conditions such as renal failure, neurological disorders and congenital defects has been well documented. Additionally some amino acids such as lysine and glycine have been reported to have antiglycating activity. During prolonged states of high blood sugar level such as diabetes, glucose forms covalent adducts with plasma proteins through a non enzymatic reaction known as the glycation reaction and the products formed are known as Advanced Glycation End Products(AGEs). The chapter titled "*N*-**(3-Aminoalkyl)proline derivatives with potent antiglycation activity**" describes the synthesis of a series of non natural *N*-(3-aminoalkyl)proline derivatives as antiglycating agents. This new series demonstrates the ability of small molecules to cause an antiglycating effect. Another advantage of stability *in vivo* was also envisaged with these non natural analogues. Various physicochemical assays such as circular dichroism (CD), fluorescence spectroscopy, MALDI-TOF and LC-MS/MS assays show that these compounds are capable of exerting good anti-oxidant properties and display low cytotoxicity, thus enhancing their value as potential anti-glycation agents. Halting the glycation reaction by forming an adduct with glucose at the initial stage was the mechanism by which these derivatives caused an antiglycation effect.
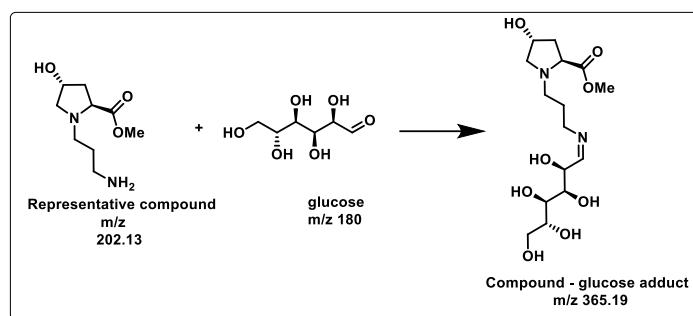
Chapter 4 titled **"Design, synthesis and evaluation of antifolates derived from Guanine"** describes describes synthesis of antifolates derived from guanine nucleobase. Folate species play an essential role in most organisms as cofactors for the biosynthesis of critical cellular components in nucleobases, proteins, and cofactors. But humans lack the *de novo* folate biosynthesis pathway and must obtain this essential nutrient from their diet in the form of the vitamin B9, folic acid. Also, the presence of this biosynthetic pathway in microorganisms, but not in humans, makes it an ideal target for drug development. Hence in connection with efforts in searching for novel antifolates, molecular docking studies were carried out for the title ligands (Ligands 1, 2, 3, 4), folic acid and a marketed drug 'Pemetrexed' against the protein Serine hydroxymethyltransferase (SHMT). SHMT is a ubiquitous enzyme found in all prokaryotes and is the only enzyme yet to be exploited as a target for cancer chemotherapy. Docking study revealed that all the designed ligands showed a good interaction energy with all isoforms of SHMT as did Pemetrexed. Additionally, the ligands when docked against FRα protein (4LRH) showed similar and in some cases higher binding energy compared to folic acid. This result was in line with the hypothesis of this study since the designed ligands were envisaged to have a competitive binding with folic acid. The figure below depicts chemical structures of the designed antifolate ligands.
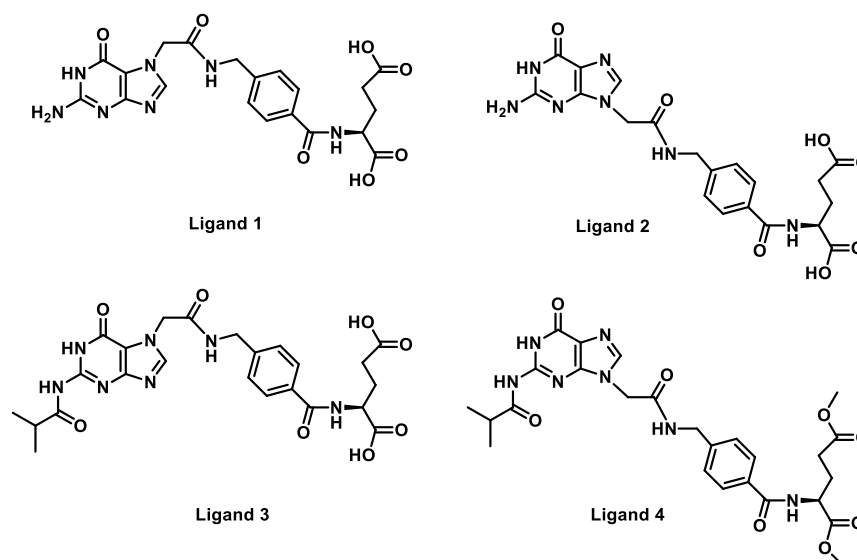
Chapter 2

Oxytocin and its analogues mediated glucose uptake in Chinese Hamster Ovary cells

2.0 Introduction

Although the past decade witnessed a rise in peptides/ proteins as therapeutics, certain limitations such as short plasma half life (stability) and negligible oral bioavailability cause a slowdown in their use.[89] To overcome these limitations, techniques like cyclisation help improve bioavailability and stability, guarding the peptides against the enzymatic degradation from aminopeptidases (N-termini) and carboxypeptidases (C-termini). Additionally cyclisation provides peptides with more organised conformations making them more selective towards their specific target.[103, 104]

Disulfide-rich peptides are found throughout nature[105] and have attracted attention due to their remarkable stability.[106] This stability is a result of the disulfide bonds; sometimes a single bond or multiple bonds acting as structural motifs to protect peptides from stress and unfolding conditions.[107] Additionally many other naturally occurring acyclic disulfide rich peptides have been cyclised to improve their stability so as to be potential drug candidates.[106] Despite the above mentioned advantages, synthesis and re-engineering of these disulfide containing peptides has been a challenge mainly due to the difficulty in achieving the correct oxidative folding; which is a major hurdle in presence of multiple disulfide bonds as the number of possible isomers increases.[108]

To achieve correct structural conformation and to overcome the difficulty in synthesis of these peptides, numerous studies have been undertaken which have concluded that all the disulfide bonds in a particular peptide do not play an important role in order to retain its biological activity.[109] Thus a new branch of disulfide bond engineering has emerged as an important strategy and many researchers have come up with novel ways to replace the much essential disulfide bonds with chemical linkages without disturbing the structure or activity of the parent peptide with amide, thioether, selenium, dicarba bridges etc.[110]

Although, many disulfide rich peptides are in clinical stage, several of such peptides have been recently approved for human use.[111]
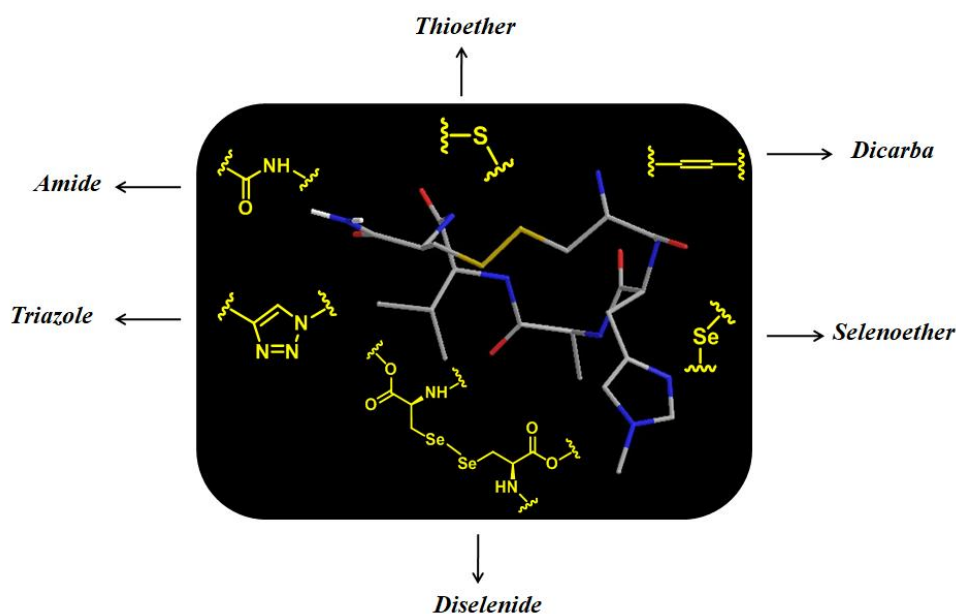
Figure 18: Representative peptide and disulfide bond surrogates (in yellow)

2.1 Oxytocin

Cyclic peptides that exhibit organised, rigid conformation act as excellent scaffolds for design of newer peptides since these peptides would possess better stability and enhanced biological activity.[112] It was the early 1900's when the physiological importance of oxytocin was first described that it assisted uterine contraction and lactation. A hypothalamic hormone peptide, discovered by H. H. Dale in 1909, it was characterised, structurally elucidated and synthesised by 1953. Vincent du Vigneaud in 1953 first sequenced and synthesised oxytocin for which he was awarded Nobel Prize in 1955.[113] Since it is a peptide hormone, oxytocin induces its effects by interacting with high affinity receptors present on the surface of cells in various organs of the body.[114] Numerous studies have displayed evidence that this peptide's function may not only be limited to inducing labour and lactation but it may also exert other physiological and pathological functions, which makes oxytocin and its receptor potential targets for therapeutics.[115]
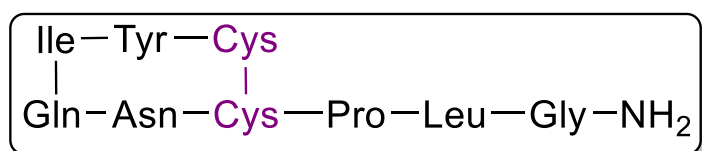
Figure 19: Amino acids in oxytocin peptide

A disulfide cyclised nonapeptide, figure 19 oxytocin is produced naturally in the hypothalamus in the brain. Evidence that this peptide hormone is found in equivalent concentrations in the

plasma of both males and females suggests a far more diverse function not limited to females as considered in the past.[116] Varied functions such as cardio protection, [115]inhibition of growth of neoplastic cells of mammary, nervous, and bone origin, [117]reduction in adipose tissue content and fatty liver have been reported for this peptide. Additionally insulin like effects, glucose intolerance, insulin homeostasis, regulation of body weight, stimulating insulin secretion, increase in β cell responsivity have also been reported by many researchers.[115,118] increase in β cell responsivity and glucose tolerance [119] have been reported by many researchers.

Above studies suggest the involvement of oxytocin in the pathophysiology of diabetes, the hypoglycemic effect being due to the glucose uptake *via* insulin like signalling pathway.[115]
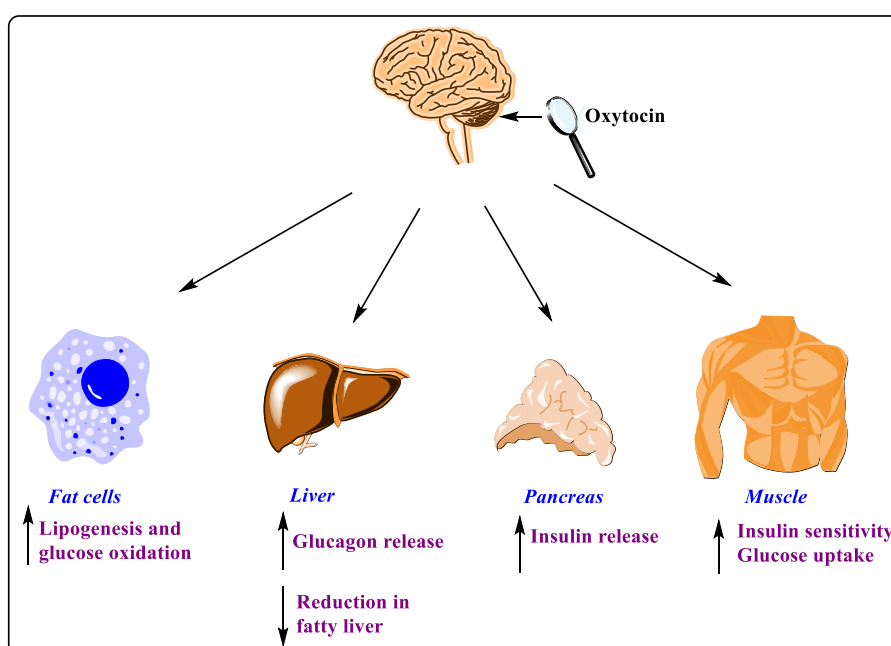
Figure 20: Metabolic effects of oxytocin : OT is secreted from the posterior lobe of the pituitary gland and binds to its receptor in peripheral tissues.

2.2 Rationale

Many disulfides in peptides/ proteins appear to have no direct functional role; rather their main purpose is to maintain the conformation of the protein.[120] Additionally the disulfide bonds are prone to reduction in extracellular environments, by glutathione and disulfide bond isomerases, that further leads to chain unfolding and many a times mismatched oxidative refolding ultimately leading to loss of biological activity. Hence replacements for the disulfide bonds in peptides would lead to analogues that are not susceptible to reduction and further oxidative refolding. [121]
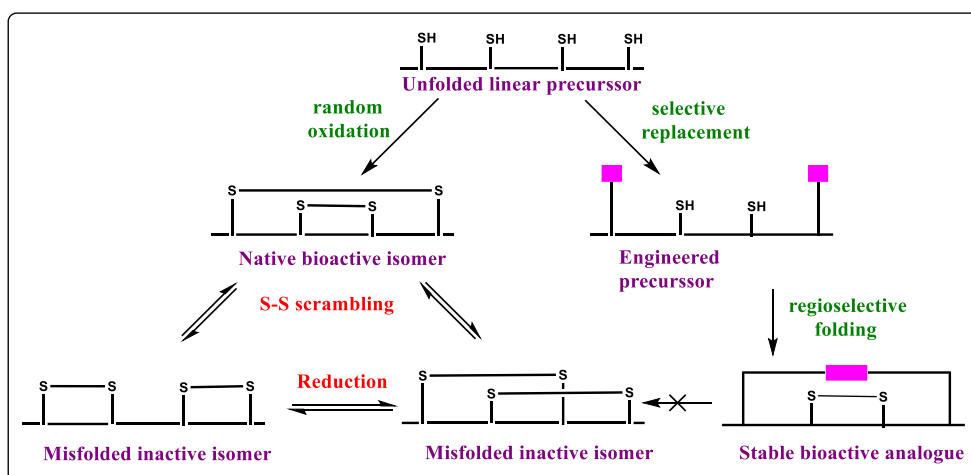
Figure 21: Disulfide bond engineering strategy

The ovary is one of the most dynamic endocrine organs which requires substantial energy source to function. Glucose is an indispensible metabolic substrate of all mammalian cells for energy demand to which ovaries are no exception.[122] Evidence of oxytocin and oxytocin receptor expression in human and subhuman primate ovaries[123] makes it a good target for glucose uptake studies. Hence, in this study, the disulfide bond in oxytocin was replaced by amide bond and the resulting peptides were evaluated for glucose uptake activity in CHO-HIRc-myc-GLUT4eGFPcells.
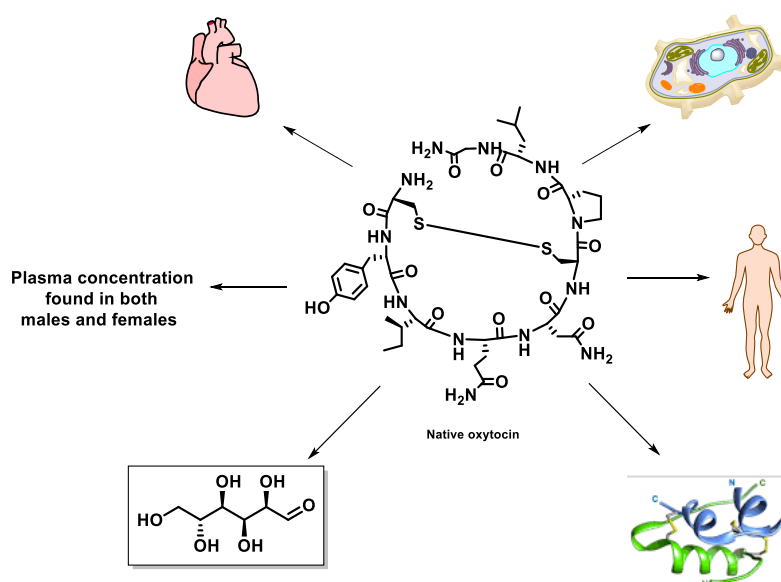
Figure 22: Oxytocin and its various effects on the human body [115, 116, 118]

2.3 Results and Discussion

2.3.1 Synthesis of linear and cyclised peptides

Six peptides, namely O1 (native oxytocin), O2, O3, O4, O5, O6 were synthesized employing solid phase method using Fmoc chemistry on MBHA resin (substitution 1.75mmol/g). Peptide chains were elongated in the consecutive cycles of deprotection and coupling. 3 equiv. of each TBTU, HOBt and DIPEA were used as coupling agents and base. Deprotection of the Fmoc protected amino groups was achieved with 20% piperidine in DMF. Deprotection of Boc group was achieved using 50% TFA in DCM followed by neutralisation of resin using 5% DIPEA in DCM. Coupling and deprotection steps were monitored by the Kaiser Test. It is a qualitative test for the presence or absence of free primary amino groups, and it can be a useful indicator of the completeness of a coupling step. The test is based on the reaction of ninhydrin with primary amines, which gives a characteristic purple colour.[97] After coupling reaction, the resin was capped with acetic anhydride 10eq with 10 eq DIPEA as base in DCM.

For the synthesis of peptide O1, all the amino acids used were Fmoc amino acids. The side chain of cysteine was protected with Acm protection. When the desired number of couplings was reached after consecutive cycles of protection and deprotection, on resin simultaneous Acm deprotection and cyclisation was achieved using I_2 in DMF. This disulfide cyclisation also led to inter cyclised peptide O2. Cleavage conditions for peptide O1were cleaved using TFA and TFMSA for 2h. No reagents containing Thiol were employed during cleavage. Similarly for the synthesis of peptide O3, consecutive cycles of protection and deprotection of Fmoc amino acids led to the desired length of peptide. For this peptide, cysteine on the fourth position was replaced by lysine and the *N*-terminal cysteine was replaced by succinic acid. On resin cyclisation was achieved by employing 3eq each of HOBt, TBTU and DIPEA after deprotection and neutralisation of Boc on lysine side chain by using 50% TFA in DCM and 5% DIPEA in DCM. Cleavage of this peptide using TFA, TFMSA, thioanisole and 1, 2- ethanedithiol for 2hr gave peptide O3. While for peptide O4, after the synthesis of desired length of O3, the anchored peptide was directly cleaved using TFA, TFMSA, thioanisole and 1, 2- ethanedithiol for 2hr. For the synthesis of peptide O5, fourth position cysteine was replaced by glutamic acid and *N*-terminal cysteine was replaced by lysine. After consecutive couplings and deprotections of Fmoc amino acids, side chains of lysine and glutamic acid (Boc and tBu respectively) were deprotected by using 50% TFA in DCM. On resin cyclisation using 3 eq each of HOBt, TBTU and DIPEA was achieved. *N*-terminal Fmoc protection was deprotected prior to cleavage. Cleavage of anchored peptide by

employing cleavage reagents TFA, TFMSA, thioanisole and 1, 2- ethanedithiol for 2hr was achieved. For the synthesis of peptide O6, after the desired length of peptide O5 was achieved, deprotection of N-terminal Fmoc followed by cleavage using TFA, TFMSA, thioanisole and 1, 2- ethanedithiol for 2hr carried out. All the synthesised peptides were purified by reverse phase HPLC and unambiguously characterised by MALDI-TOF.

Scheme 1: Detailed solid phase synthesis of native oxytocin and dimer. Reagents and conditions: 3 equivalents of amino acid, coupling reagents (HOBt and TBTU) and base (DIPEA) were used for coupling reaction. Amino acids: (1) Fmoc Gly-OH; (3) Fmoc Leu-OH; (4) Fmoc Pro-OH; (5) Boc Cys(Acm)-OH; (6) Fmco Asn(trt)-OH; (7) Fmoc Gln(trt)-OH; (8) Fmoc Ile-OH; (9) Fmoc Tyr(tBu)-OH; (10) Boc Cys(Acm)-OH; (11) 10% I$_2$ in DMF:H$_2$O (4:1) for 40 min at room temperature followed by washing with 2% ascorbic acid in DMF. (12) TFA, TFMSA, 2hr, 0°C to room temperature.
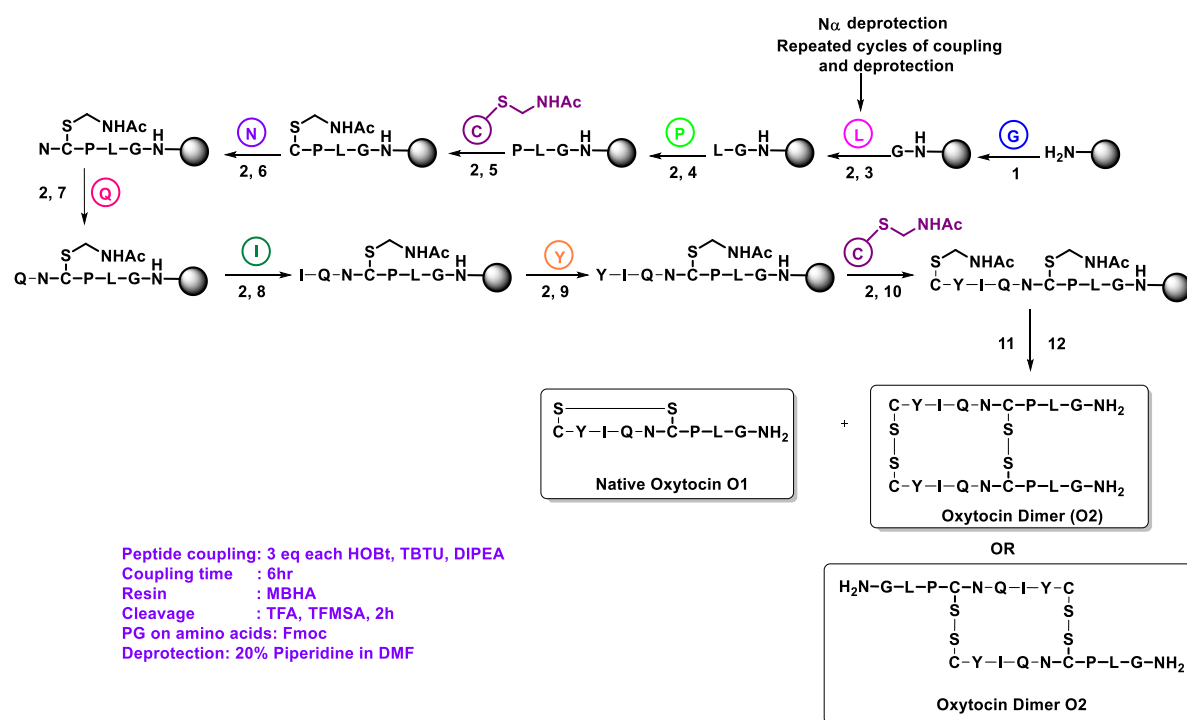
Scheme 2: Detailed solid phase synthesis of amide cyclised and linear analogues of oxytocin. Reagents and conditions: 3 equivalents of amino acid, coupling reagents (HOBt and TBTU) and base (DIPEA) were used for coupling reaction. Orthogonal deprotection (2) was achieved with 20% piperidine in DMF. Amino acids: (13) Fmoc Gly-OH; (14) Fmoc Leu-OH; (15) Fmoc Pro-OH; (16) Fmoc Lys(Boc)-OH; (17) Fmoc Asn(trt)-OH; (18) Fmoc Gln(trt)-OH; (19) Fmoc Ile-OH; (20) Fmoc Tyr(tBu)-OH; (21) Succinic acid; (22) 50% TFA in DCM; (23) 3 eq. each of HOBt, TBTU, DIPEA; (24) TFA, TFMSA, 1,2-ethanedithiol, thioanisole; (25) Fmoc

Gly-OH; (26) Fmoc Leu-OH; (27) Fmoc Pro-OH; (28) Fmoc Glu(tBu)-OH; (29) Fmoc Asn(trt)-OH; (30) Fmoc Gln(trt)-OH; (31) Fmoc Ile-OH; (32) Fmoc Tyr(tBu)-OH; (33) Fmoc Lys(Boc)-OH; (22) 50% TFA in DCM; (23) 3 eq. each of HOBt, TBTU, DIPEA; (24) TFA, TFMSA, 1,2-ethanedithiol, thioanisole, 0°C to room temperature.
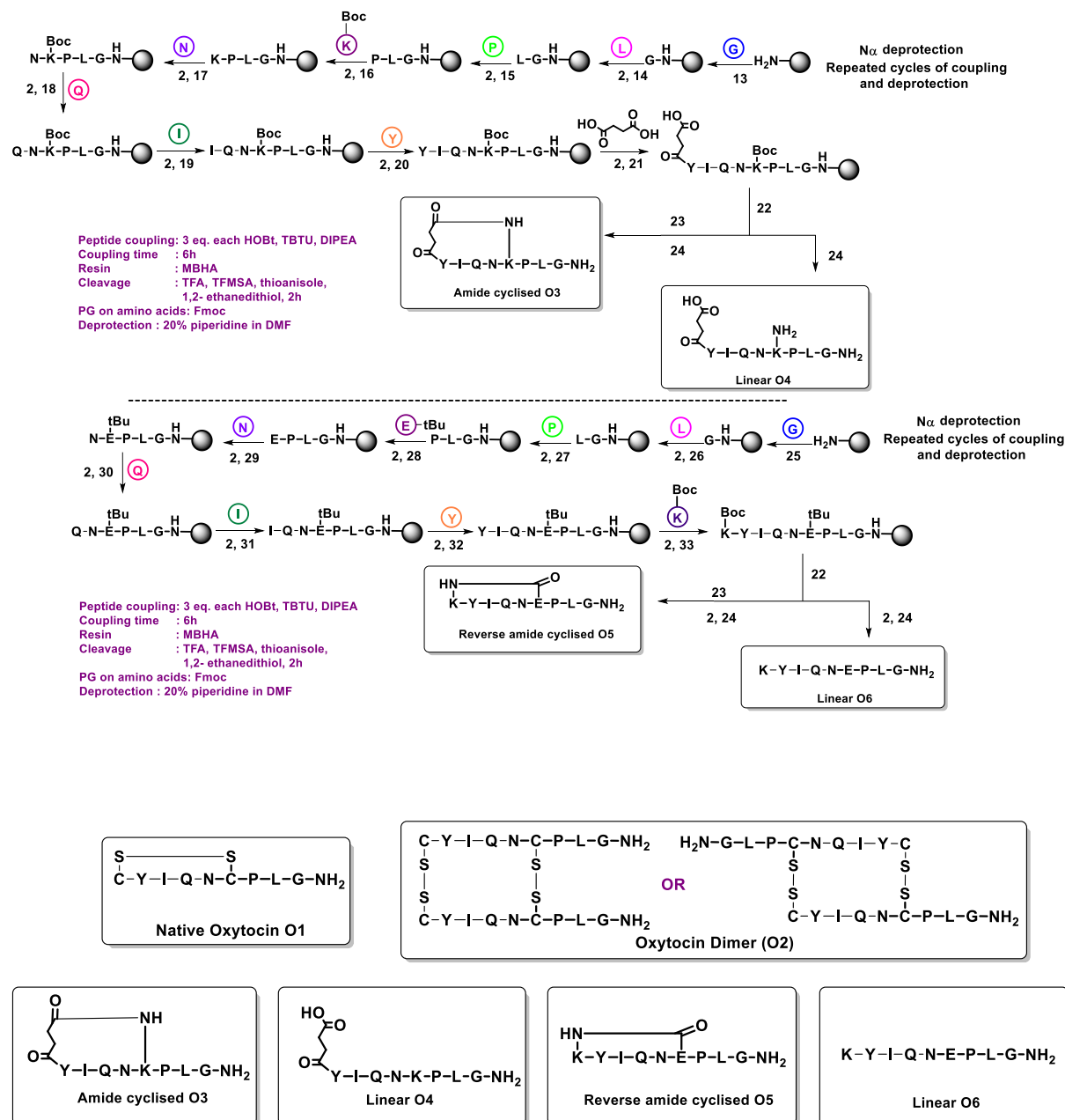
Figure 23: Single amino acid representation of synthesised peptides.

Table 4: MALDI-TOF characterisation of Oxytocin and its analogues

Sr. No	Peptide Name	MALDI-TOF	
		Calculated Mass	Observed Mass
1	O1 (Native oxytocin)	1007	1007(M+H), 1029 (M+Na), 1045 (M+K)
2	O2 (Dimer)	2014	2013, 2035 (M+Na)
3	O3(Amide cyclised)	1013	1013, 1035 (M+Na), 1051 (M+K)
4	O4 (Linear)	1031	1031, 1054 (M+Na)
5	O5 (Reverse amide cyclised)	1042	1042, 1064 (M+Na), 1080 (M+K)
6	O6 (reverse linear)	1060	1060

2.3.2 Structural confirmation by circular dichroism

A quick and correct characterisation of proteins is essential in the field of structural genomics[124] and Circular dichroism (CD) is a valuable and simple technique for studying the structure of proteins because many common conformational motifs, including poly-L-proline II-like helices and turns, α-helixes, β-pleated sheets, show a characteristic far UV (178-250 nm) CD spectra.[125] CD spectroscopic study was carried out to assess the secondary structure of the synthesized peptides. CD spectra were recorded in water at 25 ºC. The absence of the native disulfide bond was found to exert no major structural effect on the analogues of oxytocin. The CD spectrum of all the analogues is similar to the CD spectrum of oxytocin except for peptide O4. O4 exhibits a CD spectrum with two minima at ~220nm and 230 nm similar to an α helix suggesting the possibility of additional secondary structures for this peptide. Remaining peptides showed a maximum at ~230nm and a minimum at ~195 nm which is typical of a random coil nature as depicted in figure 24.
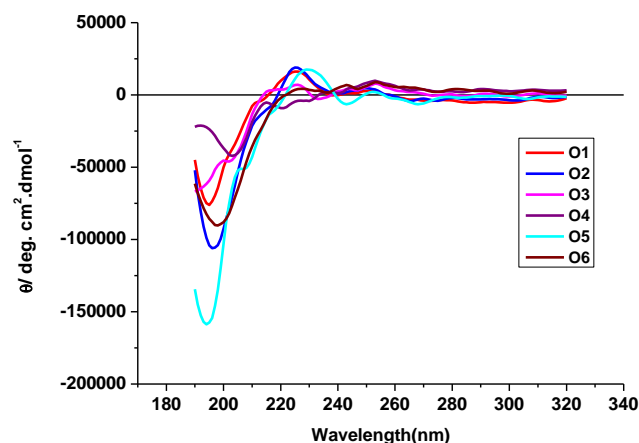
Figure 24: CD spectra of peptides O1, O2, O3, O4, O5, O6 at 60µM in water at 25ºC

2.3.3 Estimation of secondary structure by Circular Dichroism

Knowing the structure of proteins is very important to understand their biological functions.[126] Circular dichroism provides very extensive information in this regard since the UV region between 190nm to ~240nm gives signals from amide chromophores of the peptide bond.[127] CD spectra of proteins whose crystal structure has been already determined serve as reference spectra for a wide range of unknown proteins.[128] Although CD can provide information regarding the secondary structure of proteins/ peptides, since it is an average of all conformers present at the time of measurement, any quantitative estimation should be considered as a useful guideline rather than absolute figure. The secondary structure of the synthesised peptides was estimated using Reed's reference[129] in the secondary structure estimation software provided with the JASCO J-815 spectropolarimeter. The experiment was carried out in water at 25ºC. Estimation based on the CD plots of all the synthesised peptides (Table 5) revealed a 100% random coil nature for all the peptides except for peptide O4 which exhibited a 13% and 87% α helix and random coil respectively. None of the peptides exhibited a β sheet nature. This is also evident from the Figure 24.

Table 5 Secondary structure estimation by Reed's estimation on JASCO J-815 spectropolarimeter

Peptide	% Secondary Structure		
	α helix	β sheet	Random coil
O1	0	0	100

O2	0	0	100
O3	0	0	100
O4	13	0	87
O5	0	0	100
O6	0	0	100

2.3.4 Thermal Stability by Circular Dichroism

Along with estimating the secondary structure, circular dichroism also can be used to determine the thermodynamics and kinetics of protein folding and denaturation.[125] Thermal stability can be studied by a gradual increase in temperature and observing the changes in the spectrum with respect to time. Protein unfolding (at high temperature) and refolding (at lower temperature) can provide important information regarding the thermal stability of proteins/ peptides/ biomolecules.[130] A similar approach was used to see the effect of an increase in temperature from 25ºC to 55ºC (with a step-wise increment of 10ºC) and decrease from 55ºC-25ºC (with a step-wise decrease of 10ºC) in the temperature on the peptides (Figure 25).

For peptide O1, it was observed that when the temperature was increased from 25ºC to 55ºC and vice versa, there was not much change observed in the CD plot, although a small change in amplitude was observed, maximum at 225 nm appears to be more or less unchanged indicating that the increase as well as decrease in temperature did not cause any perturbation in the secondary structure of the peptide. Change in amplitude was observed Similarly when peptide O2 was heated from 25ºC to 55ºC, a slight dip in maxima was observed for temperatures 45ºC and 55ºC, while a decrease in temperature from 55ºC-25ºC resulted in an overall decrease in maxima. Although a change in amplitude was observed for peptide O2, the CD plot is indicative that the secondary structure of the peptide has been preserved throughout the experiment. Although the secondary structure of peptide O3 was preserved during the experiment, a slight dip in the maxima was observed when the temperature was brought down from 55ºC-25ºC. As for peptide O4, a similar dip in the maxima was seen when there was a decrease in temperature from 55ºC-25ºC. Peptide O5 did not exhibit noticeable difference in the CD plot during an increase or decrease in temperature whereas a slight dip in maxima at 25ºC was observed for peptide O6 when the temperature was decreased from 55ºC-25ºC. Altogether, no drastic perturbations were seen during an increase or decrease in temperature for

all the peptides studied. Further, the heating and cooling experiments indicated that the slight decrease in amplitude that was observed during heating was reversible, and cooling resulted in a minimal of the original spectrum observed at 25 ºC.

O1 heating (left) and cooling (right)

(A)
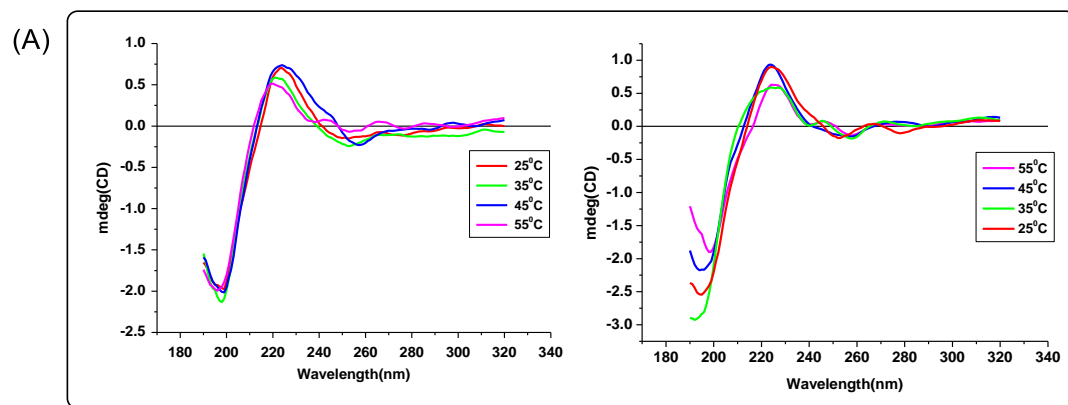
O2 heating (left) and cooling (right)

(B)
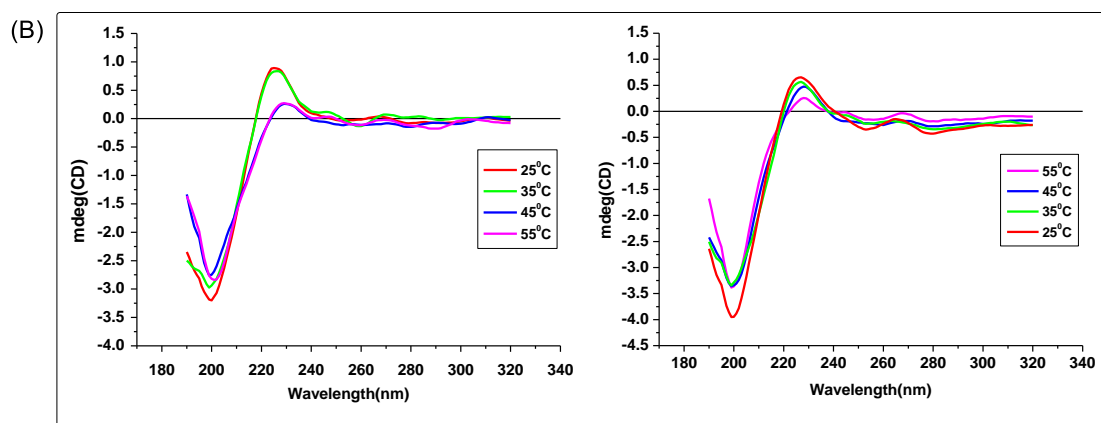
O3 heating (left) and cooling (right)

(C)
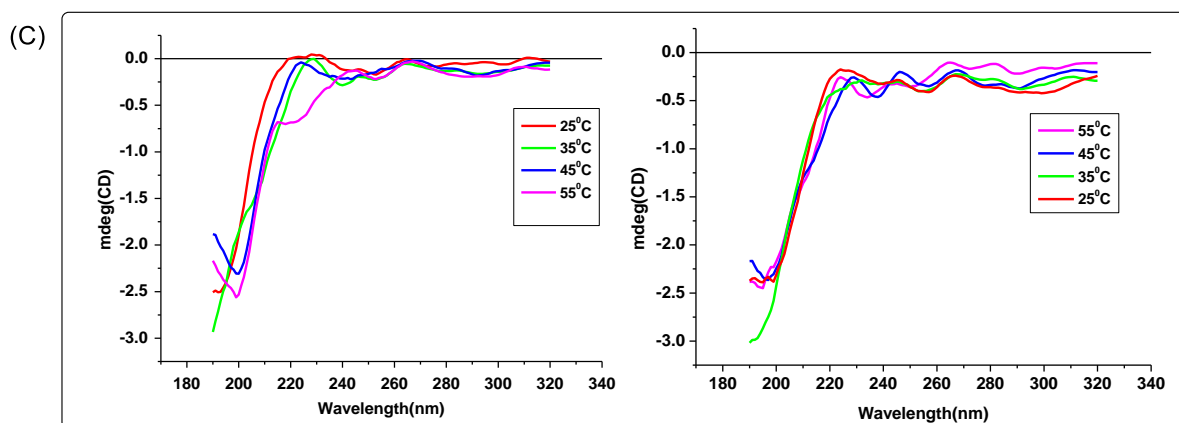
O4 heating (left) and cooling (right)

(D)
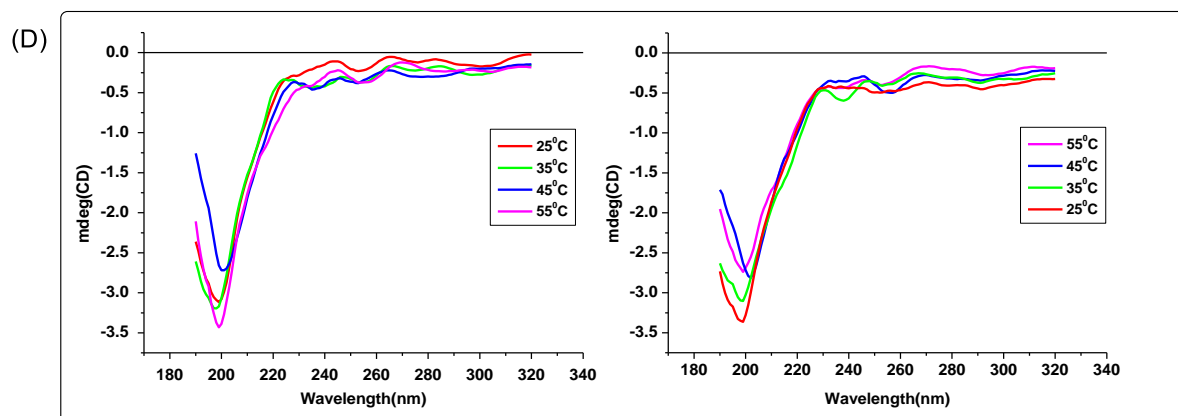
O5 heating (left) and cooling (right)

(E)
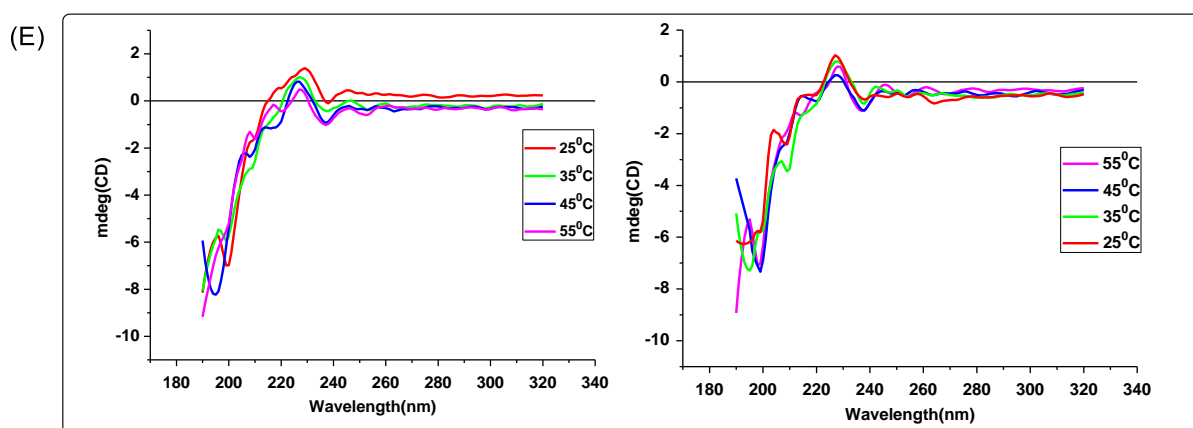
O6 heating (left) and cooling (right)
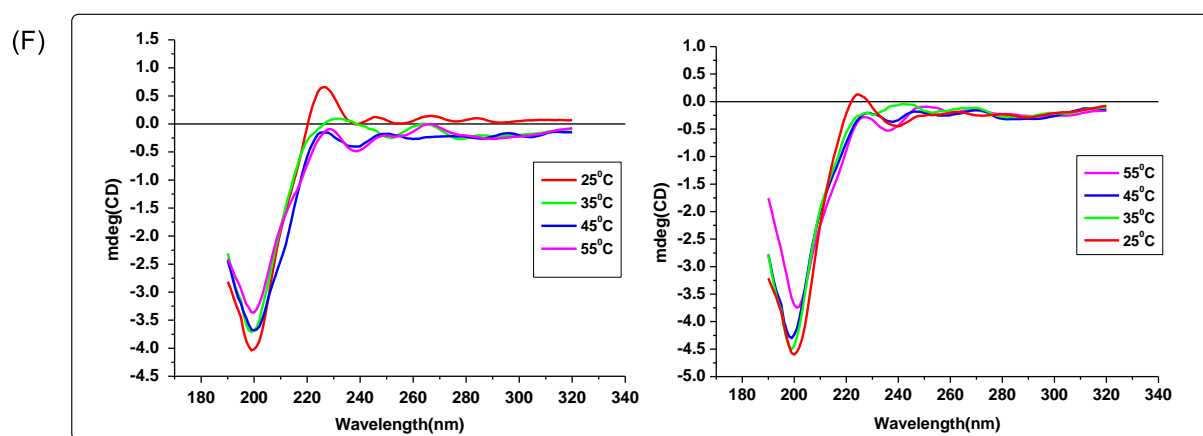
Figure 25: Heating (left panel) and cooling (right panel) plots of peptides from 25º C to 55º C with a 10º C increment and vice versa. Peptide concentration at 60μM in water; O1, O2, O3, O4, O5, O6 plots (A) (B) (C) (D) (E) (F) respectively

2.3.5 Measurement of Glucose uptake by Glucose oxidase-peroxidase assay

Glucose levels were determined by an enzymatic colourimetric glucose oxidase assay in CHO-HIRc-myc-GLUT4eGFPcells. The colorimetric detection of glucose based on the H_2O_2 during glucose oxidase catalyzed glucose oxidation.[131] All the analogues of oxytocin as well as oxytocin itself have shown appreciable glucose uptake in comparison to the control. Particularly peptide O6 at 5 μM and 10 μM concentration showed a glucose uptake very much similar to insulin, which was taken at a standard concentration of 200nM. Insulin produced glucose uptake of about 48.3% while peptide O6 at 5 μM and 10 μM concentration produced glucose uptake of about 42% and 46.5% respectively. (Figure 26)
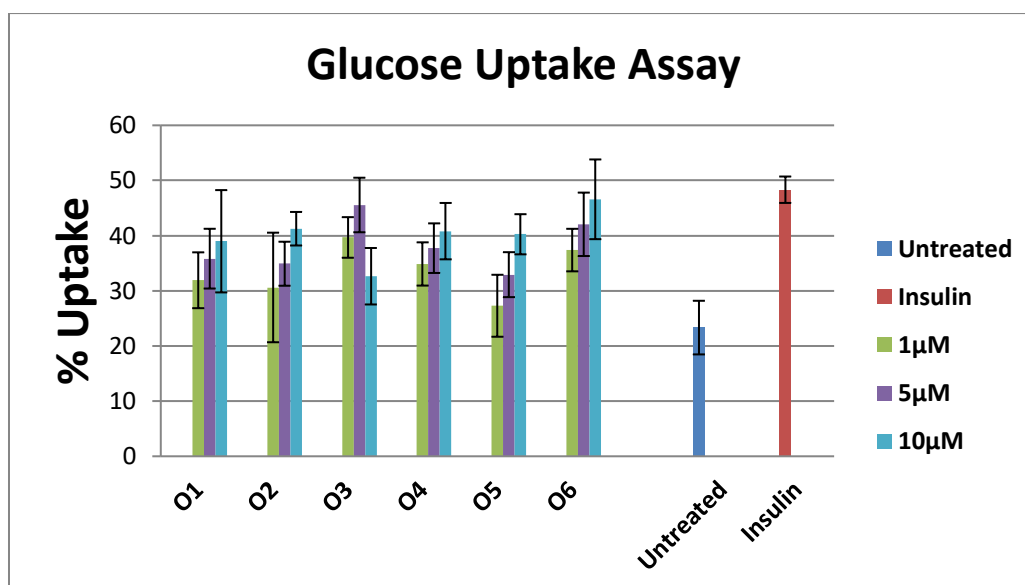
Figure 26: Glucose uptake assay of title peptides in CHO-HIRc-myc-GLUT4eGFPcells at 1 μM, 5 μM and 10 μM concentrations, insulin concentration is 200nM and untreated cells are in absence of insulin and title peptides. Insulin and title peptides were dissolved in deionised water prior to treatment to achieve the experimental concentration at 37º C. Glucose uptake of the cells was measured after 24h of treatment. Error bars represent standard deviation of three independent experiments.

2.3.6 Cell Viability assay

The synthesised peptides were evaluated for the cell viability of CHO-HIRc-myc-GLUT4eGFPcells using the standard MTT assay. Reduction of the yellow tetrazolium dye to a purple formazan dye due to mitochondrial reductase present in live cells is the principle of this assay. Therefore a low colour intensity indicates low cell viability. As seen in Fig.27, there was no evident cyctotoxicity exhibited by the peptides at all the concentrations tested (1 μM, 5 μM and 10 μM). DMSO (15%) was the positive control used in the experiment since this concentration is known to be toxic towards the cells. These results further support the hypothesis that the synthesised peptides could be employed as glucose uptake agents without causing any evident cytotoxicity towards the cells.
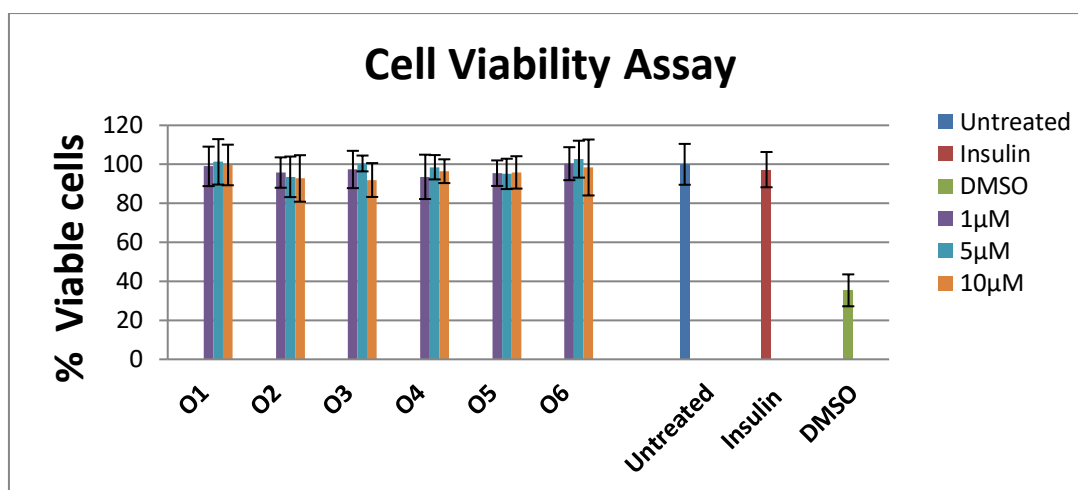
Fig 27: Cell viability assay depicting viability of CHO-HIRc-myc-GLUT4eGFPcells in presence of title peptides (O1, O2, O3, O4, O5 and O6) at 1 μM, 5 μM and 10 μM. Insulin concentration is 200nM and untreated cells are in absence of insulin and title peptides. 15% DMSO is the positive control. Insulin and title peptides and DMSO were dissolved in deionised water prior to treatment to achieve the experimental concentration at 37º C. glucose uptake of the cells was measured after 24h of treatment. Error bars represent standard deviation of three independent experiments.

2.3.7 GLUT-4 translocation assay

Cellular glucose uptake is mainly accomplished *via* facilitated transport mediated by a family of glucose transporters (GLUTs): GLUT1 is responsible for basal glucose uptake, and GLUT4 is insulin regulated GLUT.[132] Insulin-stimulated translocation of GLUT4 to cell membrane leading to glucose uptake is the rate-limiting step in diabetes and also a defined target of antidiabetic drug research. This assay describes a real-time, visual, cell-based qualitative GLUT4 translocation assay wherein GLUT4 translocation is visualized by live cell imaging based on GFP fluorescence.[133] CHO-HIRc-myc-GLUT4eGFP cells overexpressing insulin receptor and GFP tagged GLUT4 were used as insulin responsive cell line model. Figure 28 depicts the translocation of GLUT-4 and its quantification in the presence of insulin, peptide O1 and peptide O2 after 30 min. Native oxytocin (O1) and the dimer (O2) were used as representative peptides for this study. Panel A shows the GLUT-4 translocation from the nuclear region to cell membrane. In the control cells (con panel), there is no movement of GLUT-4 from the nuclear region to the cell membrane. Whereas for cells treated with insulin (Ins panel), there is a movement of GLUT-4 after 30 min of treatment to the cell membrane. Similarly for peptides O1 and O2 there is a movement of GLUT-4 from nuclear region to cell membrane after 30 min. Panel B shows the quantification of GLUT-4 translocation for control/

untreated cells, cells treated with insulin and peptides O1 and O2. When compared with control cells, peptides O1 and O2 show significant glucose uptake.
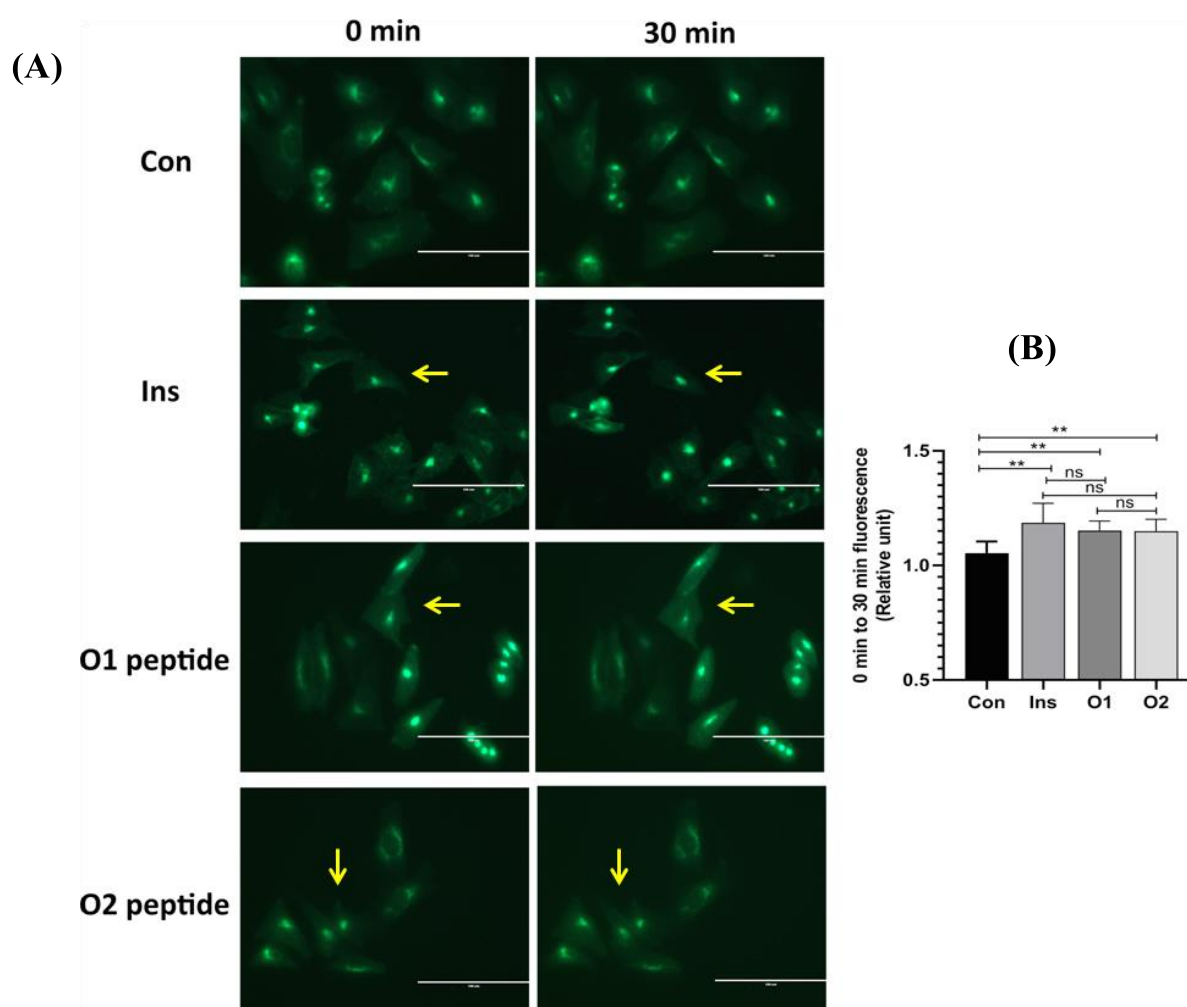
Figure 28: (A) Glut-4 translocation assay for peptides O1, O2 (500nM) and Insulin (200nM) at 0 min and 30 min intervals at 37°C and (B) quantification of GLUT4 translocation by taking the ratio of the membrane fluorescence intensity to total integrated fluorescence of the cell using ImageJ software. Control/ untreated cells; con, Insulin treated cells; ins. Asterisks indicate significant GLUT-4 translocation analysed by one way ANOVA and unpaired t-test (p<0.05). Error bars represent standard deviation of the experiment.

2.4 Conclusion

Native oxytocin, it's linear and amide-cyclised analogues were synthesized by solid phase peptide synthesis. The secondary structure of all the analogues was of a random coil except for peptide O4 which displayed a small α-helix component additional to the random coil structure studied by circular dichroism. *In vitro* cyctotoxy of the peptides towards Chinese Hamster Ovary (CHO) cells showed almost no cytotoxicity upto 10μM concentration. Glucose uptake efficiency of the synthesized peptides in CHO cells was analysed by glucose oxidase peroxidase assay. All the peptides showed appreciable glucose uptake. Peptides O6, O3 and O5 showed similar glucose uptake as insulin. Furthermore, GLUT-4 translocation assay for peptides O1 and O2 was also carried out wherein it was confirmed that the synthesized peptides caused signaling that lead to GLUT-4 translocation to the cell membrane causing glucose uptake. Temperature dependent structural changes were also studied by circular dichroism wherein gradual temperature rise from 25ºC to 55ºC with an increment of 10ºC and vice versa showed no significant structural change. Altogether, the above studies show that replacement of the disulfide bridge in native oxytocin led to analogues with better activity. Hence like oxytocin, many other peptides containing disulfide bonds may act as starting points for the synthesis of disulfide-depleted peptides with potential biological activities.

2.5 Methods

2.5.1 Synthesis of native oxytocin (O1) and dimer (O2)

Synthesis of native oxytocin and the dimer was carried out on MBHA resin with a loading of 1mmol/g. All the amino acids used for the synthesis were Fmoc amino acids. For coupling of amino acids, HOBt, TBTU were the coupling reagents used and base used was DIPEA. After coupling reaction, the resin was capped with acetic anhydride 10eq with 10 eq DIPEA as base in DCM. During the synthesis, after coupling of amino acids to the resin, deprotection of Fmoc was carried out by 20% piperidine in DMF(5 mins * 3 times). After deprotection, thorough washing of resin was done with DCM and DMF and the next amino acid was coupled to the resin. Coupling and deprotection steps were monitored by Kaiser Test. After attachment of the last amino acid to the resin, simultaneous deprotection of S-Acm group and cyclization was carried out in presence of 10% Iodine in DMF: H_2O (4:1) for 40min at rt. After cyclization, the resin was washed with 2% ascorbic acid in DMF. Final washing with DMF and DCM was carried out and cleavage of the resin was done using TFA and TFMSA for 2hr, 0ºC to room

temperature. No thiol reagents like thioanisole and 1, 2- ethanedithiol were used during cleavage. The cleaved peptide was purified using RP-HPLC and characterised by MALDI-TOF. During MALDI-TOF analysis, the presence of a dimerised peptide was evidenced. Though ^{1}H NMR analysis was carried out, the direction of dimerization could not be identified.

2.5.2 Synthesis of linear (O4 and O6) and cyclic analogues (O3 and O5)

Synthesis of linear and amide cyclised peptides were as carried out on MBHA resin with a loading of 1mmol/g. All the amino acids used for the synthesis were Fmoc amino acids. For coupling of amino acids, HOBt, TBTU were the coupling reagents used and base used was DIPEA. After coupling reaction, the resin was capped with acetic anhydride 10eq with 10 eq DIPEA as base in DCM. During the synthesis, after coupling of amino acids to the resin, deprotection of Fmoc was carried out by 20% piperidine in DMF (5min * 3 times). After deprotection, thorough washing of resin was done with DCM and DMF and the next amino acid was coupled to the resin. Coupling and deprotection steps were monitored by Kaiser Test. For the synthesis of linear analogue O4, Fmoc Lys (Boc) OH was used instead cysteine at the fourth position and the terminal cysteine was replaced by succinic acid. After attachment of succinic acid, the peptide was cleaved from the resin using TFA, TFMSA, thioanisole and 1, 2-ethanedithiol. Cyclisation of this peptide to give an amide cyclised analogue (O3) of oxytocin was carried out while the peptide was still attached to the resin. Boc group on lysine was deprotected using 50% TFA in DCM (15 min * 2 times) and neutralised with 5% DIPEA in DCM (3 min * 3 times), after which cyclisation reaction was carried out using 3 equiv. of each TBTU, HOBt and DIPEA at room temperature. Following cyclisation, the peptide was cleaved from the resin using TFA, TFMSA, thioanisole and 1, 2- ethanedithiol.

Similarly for the synthesis of linear analogue O5, the amino acids to replace cysteine used were Fmoc Lys(Boc) OH and Fmoc Glu (tBu) OH where Fmoc Glu (tBu) OH was coupled at the N terminus. Deprotection of Fmoc group was carried out by 20% piperidine in DMF (5min * 3 times) following cleavage using TFA, TFMSA, thioanisole and 1, 2-ethanedithiol afforded linear peptide O5. Simultaneous deprotection of both Boc and tBu was done using 50% TFA in DCM (15 min * 2 times) and neutralisation with 5% DIPEA in DCM (3 min * 3 times) followed by cyclisation in presence of 3 equiv. of each TBTU, HOBt and DIPEA was carried out. After cyclisation, Fmoc group on terminal lysine was deprotected using Fmoc was carried out by 20% piperidine in DMF (5min * 3 times) to afford cyclised analogue O6.
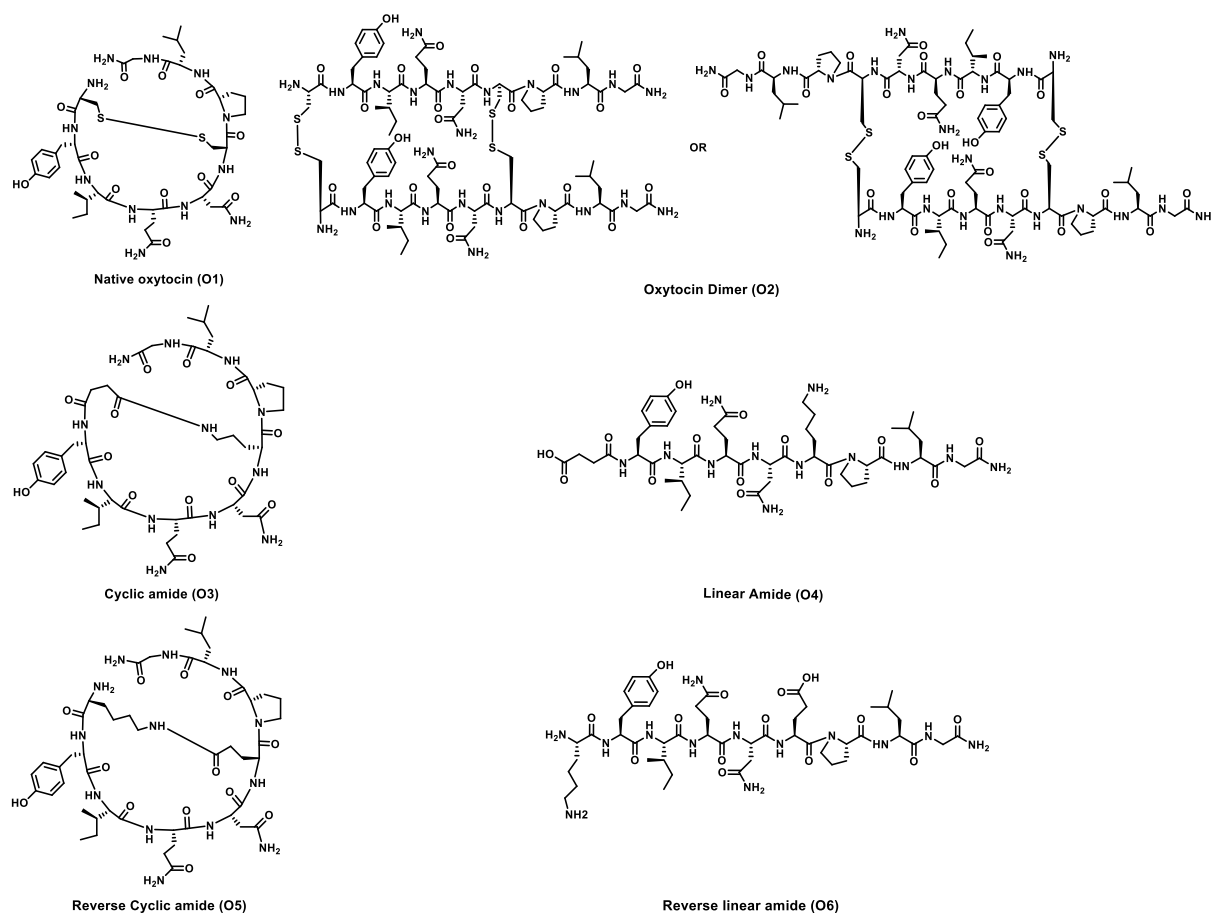
Figure 29: Chemical structures of oxytocin analogue peptides

2.5.3 Picric acid estimation

After first coupling, the functionalised resin was swollen in dry DCM for 30min and then treated with 20% piperidine in DMF for 15min (3*5min) to deprotect the protecting group. After deprotection reaction, the resin was washed with DMF followed by DCM. The free amine thus formed was treated with 0.1M picric acid solution in DCM for 5 mins. Excess picric acid solution was drained off from the reaction vessel. This step allows the formation of picrate salt of the free amine bound to the resin. 1mL 5% DIPEA in DCM was added to the reaction vessel to displace the picric acid. The eluant was collected and the volume was made upto 10mL. Absorbance was recorded at 358nm in ethanol with coefficient of extinction as $14500cm^{-1}M^{-1}$. The following formula was used for the estimation of loading value:

$$\text{Loading of resin} = \frac{\text{O.D X Volume of stock solution X Dilution factor}}{\epsilon\ (14500)}$$

2.5.4 Kaiser Test

Few resin beads were taken in a test-tube and 2 drops of each of ninhydrin, phenol and 0.1% potassium cyanide solution were added to the test-tube and heated for 2 minutes on a hot gun. The presence of free amine groups was confirmed by the appearance of dark blue coloured resin beads in the test tube.

2.5.5 General procedure for cleavage of peptides from resin

After *N*-terminal Fmoc deprotection, cleavage of the peptide from the resin using TFA/thioanisole/EDT/TFMSA procedure was carried out. In a typical cleavage procedure, 5mg resin was treated with 10µL thioanisole and 4µL 1, 2- ethanedithiol. This reaction mixture was kept in ice bath for 10mins followed by addition of 80µL TFA and 8µL TFMSA. This reaction mixture was kept at 0ºC initially and was allowed to warm to room temperature for 2 hr. After the completion of reaction, the resin was filtered using 400 µL TFA followed by evaporation under pressure. The peptides were precipitated by adding ice cold ether. The precipitated peptides were re-dissolved in deionised water and subjected for purification by Reverse Phase High Performance Liquid Chromatography. (RP-HPLC) For the cleavage of peptides containing disulfide bond, only TFA and TFMSA were employed, no thiol containing reagents were employed for cleavage.

2.5.6 Purification of synthesized peptide by Reverse Phase High Performance Liquid Chromatography (RP-HPLC)

The crude peptides were purified by reverse phase HPLC on a C18 column. Following linear gradient, the solvent system A was 95%-5%-0.1% H_2O-ACN-TFA and solvent system B was 50%-50%-0.1% H_2O-ACN-TFA. All the peptides were monitored at 220 nm while purification. All the peptides were more than 95% pure. The purity of the peptides were re-confirmed in solvent system A- 95%-5%-0.1% H_2O-ACN-TFA and B- 50%-50%-0.1% H_2O-ACN-TFA or B- 20%-80%-0.1% H_2O-ACN-TFA. Although some of the peptides show a broad and/or shoulder peak after purification, MALDI-TOF analysis of the broadened peaks and/or shoulder peaks showed a single molecular weight peak only.

2.5.7 Characterisation of synthesised peptides by MALDI-TOF

Synthesised peptides were characterised by Matrix Assisted Laser Desorption Ionisation Time of Flight. (MALDI-TOF AB Sciex TOF/TOF 5800 spectrometer) To facilitate ionisation of peptides, α-cyano 4-hydroxy cinnammic acid as the matrix was used. Peptides to matrix ratio used was 1:1 and ionisation was recorded in the positive reflective mode.

2.5.8 Structure determination by Circular Dichroism

CD spectra were recorded on Jasco J-815 CD Spectrometer equipped with a Jasco PTC-424S/15 Peltier system. Peptides samples in deionised water were prepared at a concentration of 60µM. All the CD spectra from 320nm to 190 nm were recorded at room temperature and represent an average of 3 scans and were corrected for respective blanks. 1mm pathlength cuvette was used for all the accumulations.

2.5.9 Thermal Stability by Circular Dichroism

Thermal stability of the synthesised peptides was evaluated on a JASCO J-815 spectropolarimeter by recording the CD spectra at 25ºC, 35ºC, 45ºC and 55ºC from 320nm-190nm. Peptides samples in deionised water were prepared at a concentration of 60µM. The spectra were collected as accumulations of three scans and were corrected for respective blanks. 1mm pathlength cuvette was used for all the accumulations.

2.5.10 Measurement of Glucose uptake by Glucose oxidase-peroxidase assay

Glucose oxidase enzyme (GOD) oxidizes D-glucose to gluconic acid along with liberation of hydrogen peroxide. CHO-HIRc-myc-GLUT4eGFP cells were seeded at a cell density of 8000 cells per well in a 96 well plate. After the cells adhered and attained their morphology, they were serum starved for 4 h and treated with various concentrations of the title peptides (O1, O2, O3, O4, O5, O6) in triplicate for 24 h, while only serum starved cells served as control. After 24h treatment, the media from each well was collected and analysed for residual glucose. Residual media when combined with standard Glucose oxidase peroxidase (GOD-POD) reagent produces a pink colour proportional to the glucose present in the media. Absorbance was measured at 505 nm using Thermoscientific Multiscan Go plate reader and Glucose oxidase peroxidise kit from Agapee was for the assay. Peroxidase enzyme reacts with hydrogen peroxide to liberate nascent oxygen which in turn reacts to form red dye. The intensity of

colour is directly proportional to concentration of glucose. The intensity of colour is measured colorimetrically at 505 nm.

2.5.11 Cell Viability assay

CHO-HIRc-myc-GLUT4eGFP cells seeded at a cell density of 8000 cells per well in a 96 well plate. After the cells adhered and attained their morphology, they were serum starved for 4 h and treated with various concentrations of the title peptides (O1, O2, O3, O4, O5, O6) in triplicate for 24 h, while only serum starved cells served as control. 25µg of MTT reagent (dissolved in serum starvation media) was added to each well and incubated in dark at 37^0C for 3h wherein violet formazan crystals were observed. Media from each well was discarded and crystals were dissolved in 50 mL DMSO. Absorbance was measured at 575 nm using Thermoscientific Multiscan Go plate reader. This assay was performed using the MTT molecular probes by Life Technologies

2.5.12 GLUT-4 translocation assay

CHO-HIRc-myc-GLUT4eGFP cells were used for live cell imaging to determine the effect of the title peptides and insulin on GLUT4 translocation. In brief, the cells were grown in a 35 mm petri dish until 70–80% confluency was achieved. The slides were placed on a fluorescent microscope (Olympus, Shinjuku-ku, Tokyo, Japan) and the cells were focused under a 40 objective followed by treatment. Insulin (200 nM) and peptides O1 and O2 (500nM) treatment was given to the cells. The images were captured at room temperature at the rate of one frame per min for 10–30 min using a CCD (cooled charge-coupled device), 1.4 megapixel, 12 bit camera (Olympus) attached to a microscope. GLUT-4 quantification was performed using Image J software. A ration of 30 min to 0 min membrane fluorescence was calculated and normalised and relative intensities were plotted in Graphpad Prism software.

Appendix

Characterisation	Page number
MALDI-TOF spectra of peptides	43-45
HPLC chromatograms of peptides	46- 47

TOF/TOF™ Reflector Spec #1[BP = 1029.5, 9635]

Peptide- O1
Calculated. mass- 1007
Observed mass- 1007, 1029 (M + Na), 1045 (M + K)

TOF/TOF™ Reflector Spec #1[BP = 2036.9, 2303]

Peptide- O2
Calculated. mass- 2014
Observed mass- 2013, 2035 (M + Na)

TOF/TOF™ Reflector Spec #1[BP = 1035.6, 1041]

Peptide- O3
Calculated. mass- 1013
Observed mass- 1035 (M + Na),1051 (M + K)

TOF/TOF™ Reflector Spec #1[BP = 1031.6, 924]

Peptide- O4
Calculated. mass- 1031
Observed mass- 1031, 1054 (M + Na)

TOF/TOF™ Reflector Spec #1[BP = 1064.6, 3386]

Peptide- O5
Calculated. mass- 1042
Observed mass- 1042, 1064 (M +Na), 1080 (M + K)

TOF/TOF™ Reflector Spec #1[BP = 1060.7, 2912]

Peptide- O6
Calculated. mass- 1060
Observed mass- 1060

Peptide O1

Peptide O2

Peptide O3
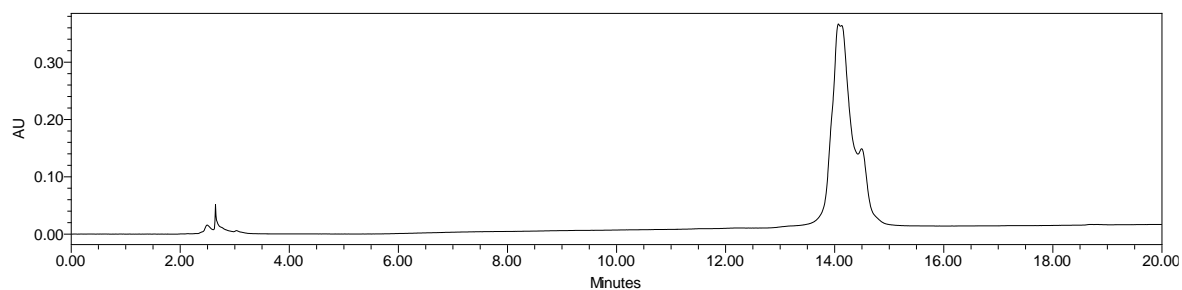
Peptide O4
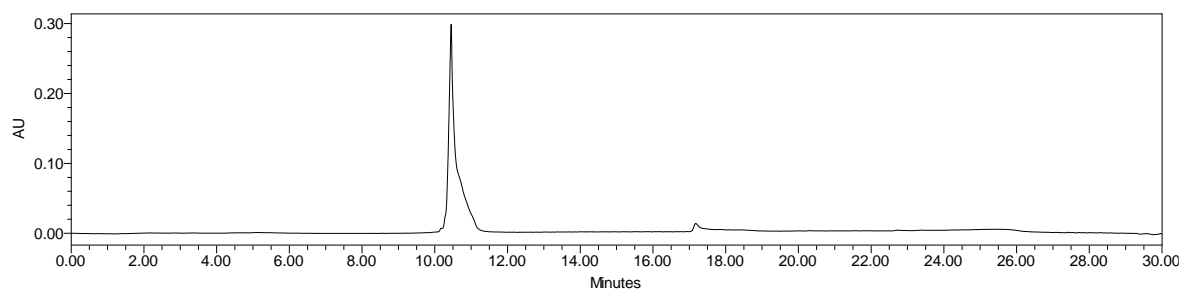
Peptide O5
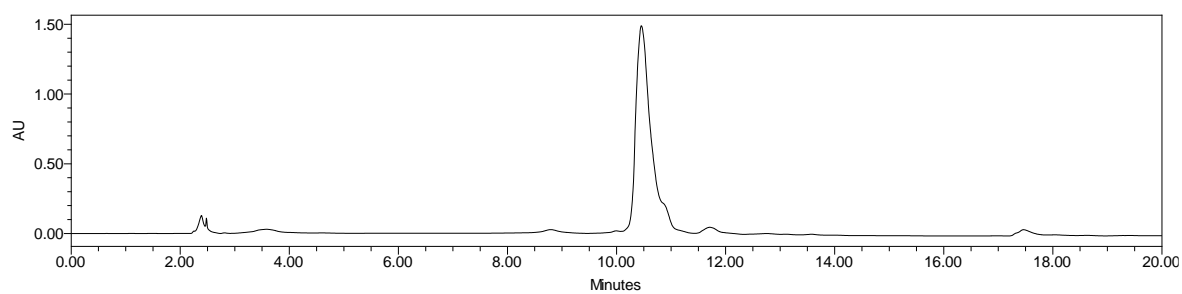
Peptide O6
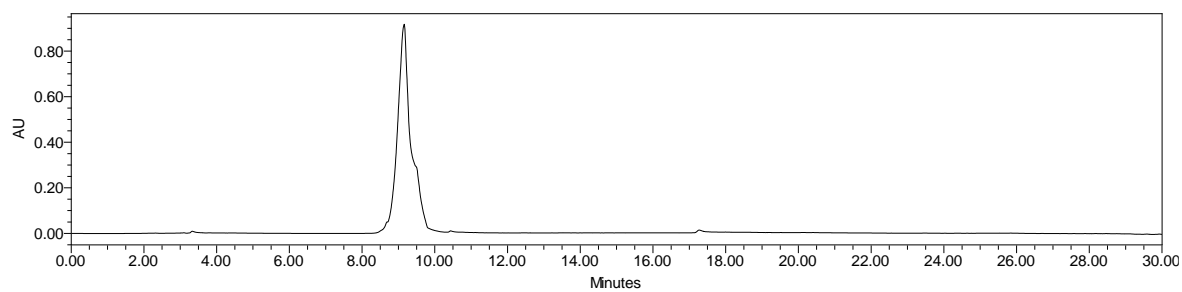
Chapter 3

N-(3-Aminoalkyl)proline derivatives

with potent antiglycation activity

3.0 Introduction

One of the oldest known disease to man is Diabetes Mellitus (DM) which is very well identified by elevated blood glucose levels that warrant frequent monitoring and proper control.[134] Relative or absolute deficiency of insulin secretion by pancreatic β-cells and /or resistance to insulin action leads to a condition of hyperglycemia or high amount of blood glucose level.[5] If left untreated, chronic hyperglycemia leads to serious complications concerning the heart, kidney, eyes etc. An estimated 171 million people in the world suffer from diabetes and it is estimated to reach 366 million by 2030.[135]

3.1 Glucose, Hyperglycemia, Glycation and AGEs

Glucose is the main energy source and is required for cells to function properly and to maintain good health. It consists of six carbon atoms and exists in two isomeric forms as shown in figure 30. Hyperglycaemia is the defining common feature of all types of diabetes[4] which develops when there is an imbalance between glucose utilisation and glucose intake.[136] Non enzymatic protein glycation is the first outcome of high glucose concentration typical of a diabetic state.[137] Reducing sugars react with amino/ guanidino groups of proteins and peptides in the Maillard or glycation reaction and the products obtained thereof are known as Advanced Glycation End products (AGEs).[138] In regards to glycation, the ring structure of glucose makes it highly stable and less reactive with amine residues in proteins. On the contrary, the open configuration accounts for its high reactivity.[139] Compared to other sugars such as ribose, glucose shows the slowest glycation rate because of the prevalence of cyclic form rather than open chain in the body.[140] Although glucose is less reactive compared to other reducing sugars in the body, its level in the body is around 3-6 mM whereas the concentration of other reducing sugars in the body is maintained at much lower concentrations (μM) and hence being the most abundant sugar, contributes the highest to glycation of proteins.[139]
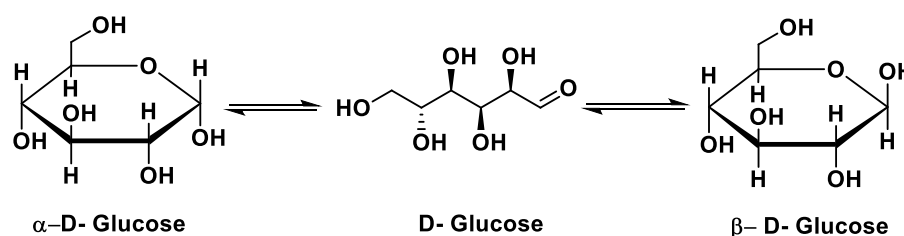
Figure 30: Chemical structures of glucose molecules. Acyclic form of D-glucose (centre) and cyclic forms of α-D-glucose (left) and β-D-glucose (right)

The occurrence of AGEs is high in state of prolonged and chronic hyperglycemia with accumulation in circulating blood and tissues.[141] Several adverse effects are seen due to AGEs like altered structure and function of extracellular matrix caused due to protein cross linking, the interaction with specific receptors and the production of reactive oxygen species (ROS) are some of them.[142] Conversely, this increase in ROS production can in turn lead to additional AGE formation.[143] Along with endogenously produced AGEs, many foods also act as sources and approximately 6-7% enter the body through food and stay in the body for some time.[141]

Different tissues such as arterial collagen, retinal blood vessel walls and renal tissue show evidence of more than a dozen AGE formation, which are formed at a slow but constant rate in the body.[42] These AGEs that form can in turn react with other proteins ultimately leading to crosslinking reactions.[137] Many chronic diseases arise due to the presence of AGEs such as diabetes and diabetic complications, neurodegenerative diseases, and cardiovascular and kidney diseases.[144]

Protein glycation: Initiation stage

The glycation reaction is a type of condensation reaction between amino group of biopolymers with the reactive carbonyl groups of reducing sugars (glucose) to produce a Schiff's adduct such as glycosylamine. The product once formed rearranges itself further to form Amadori product or ketomine compounds that are more stable and almost irreversible.

Protein glycation: Propagation stage

A series of chemical rearrangements, dehydration, degradation and oxidation reactions of the Amadori products give rise to carbonyl compounds that have a potential to react with other amino molecules. The reactivity of the late stage Amadori products is higher than the reactivity of the original carbonyl functionality in the sugar from which they are derived.

Protein glycation: Advanced stage

Advanced Glycation End products are formed in this step. Further cross linking can also be seen in this step. The AGEs formed are of different types such as cross-linked or non-cross-linked and fluorescent or non-fluorescent forms.[139]
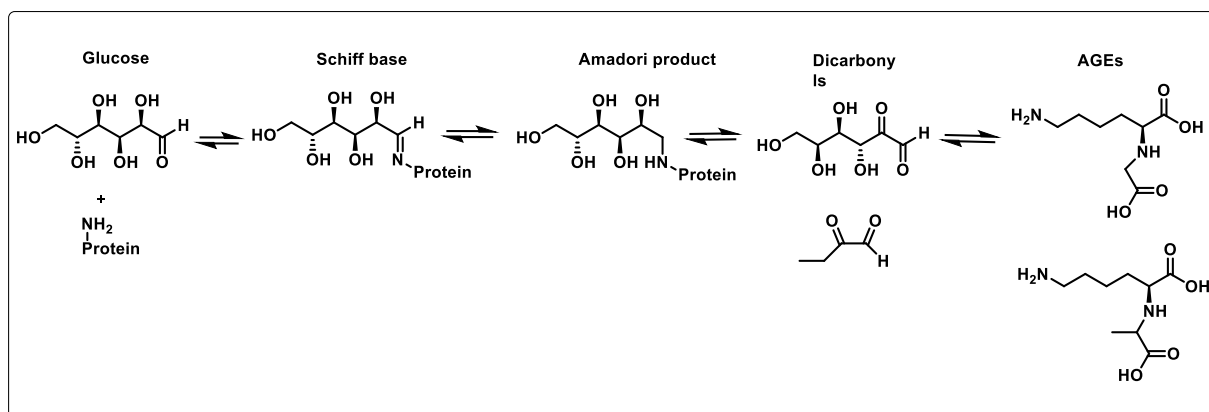
Figure 31: Formation of advanced glycation end-products (AGEs)

3.2 Antiglycation

Since glycation is a reaction between amino groups and carbonyl groups, several macrophages contain receptors that recognize these AGEs and remove them by endocytosis.[145] Although many compounds exist having antiglycating properties, they are still not in use for humans. There require various strategies to reduce the formation and accumulation of glycation products at every stage such as:

- Block free amino groups on proteins, preventing glycation by free sugars.
- Block reactive carbonyl groups on reducing sugars, Amadori products and dicarbonyl intermediates (3- deoxyglucosone, methyl glyoxal, etc.) effectively reducing the formation of products from glycation and/or AGEs.
- Antibodies specific to Amadori products may be used to block their reactivity. This method has the advantage of increased specificity compared to the non-specific blocking of reactive carbonyl groups.
- Chelation of transition metals may reduce glycation-derived free radicals.
- Antioxidants may inhibit the reactivity of free radicals *via* auto-oxidative glycation, glycoxidation.
- AGEs-cross-link breakers offer the potential of reversing diabetic complications.[140,146]

3.3 Rationale

Hypoglycaemic antidiabetic therapeutic agents are known to work by mediating an increased glucose disposal within the body.[147] Some amino acids such as lysine and glycine have also been reported to have antiglycating activity. We therefore surmised that the use of non-natural analogues such as the title compounds would have advantages in terms of stability *in vivo*.

3.4 Results and discussion

Scheme 3: Synthesis of title *N*(3-aminoalkyl)prolyl derivatives.
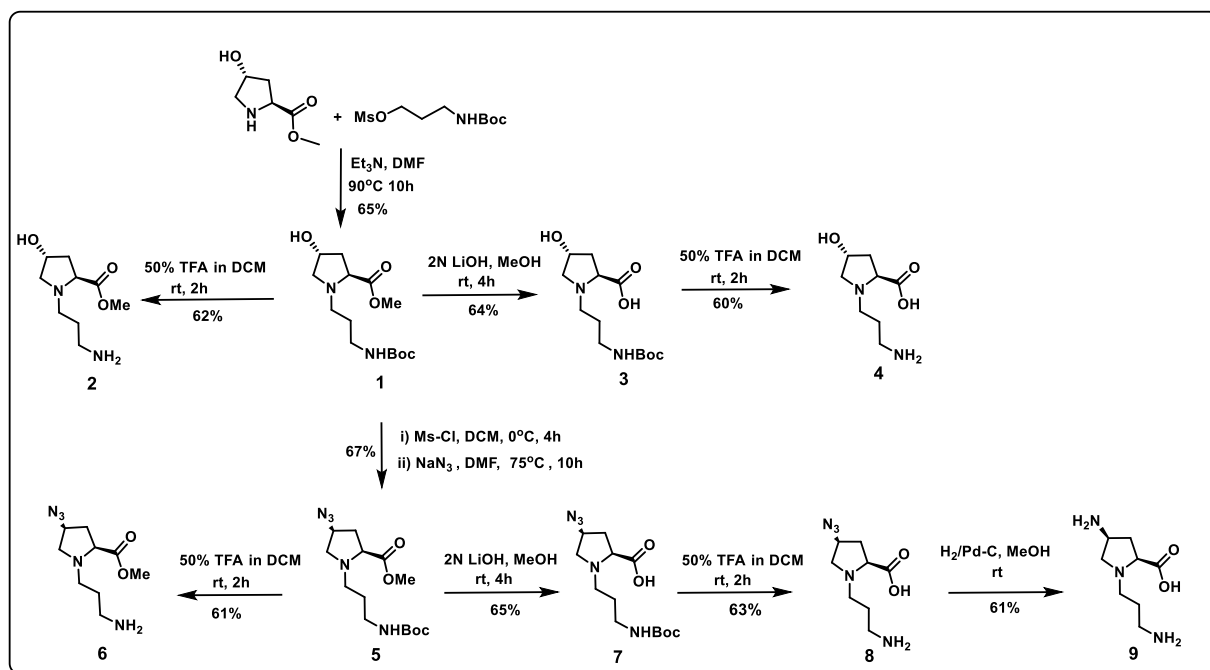
3.4.1 Synthesis of title *N*(3-aminoalkyl)prolyl derivatives.

The title compounds were synthesized using a simple strategy outlined in Scheme 1. Accordingly, 4(*R*)-hydroxy-2(*S*)-proline methylester was *N*-alkylated by treating it with 3-((*tert*-butoxycarbonyl) amino)propyl methanesulfonate in CH_2Cl_2 in the presence of triethylamine to yield compound **1** which was subjected to acidolytic removal of the Boc protecting group to afford compound **2** in 62% yield. Compound **3** was obtained from **1** upon saponification with lithium hydroxide and subsequent purification by column chromatography on neutral alumina. The removal of the Boc protecting group in **3** resulted in

compound **4**, which was obtained in 60% yield. Mesylation of the 4(*R*)-OH group in compound **1** yielded the 4(*R*)-*O*-mesyl derivative, which was further converted to its 4(*S*)-azide counterpart **5** in 67% yield over two steps after purification by column chromatography. Compound **6** was obtained by acidolytic removal of the Boc protecting group in **5**, and purified by column chromatography. Compound **5**, upon saponification with lithium hydroxide yielded **7** in 65% yield after column purification. Compound **7** was further subjected to removal of the Boc protecting group by treating it with TFA in CH_2Cl_2 to yield compound **8** in 63% yield. Reduction of the azide function in **8** by hydrogenation in the presence of Pd–C further gave

compound **9** in 61% yield. All the compounds were unambiguously characterized by appropriate spectroscopic and spectrometric techniques.

3.4.2 Circular Dichroism (CD)

CD is a powerful tool for investigating the structure and conformational changes of proteins, such as those occurring upon glycation, where an increase in the beta sheet conformation is observed. A decrease in the beta sheet conformation is therefore, indicative of antiglycation ability. Analysis of the CD spectra of BSA upon glycation in the presence of the title compounds revealed a decrease in the beta sheet percentage in comparison to glycated BSA, when no compound was present. In particular, the beta sheet conformation in the presence of compounds **2**, **4**, **6**, **8** and **9** was 3%, 1.7%, 2.5%, 2% and 5.2% respectively, while that in glycated BSA was 18.7% and 5.7% in the presence of known anti-glycating agent, aminoguanidine (Amg). Thus, the title compounds are able to protect the conformation of BSA and inhibit beta sheet formation significantly, even in comparison to aminoguanidine. (figure 32)
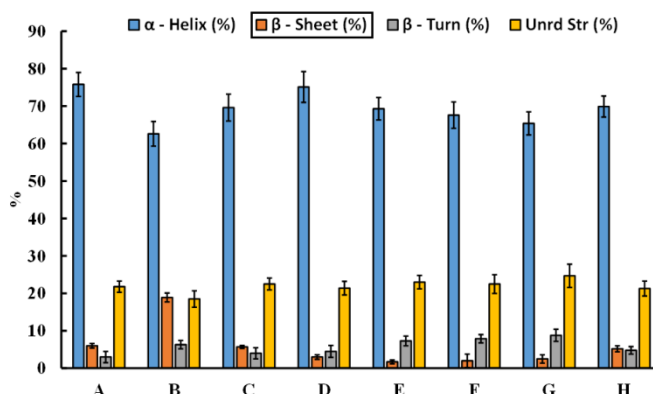
Figure 32: CDPro analysis of (A) native BSA, (B) glycated BSA and (C-H) glycated BSA in the presence of 20mM aminoguanidine and compounds 2, 4, 8, 6 and 9 respectively. (Unrd= unordered.)

3.4.3 *In vitro* glycation inhibition by fluorescence spectroscopy

The degree of glycation can be measured by measuring the fluorescence intensity at 440nm, using an excitation wavelength of 370nm, since most AGEs have a characteristic fluorescence with an excitation maximum approximately at 370nm, and emission around 440nm.[148] A decrease in the fluorescence emission intensity is thus, indicative of inhibition of AGE formation. BSA was incubated with glucose in the presence and absence of title compounds and the fluorescence emission of the compounds was monitored after excitation at 370nm. The formation of AGEs was monitored after 14 days by measuring the fluorescence emission

intensity at 440nm. The fluorescence-based assay for AGEs was also used to determine the IC_{50} values of representative title compounds **2** and **6**. Accordingly, these were found to be 19.2mM and 29.5mM respectively. In comparison, the IC_{50} value of aminoguanidine, considering the fluorescence excitation/emission as 370nm/440nm is reported to be 10mM. However, this fluorescence-based method may be less specific because some fluorescent AGEs differ in their excitation-emission wavelengths. (figure 33)
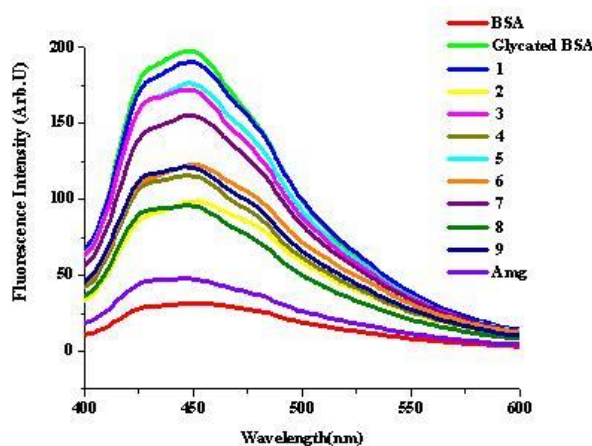
Figure 33: AGE fluorescence spectra of BSA, glycated BSA, and glycated BSA treated with 20mM title compounds and aminoguanidine.

3.4.4 Western Blot Assay

Western blot analyses of the glycation products of BSA in the presence and absence of the title compounds were carried out using both anti-AGE as well as anti-CML (anti-Carboxymethyl lysine) antibodies, since CML is known to be the most abundant non-fluorescent AGE. In both the cases, compounds **2** and **4** were found to exhibit potent anti-glycation activity, which was even superior to aminoguanidine. Ponceau staining of the gels was also carried out to illustrate the protein (BSA) content in the samples. (figure 34) Ponceau S is a diazo dye that may be used to prepare a stain for rapid reversible detection of protein bands on nitrocellulose or PVDF membranes (western blotting), as well as on cellulose acetate membranes.[149]

Anti CML Blot

Anti AGE Blot

Anti CML Blot

Anti AGE Blot

Lane (L-R) compound 9, 6, 8, 4, 2, aminoguanidine, glycated BSA, BSA , ladder

Lane (L-R) compound 9, 6, 8, 4, 2, aminoguanidine, glycated BSA, BSA
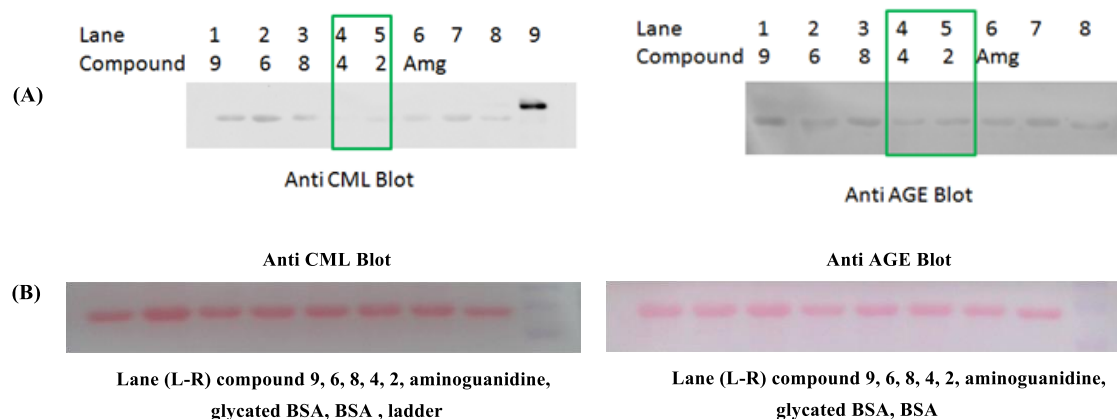
Figure 34 (A) Western blot analysis: Anti-CML Blot: Lane (1-9) compound **9, 6, 8, 4, 2**, aminoguanidine, glycated BSA, BSA, ladder; Anti-AGE Blot: Lane (1-8) compound **9, 6, 8, 4, 2**, aminoguanidine, glycated BSA, BSA **(B)** Ponceau staining: Anti-CML Blot: Lane (L-R) compound **9, 6, 8, 4, 2**, aminoguanidine, glycated BSA, BSA, ladder; Anti-AGE Blot: Lane (L-R) compound **9, 6, 8, 4, 2**, aminoguanidine, glycated BSA, BSA

3.4.5 MALDI-TOF-MS based insulin glycation assay

Following a previously reported MALDI-TOF assay, insulin (*m/z* 5808) was glycated in presence of glucose to form Amadori modifications (*m/z* 5970). The intensity of glycated insulin was monitored in the presence or absence of inhibitors. Figure 35 shows that all the title compounds were able to inhibit Amadori modification of insulin which is indicated by the decreased intensity of the glycated insulin peak, with compound **9** showing the highest inhibition.
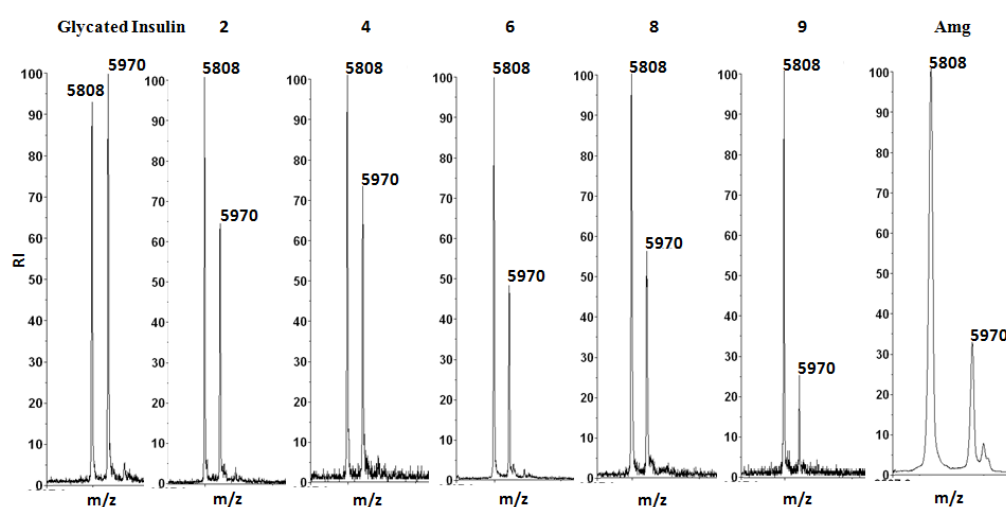
Figure 35: MALDI-TOF assay for glycation inhibition for glycated insulin, insulin in presence of compounds 2, 4, 6, 8, 9 and aminoguanidine.

3.4.6 Fructosamine Assay

In the early stages of glycation, unstable Schiff bases are formed and turned into Amadori products such as fructosamine, which is clinically used as an indicator for short-term control of blood sugar in diabetic patients. Reduction of fructosamine, therefore, is a therapeutic strategy to delay incident vascular complications. The fructosamine assay is a simple colorimetric test that measures glycated serum protein concentrations. Colour change is based on the reduction of nitroblue tetrazolium (NBT) to monoformazan (MF) by Amadori rearrangement products. The fructosamine levels in presence of the title compounds were found to be significantly lower than that in glycated BSA (figure 36). Specifically, the fructosamine levels were 539, 551, 476, 592 and 526μmol L^{-1} in the presence of compounds **2**, **4**, **6**, **8** and **9** respectively, which were even lower than that in the presence of aminoguanidine (735μmol L^{-1}), while the fructosamine level in glycated BSA was 976μmol L^{-1}. This data is thus, in accordance with that obtained from the MALDI-TOF assay, which suggested that the title compounds inhibited Amadori product formation.
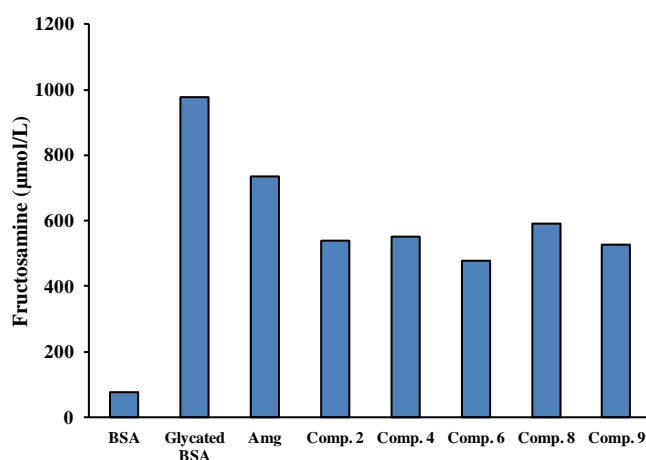
Figure 36: Fructosamine levels of BSA, glycated BSA and glycated BSA in presence of aminoguanidine or title compounds.

3.4.7 Adduct formation of title compounds with glucose

The title compounds were incubated in the presence of glucose and the reaction mixture was subjected to LC-MS analysis. All the compounds were found to form adducts with glucose. These are probably Schiff base type of adducts formed by the reaction of the free amino group in the title compounds with the aldehyde group of the sugar. The

formation of adduct with glucose could thus, be one of the ways by which the title compounds inhibit glycation. (figure 37)
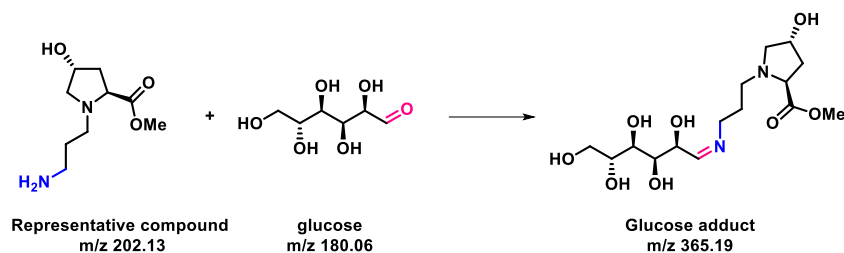
Figure 37: Adduct formation between representative title compound and glucose

3.4.8 Cytotoxicity assay

The cytotoxicity of the title compounds was evaluated in L6 rat muscle cells by measuring the cell viability using the standard MTT assay. It is a colourimetric assay that is based on the reduction of the tetrazolium dye, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide to a purple formazan derivative, that is achieved by enzymes present in viable cells. A decreased colour intensity is therefore indicative of low cell viability as a result of increased cytotoxicity. As seen in Figure 38, all the title compounds were non-toxic at almost all concentrations tested, except at 40mM concentration, where a slight decrease in cell viability was observed, this being more pronounced in compounds **6** and **8**. However, this concentration is double that used for the glycation inhibition studies described herein.
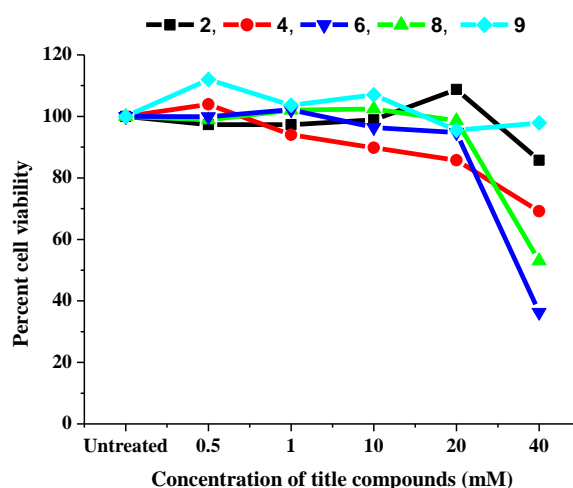
Figure 38: MTT assay depicting viability of L6 rat muscle cells in presence of title compounds.

3.4.9 Anti-oxidant properties of title compounds

AGEs, along with advanced lipoxidation products, have also been implicated in the generation of free radicals and reactive oxygen species (ROS), that are known to be involved in a variety of cellular processes ranging from apoptosis and necrosis to carcinogenesis.[16] It therefore follows that any candidate, that is capable of inhibiting AGE formation and additionally possess anti-oxidant properties, would have increased therapeutic value. Keeping this in mind, the title compounds were evaluated for their ability to reduce the concentration of intracellular ROS, using ascorbic acid as a control anti-oxidant, in a fluorescence-based assay. The anti-oxidant activity of a compound is evidenced by a decrease in the fluorescence intensity. Figure 39 shows a comparative study of the title compounds for anti-oxidant activity, wherein, it was found that compounds **2** and **4** exhibited higher anti-oxidant activity than other compounds, including ascorbic acid. More specifically, the percent fluorescence values were 57, 56, 83, 62 and 70 for compounds **2**, **4**, **6**, **8** and ascorbic acid respectively. Compound **9** did not show anti-oxidant activity as observed for other compounds.
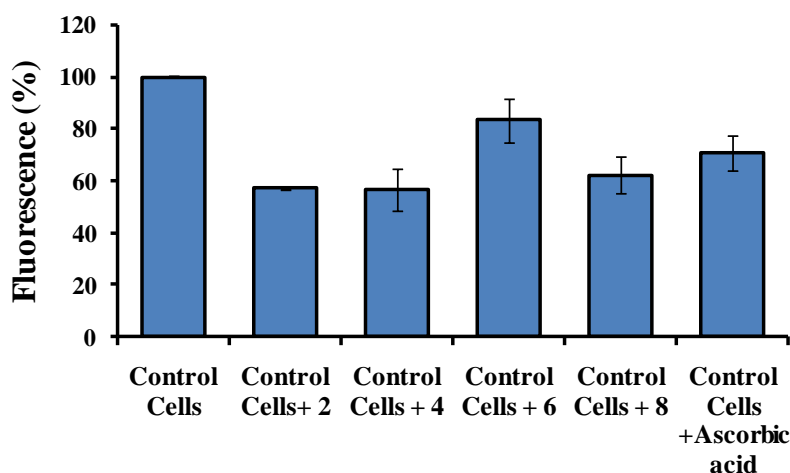
Figure 39: Anti-oxidant properties of title compounds in comparison to ascorbic acid.

3.5 Conclusion

The studies reported herein represent an important contribution towards the search for new molecules that not only inhibit glycation and AGE formation, but are also effective at controlling the concentration of intracellular reactive oxygen species. Moreover, the title compounds are easily accessible synthetically through simple chemical transformations, that should be amenable to scale-up. The superior antiglycation properties of the title compounds

have been demonstrated by circular dichroism, fluorescence spectroscopy, Western blot assay and also mass spectrometry. Mass spectrometric analysis and the fructosamine assays suggest that the title compounds form adducts with glucose, indicating that they probably act by inhibiting the formation of Amadori products. The low cytotoxicity of the title compounds is yet another favourable attribute. On the whole, the compounds-methyl (2*S*,4*R*)-1-(3-aminopropyl)-4-hydroxypyrrolidine-2-carboxylate (2) and (2*S*,4*R*)-1-(3-aminopropyl)-4-hydroxypyrrolidine-2-carboxylic acid (4) exhibited the best activity in the present study. These results suggest the contribution of the C4-hydroxyl group towards the antiglycation activity of the title compounds, in addition to the free amino function. The hydroxyl group offers possibilities of hydrogen-bonding and/or metal ion chelation, which could influence the activity of the respective title compound. Further, C4 being a chiral centre, the role of stereochemistry on the observed activity also cannot be ruled out.

3.6 Methods

3.6.1 General Information

All reagents were purchased from Sigma-Aldrich. DMF, CH_2Cl_2 (dichloromethane), were dried following standard procedures. DMF, CH_2Cl_2 were stored in 4 Å molecular sieves. Column chromatography was performed for purification of compounds on silica gel (100- 200 mesh or 60-120 mesh, Merck). TLCs were performed on Merck 5554 silica 60 pre-coated aluminium sheets. Compounds were visualized by spraying with Ninhydrin reagent and heating. ^1H and ^{13}C NMR spectra were obtained using Bruker AC-200, AC-400 or AC-500 NMR spectrometers. The chemical shifts are reported in delta (δ) values and referred to internal standard TMS for ^1H. ^1H NMR data are reported in the order of chemical shift, multiplicity (s, singlet; d, doublet; t, triplet; q, quartet; br, broad; br s, broad singlet; m, multiplet and/ or multiple resonance), number of protons. MALDI-TOF spectra were obtained from MALDI-TOF AB Sciex TOF/TOF 5800 spectrometer with CHCA (α-Cyano-4-hydroxycinnamic acid) as the matrix. Circular Dichroism (CD) analyses were performed on a JASCO J-815 spectrophotometer, and spectra were recorded at 25 °C in a 1 mm path length cuvette as accumulations of 3 scans. Specific rotations of samples were recorded on a Bellingham Stanley ADP220 Polarimeter, IR was recorded on Bruker Alpha spectrophotometer. Primary anti-AGE antibody, anti-CML antibody and secondary antibody conjugate were purchased from Merck Millipore (India).

3.6.2 Circular Dichroism (CD)

The CD spectra were measured as described earlier. The protein concentration was constant in all the samples (0.02 mg ml^{-1}). CD spectra were acquired on a JASCO J-815 spectropolarimeter at room temperature. The spectra were collected as accumulations of three scans and were corrected for respective blanks. Results are expressed as molar ellipticity, [θ] (deg cm^2 dmol^{-1}). The CD spectra of the protein samples were analysed to calculate the secondary structure content using CDPro software that has three algorithms: CONTINLL, CDSSTR and SELCON3.

3.6.3 *In vitro* glycation inhibition by fluorescence spectroscopy and IC$_{50}$ determination

The BSA glycation reaction was performed as described earlier by Kanska and Boratyński with minor modification described recently by Kolekar *et al* [148] with or without title compounds (20mM) and kept at 37°C for 10-15 days. *In vitro* glycation was monitored by fluorescence spectroscopy, excitation at 370nm and emission at 440nm by using a spectrofluorometer (Thermo, Varioskan Flash Multimode Reader).

The percent inhibition of AGE formation was calculated using the formula: %Inhibition = $(1 - F_i/F_c) \times 100$, where F_i = fluorescence intensity of glycated BSA treated with inhibitor and F_c = fluorescence intensity of glycated BSA in the absence of any inhibitor. The apparent IC$_{50}$ was determined by plotting the percent glycation inhibition Vs inhibitor concentration.

3.6.4 Western Blot Assay

BSA, glycated BSA or glycated BSA with 20mM of aminoguanidine or 20mM of compounds 2, 4, 6, 8, and were incubated at 37°C for 10-15 days. 5µg of each protein sample was separated by 12% SDS-PAGE and transferred onto a PVDF membrane. The membranes were blocked with 5% skimmed milk powder dissolved in PBS. The proteins were probed by anti-AGE and anti-CML antibodies at the antibody dilution of 1:2000 each. Secondary anti-rabbit antibody conjugated with HRP was used at a dilution of 1:5000 for both the blots. Immunoreactive bands were visualized using Western Bright Quantum western blotting detection kit (Advansta) and documented by Syngene Imaging System.

3.6.5 MALDI-TOF-MS based insulin glycation assay

The reaction mixture (100 µl) comprising phosphate buffer (pH 7.4, 0.1M) title compounds (20mM), insulin (1.8 mg/ml, 25µl) and glucose (250mM, 25µl) was incubated at 37°C. The reaction was monitored till the relative intensity of glycated insulin reached 100% as seen by MALDI-TOF-TOF analysis (AB SCIEX TOF/TOF™ 5800) (5 days). The reaction mixture was mixed with sinapinic acid (ratio 1:5) and analysed by MALDI-TOF-TOF in linear mode using Anchor Chip 384 targets as described.

3.6.6 Fructosamine Assay

The fructosamine level was measured by the NBT Labkit (Chemelex, S.A.) protocol. 300 µl of 0.75 mM NBT was added to a 96-well microplate containing 30 µl of 0.50µg BSA, glycated BSA, or glycated BSA with title compounds (20mM). Glycated BSA with aminoguanidine (20mM) was taken as a control. The reduction of NBT absorbance by fructosamine was measured at 520nm immediately after additions (considered as Abs1) and after incubation at 37°C for 15 min (considered as Abs2). The absorbance was monitored by using a UV spectrophometer (UV 1800, Shimadzu).

3.6.7 Adduct formation of title compounds with glucose

To elucidate the AGE inhibition mechanism of the title compounds, **2**, **4**, **6**, **8**, **9** (20mM each) were incubated with glucose (0.5M) in phosphate buffer pH 7.4 at 37°C for 6 days. The reaction was analysed by LC-MS on Q-Exactive Orbitrap to study and detect the glucose adduct formation with the title compounds.

3.6.8 Cytotoxicity assay

L6 rat muscle cells were seeded at a cell density of $1x10^4$ cells per well in a 96 well plate. After the cells adhered and attained their morphology, they were serum starved for 4h and treated with various concentrations of the title compounds (2, 4, 6, 8, and 9) in triplicate for 16h, while only serum starved cells served as control. After incubation, cells were given one wash with PBS and 100µL fresh serum free media was added. 6µL of 5mg/ml MTT (dissolved in PBS) was added to each well and incubated in dark at 37°C until violet formazan crystals were observed. Media from each well was discarded and crystals were dissolved in 100µL DMSO. Absorbance was measured at 555nm

using Varioskan flash multimode plate reader. This assay was performed using the Vybrant MTT Proliferation Assay Kit from Invitrogen.

3.6.9 Anti-oxidant properties of title compounds

Levels of intracellular ROS were measured in the presence of title compounds using ascorbic acid as control with 2', 7'-dichlorodihydrofluorescein diacetate (Molecular Probes). Cells were grown in SD medium with or without compounds for 3 days. Briefly, aliquots were taken after 3 days; cells were pelleted down and washed with PBS (137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, and 2mM KH_2PO_4). Cells were incubated with the dye (10µM) for 90min at 30°C. Cells were then washed twice in PBS and analyzed by flow cytometry analysis on BD Accuri C6 flow cytometer equipped with a 50mW argon laser emitting at 488nm. The green fluorescence was collected through a 488nm blocking filter. Data acquired from a minimum of 10,000 cells per sample at low flow rate were analyzed using BD accuri C6 software.

3.6.10 Synthesis of monomers

Methyl (2*S*,4*R*)-1-(3-((*tert*-butoxycarbonyl)amino)propyl)-4-hydroxypyrrolidine-2-carboxylate (1). Methyl (2*S*,4*R*)-4-hydroxypyrrolidine-2-carboxylate (1.00g, 6.8mmol)was dissolved in dry DMF; to it dry Et₃N (3.6mL, 20.0mmol) was added and stirred for 5 min at RT. Then to the above solution, 3-((*tert*-butoxycarbonyl)amino)propyl methanesulfonate (2.09g, 8.0mmol) was added drop-wise with continuous stirring. The reaction mixture was heated at 90°C for 10h. The reaction was monitored by TLC and after completion of the reaction, DMF was removed under vacuum and the compound was extracted from water with ethyl acetate. The organic layer was concentrated to get the crude compound, which was purified by column chromatography to get yellow coloured gum. (1.3g, 65%)

Specific rotation: [α]$_D$ = -48.31 (c= 0.135 in MeOH)

¹H NMR (200 MHz, CDCl₃) δ: 1.44 (s, 9H), 1.59-1.69 (m, 2H), 2.08-2.17 (m, 2H), 2.38-2.56 (m, 2H), 2.69-2.83 (m, 1H), 3.08 (br s, 1H), 3.19 (br m, 2H), 3.36-3.43 (dd, *J* = 5.4, 10.1Hz, 1H), 3.50-3.58 (t, *J* = 7.6Hz, 1H), 3.72 (s, 3H), 4.43-4.46 (m, 1H), 5.38 (br s, 1H) ppm.

¹³C NMR (50 MHz, CDCl₃) δ: 28.0, 28.4, 39.3, 51.8, 60.9, 64.4, 70.0, 78.9, 156.3, 174.4 ppm

13C DEPT (50 MHz, CDCl3) δ: 28.0 (*C*H$_2$), 28.4, 38.5 (*C*H$_2$), 39.3 (*C*H$_2$), 51.8 (*C*H$_2$), 51.8, 60.9 (*C*H$_2$), 64.4, 70.0 ppm.

HRMS: Calculated Mass for C$_{14}$H$_{26}$N$_2$O$_5$: 302.1842, Observed Mass (M+H): 303.1915, (M+Na): 325.1713.
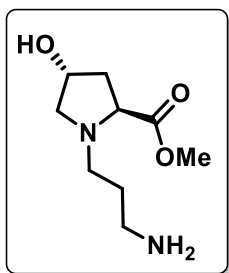
Methyl (2*S*,4*R*)-1-(3-aminopropyl)-4-hydroxypyrrolidine-2-carboxylate (2). To compound **1** (0.10g, 0.33 mmol), a 50% solution of TFA in DCM (5mL) was added with vigorous stirring at RT. After completion of the reaction, solvents were removed under vacuum and crude compound was purified by column chromatography. The resultant TFA salt of the compound was obtained in 62% yield (0.065g).

Specific rotation: [α]$_D$ = -24.47 (c= 0.11 in MeOH).

IR (neat)ν: 3352, 2946, 2833, 1749, 1683 cm^{-1}.

1H NMR (200 MHz, D$_2$O) δ: 2.09 - 2.21 (m, 2H), 2.33-2.41 (m, 1H), 2.49 - 2.60 (m, 1H), 3.05-3.13 (m, 2H), 3.28-3.41 (m, 2H), 3.54-3.60 (m, 1H), 3.86 (s, 3H), 3.91-3.99 (m, 1H), 4.66 (br m, 2H) ppm.

13C NMR (50 MHz, D$_2$O) δ: 24.6, 38.0, 38.8, 52.7, 53.2, 60.2, 64.8, 69.2, 175.5 ppm.

13C DEPT (50 MHz, D$_2$O) δ: 24.6 (*C*H$_2$), 38.0 (*C*H$_2$), 38.8 (*C*H$_2$), 52.7, 53.2 (*C*H$_2$), 60.2 (*C*H$_2$) 64.8, 69.2 ppm.

HRMS: Calculated Mass for C$_9$H$_{18}$N$_2$O$_3$: 202.1317, Observed Mass (M+H): 203.1392.

(2*S*,4*R*)-1-(3-((*tert*-Butoxycarbonyl)amino)propyl)-4-hydroxypyrrolidine-2-carboxylic acid (3). Compound **1** (0.2g, 0.66 mmol) was dissolved in methanol and to it, 2N LiOH (3mL) was added with vigorous stirring. The reaction was monitored by TLC and after completion of reaction, the reaction mixture was neutralised using Dowex H$^+$ resin; the resin was subsequently filtered off. The filtrate was concentrated under vacuum and the compound was purified by column chromatography (0.12g, 64%).

Specific rotation: [α]$_D$ =-35.06, (c=0.09 in MeOH).

¹H NMR (200 MHz, D₂O) δ: 1.39 (s, 9H), 1.82-1.86 (m, 2H), 2.11-2.51 (m, 2H), 3.16 (br m, 2H), 3.23 (br s, 1H), 3.31 (br. m, 2H), 3.85-3.93 (dd, J = 4.2, 13.0Hz, 1H), 4.14-4.23 (m, 1H), 4.59 (br s, 1H) ppm.

¹³C NMR (50 MHz, D₂O) δ: 28.3, 30.1, 40.5, 58.0, 64.2, 71.1, 71.8, 83.6, 160.7, 175.6 ppm.

¹³C DEPT (50 MHz, D₂O) δ: 28.3 (CH₂), 30.1, 39.3 (CH₂), 40.5 (CH₂), 58.0 (CH₂), 64.2 (CH₂) 71.1, 71.8 ppm.

HRMS: Calculated Mass for $C_{13}H_{24}N_2O_5$: 288.1685, Observed Mass (M+H): 289.1757.
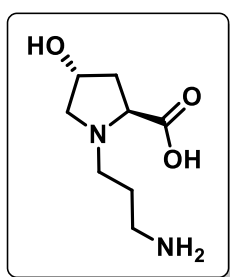
(2*S*,4*R*)-1-(3-Aminopropyl)-4-hydroxypyrrolidine-2-carboxylic acid (4). Compound **3** (0.30g, 1.04mmol) was dissolved in a 50% solution of TFA in DCM (5mL) with vigorous stirring at RT. The reaction was monitored by TLC and after the completion of reaction, solvents were removed under vacuum. Yield 0.19g (60%).

Specific rotation: $[\alpha]_D$ = -12.18 (c= 0.11 in MeOH).

IR (neat) ν: 3375, 2948, 2834, 1684 cm⁻¹.

¹H NMR (400 MHz, D₂O) δ: 2.12 - 2.13 (m, 2H), 2.25 - 2.32 (m, 1H), 2.52 - 2.56 (m, 1H), 3.06-3.10 (m, 2H), 3.27-3.38 (m, 2H), 3.52-3.54 (m, 1H), 3.94-3.97 (m, 1H), 4.51-4.55 (dd, J = 7.1, 11.3Hz, 1H), 4.64 (br s, 1H) ppm.

¹³C NMR (100 MHz, D₂O) δ: 23.4, 36.3, 38.0, 48.8, 54.5, 61.8, 69.3, 172.9 ppm.

¹³C DEPT (100 MHz, D₂O) δ: 23.4 (CH₂), 36.3 (CH₂), 38.0 (CH₂), 54.5 (CH₂), 61.8 (CH₂), 68.6, 69.3 ppm.

HRMS: Calculated Mass for $C_8H_{16}N_2O_3$: 188.1161, Observed Mass (M+H): 189.1235.

Methyl (2*S*,4*S*)-4-azido-1-(3-((*tert*-butoxycarbonyl)amino)propyl)pyrrolidine-2-carboxylate (5). To a solution of **1** (0.7g, 2.30mmol) in dry DCM, dry Et₃N (0.9mL, 6.90mmol) was added and stirred for 5min at RT. Ms-Cl (0.3mL, 0.003mol) was added dropwise at 0°C, with continuous stirring. After completion of the reaction, DCM was evaporated under vacuum and the crude mesylated compound was used for further reaction without further purification. The mesylated compound was dissolved in dry DMF, to it NaN₃ was added with continuous stirring (1.5g, 23.1mmol). The reaction was

heated for 10h at 75°C. After completion of the reaction, DMF was removed under vacuum and the compound was purified by column chromatography (0.58g, 67%).

Specific rotation: $[\alpha]_D$ = -0.0686 (c= 0.1 in $CHCl_3$).

^{1}H NMR (200 MHz, CDCl$_3$) δ: 1.44 (s, 9H), 1.65-1.68 (m,2H), 2.02-2.15 (br m, 1H), 2.43 - 2.53 (m, 3H), 2.78 - 2.88 (m, 1H), 3.15 - 3.25 (br m, 4H), 3.77 (s, 3H), 3.96 (br s, 1H), 5.47 (br s, 1H) ppm.

^{13}C NMR (50 MHz, CDCl$_3$) δ: 27.8, 28.4, 35.4, 38.4, 51.7, 52.1, 58.0, 58.7, 64.8, 78.6, 156.2, 173.3 ppm.

^{13}C DEPT (50 MHz, CDCl$_3$) δ: 27.7 (CH_2), 28.4, 35.4 (CH_2), 38.4 (CH_2), 51.7 (CH_2), 52.1, 58.0 (CH_2), 58.7, 64.8 ppm.

HRMS: Calculated Mass for $C_{14}H_{25}N_5O_4$: 327.1907, Observed Mass (M+H): 328.1979, (M+Na):350.1797.
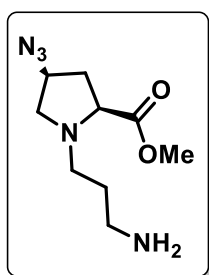
Methyl (2*S*,4*S*)-1-(3-aminopropyl)-4-azidopyrrolidine-2-carboxylate (6). To compound **5** (0.3g, 0.91mmol), a 50% solution of TFA in DCM (5mL) was added with vigorous stirring at RT. The reaction was monitored by TLC and after the completion of reaction, solvents were removed under vacuum and the product was purified by column chromatography (0.19g, 61%).

Specific rotation: $[\alpha]_D$ = -8.83 (c= 0.105 in MeOH).

IR (neat) *v*: 3376, 2188, 1745, 1678 cm^{-1}.

^{1}H NMR (200 MHz, D$_2$O) δ: 0.97-1.11 (m, 1H), 1.89-2.02 (m, 2H), 2.30-2.37 (m, 1H), 2.63-2.68 (m, 1H), 2.87-2.94 (m, 2H), 3.24-3.34 (m, 2H), 3.71 (s, 3H), 4.47-4.52 (m, 2H) ppm.

^{13}C NMR (50 MHz, D$_2$O) δ: 23.1, 33.7, 36.2, 52.9, 54.2, 58.6, 59.9, 65.7, 169.2 ppm.

^{13}C DEPT (50 MHz, D$_2$O) δ: 23.0 (CH_2), 33.6 (CH_2), 36.1 (CH_2), 52.8 (CH_2), 54.1, 58.5, 59.8 (CH_2), 65.7 ppm.

HRMS: Calculated Mass for $C_9H_{17}N_5O_2$: 227.1382, Observed Mass (M+H): 228.1457.
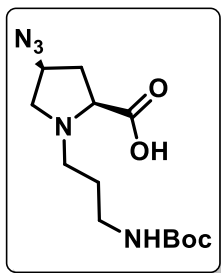
(2*S*,4*S*)-4-Azido-1-(3-((*tert*-butoxycarbonyl)amino)propyl) pyrrolidine-2-carboxylic acid (7). Compound **5** (0.2g, 0.61 mmol) was dissolved in methanol and to it 2N LiOH (3mL) was added with vigorous stirring. After completion of the reaction, the reaction mixture was neutralised using Dowex H$^+$ resin; the resin was subsequently filtered off. The filtrate was concentrated under vacuum and the compound was purified by column chromatography (0.125g, 65%).

Specific rotation: [α]$_D$ = -4.54 (c= 0.11 in MeOH).

^{1}H NMR (200 MHz, D$_2$O) δ: 1.42 (s, 9H), 1.87-1.89 (br m, 2H), 2.34-2.41 (m, 1H), 2.68 - 2.79 (m, 1H), 3.06 - 3.39 (br m, 6H), 3.76-3.82 (m, 1H), 4.08-4.12 (m, 1H), 4.57 (br s, 1H) ppm.

^{13}C NMR (50 MHz, D$_2$O) δ: 25.7, 27.6, 34.3, 53.5, 59.0, 59.6, 67.2, 81.1, 158.2, 173.0 ppm.

^{13}C DEPT (50 MHz, D$_2$O) δ: 25.7 (*CH$_2$*), 27.6, 34.3 (*CH$_2$*), 36.8 (*CH$_2$*), 53.5 (*CH$_2$*), 59.0, 59.6 (*CH$_2$*), 67.2 ppm.

HRMS: Calculated Mass for C$_{13}$H$_{23}$N$_5$O$_4$: 313.1750, Observed Mass (M+H): 314.1817, (M+Na):336.1635.
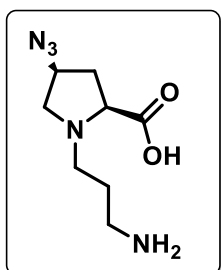
(2*S*,4*S*)-1-(3-Aminopropyl)-4-azidopyrrolidine-2-carboxylic acid (8). To compound **7** (0.3g, 0.95mmol), a 50% solution of TFA in DCM (5mL) was added with vigorous stirring at RT. The reaction was monitored by TLC and after the completion of reaction, solvents were removed under vacuum. The crude compound was used without further purification (0.132g, 63%).

**Specific rotation:** [α]$_D$ = -13.90 (c= 0.105 in MeOH).

IR (neat) ν: 3377, 2953, 2122, 1683 cm^{-1}.

^{1}H NMR (200 MHz, D$_2$O) δ: 2.08-2.16 (m, 2H), 2.38-2.46 (m, 1H), 2.70-2.85 (m, 1H), 3.04-3.11 (m, 2H), 3.37-3.41 (m, 2H), 3.80-3.86 (m, 1H), 4.25-4.29 (dd, *J* = 3.5, 10.8Hz, 1H), 4.56-4.59 (m, 1H) ppm.

^{13}C NMR (50 MHz, D$_2$O) δ: 23.3, 34.1, 36.4, 52.8, 58.8, 60.0, 66.4, 162.7, 171 ppm.

^{13}C DEPT (50 MHz, D$_2$O) δ: 23.3 (*CH$_2$*), 34.1 (*CH$_2$*), 36.3 (*CH$_2$*), 52.8 (*CH$_2$*), 58.8, 60.0 (*CH$_2$*), 66.4 ppm.

HRMS: Calculated Mass for C$_8$H$_{15}$N$_5$O$_2$: 213.1226, Observed Mass (M+H): 214.1299.
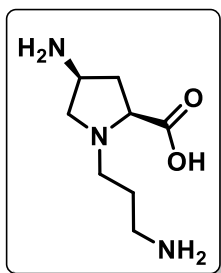
(2*S*,4*S*)-4-Amino-1-(3-aminopropyl)pyrrolidine-2-carboxylic acid (9). Compound **8** (0.2g, 0.61 mmol) was dissolved in methanol, to it Palladium on charcoal (0.02g) was added and the reaction was allowed to proceed under H_2 pressure (35-40psi). After completion of reaction, the solvent was removed under vacuum and the crude compound (0.112g, 61%) was purified by HPLC.

Specific rotation: $[\alpha]_D$ = -13.9 (c= 0.1 in MeOH).

¹H NMR (400 MHz, D₂O) δ: 1.99-2.23 (m, 3H), 2.95-3.09 (br. m, 3H), 3.2-3.35 (m, 1H), 3.43-3.58 (m, 1H), 3.7-3.81 (m, 1H), 3.91-4.0 (m, 1H), 4.24-4.34 (m, 2H) ppm. **¹³C NMR (100 MHz, D₂O) δ:** 23.2, 32.7, 36.2, 47.2, 52.7, 55.8, 67.6, 163, 170.1 ppm. **¹³C DEPT (50 MHz, D₂O) δ:** 23.2 (*C*H₂), 32.7 (*C*H₂), 36.2 (*C*H₂), 47.2, 52.7 (*C*H₂), 55.8 (*C*H₂), 67.6 ppm.

HRMS: Calculated Mass for $C_8H_{17}N_3O_2$: 187.1321, Observed Mass (M+H): 188.1392.

Appendix

Characterization	Page number
Compound 1 ^{1}H NMR, ^{13}C NMR	68
Compound 1 DEPT, HRMS	69
Compound 2 ^{1}H NMR, ^{13}C NMR	70
Compound 2 DEPT, HRMS	71
Compound 3 ^{1}H NMR, ^{13}C NMR	72
Compound 3 DEPT, HRMS	73
Compound 4 ^{1}H NMR, ^{13}C NMR	74
Compound 4 DEPT, HRMS	75
Compound 5 ^{1}H NMR, ^{13}C NMR	76
Compound 5 DEPT, HRMS	77
Compound 6 ^{1}H NMR, ^{13}C NMR	78
Compound 6 DEPT, HRMS	79
Compound 7 ^{1}H NMR, ^{13}C NMR	80
Compound 7 DEPT, HRMS	81
Compound 8 ^{1}H NMR, ^{13}C NMR	82
Compound 8 DEPT, HRMS	83
Compound 9 ^{1}H NMR, ^{13}C NMR	84
Compound 9 DEPT, HRMS	85
HRMS- Adduct formation of compound 2 with glucose	86
HRMS- Adduct formation of compound 4 with glucose	86
HRMS- Adduct formation of compound 6 with glucose	87
HRMS- Adduct formation of compound 8 with glucose	87
HRMS- Adduct formation of compound 9 with glucose	88

¹H NMR Compound -1

¹³C NMR Compound 1

DEPT Compound 1

H-4 #78 RT: 0.35 AV: 1 NL: 3.90E9
T: FTMS + p ESI Full ms [86.00-1290.00]

303.1915
R=65307
$C_{14} H_{27} O_5 N_2 = 303.1914$
0.0623 ppm

HRMS Compound -1
Calculated Mass: 302.1842
Observed Mass (M+H): 303.1915
(M+Na): 325.1713

325.1731
R=62307
$C_{14} H_{26} O_5 N_2 Na = 325.1734$
-0.8207 ppm

317.2068
R=62407

331.2845
R=51000

339.1901
R=48200

m/z

¹H NMR Compound 2

¹³C NMR Compound 2

DEPT Compound 2

69.26
64.85
52.77
60.23
53.25
38.06
38.85
24.61

Chemical Shift (ppm)

190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0

G-H8 #496 RT: 2.21 AV: 1 NL: 6.16E9
T: FTMS + p ESI Full ms [100.00-1500.00]

203.1392
R=79407
C$_9$ H$_{19}$ O$_3$ N$_2$ = 203.1390
0.9955 ppm

HRMS Compound -2
Calculated Mass: 202.1317
Observed Mass (M+H): 203.1392

259.2017
R=70307

429.2025
R=51706

617.3183
R=42406

714.1237
R=35500

835.3708
R=36700

932.8889
R=23400

1061.4417
R=22400

1173.5345
R=20500

1271.5421
R=20400

100 200 300 400 500 600 700 800 900 1000 1100 1200 1300 1400 1500

m/z

¹H NMR Compound 3

¹³C NMR Compound 3

30.14

DEPT Compound 3

71.88
71.13

58.01

64.20

39.32

40.56

28.31

190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
Chemical Shift (ppm)

H-6 #96 RT: 0.43 AV: 1 NL: 6.01E9
T: FTMS + p ESI Full ms [85.40-1281.00]

289.1757
R=66307
$C_{13} H_{25} O_5 N_2$ = 289.1758
-0.4813 ppm

233.1130
R=73707
$C_9 H_{17} O_5 N_2$ = 233.1132
-0.7561 ppm

HRMS Compound -3
Calculated Mass: 288.1685
Observed Mass (M+H): 289.1757

347.1979
R=59507

373.1277
R=52506

391.2047
R=50202

417.2016
R=47200

452.2334
R=46500

100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0

200 220 240 260 280 300 320 340 360 380 400 420 440 460 480 500
m/z

¹H NMR Compound 4

DEUTERIUM OXIDE

4.79

4.64
4.55
4.51
3.97
3.94
3.54
3.52
3.38
3.31
3.27
3.10
3.06
2.56
2.52
2.32
2.13
2.12

1.00 0.92 0.90 0.92 1.89 1.83 0.91 0.93 1.96

Chemical Shift (ppm)

¹³C NMR Compound 4

172.99

69.36
61.89
54.55
48.85
38.04
36.39
23.42

Chemical Shift (ppm)

DEPT Compound 4

69.35
68.67
61.89
54.55
38.03
36.37
23.41

Chemical Shift (ppm)

G-H9 #483   RT: 2.15   AV: 1   NL: 1.66E8
T: FTMS + p ESI Full ms [100.00-1500.00]

189.1235
R=82007
$C_8 H_{17} O_3 N_2$ = 189.1234
0.6370 ppm

HRMS Compound -4
Calculated Mass: 188.1161
Observed Mass (M+H): 189.1235

143.0017
R=93502
$C_8 H O_2 N$ = 143.0002
10.4799 ppm

214.9177
R=76802

273.0758
R=68307

246.9439
R=70702

305.1670
R=62807

331.0976
R=60507

357.0281
R=59007

409.0420
R=54307

447.2779
R=53207

486.9868
R=49502

513.0699
R=48107

539.0004
R=46106

m/z

¹H NMR Compound 5

¹³C NMR Compound 5

DEPT Compound 5

64.81
58.74
52.15
28.43

51.79
58.03
38.49
35.42
27.79

190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
Chemical Shift (ppm)

H-5 #88 RT: 0.39 AV: 1 NL: 1.55E9
T: FTMS + p ESI Full ms [86.00-1290.00]

328.1979
R=62107
C 14 H 26 O 4 N 5 = 328.1979
-0.1670 ppm

HRMS Compound -5
Calculated Mass: 327.1907
Observed Mass (M+H): 328.1979
                (M+Na):350.1797

350.1797
R=60307
C 14 H 25 O 4 N 5 Na = 350.1799
-0.6239 ppm

342.2134
R=61307

313.2739
R=61800

321.1576
R=61206

345.1367
R=39700

353.2658
R=57207

359.3158
R=48600

364.1952
R=57807

310   315   320   325   330   335   340   345   350   355   360   365   370   375
m/z

¹H NMR Compound 6

¹³C NMR Compound 6

DEPT Compound 6

H-6 #96 RT: 0.43 AV: 1 NL: 6.01E9
T: FTMS + p ESI Full ms [85.40-1281.00]

HRMS Compound -6
Calculated Mass: 288.1685
Observed Mass (M+H): 289.1757

¹H NMR Compound 7

¹³C NMR Compound 7

DEPT Compound 7

H-10 #93   RT: 0.42   AV: 1   NL: 3.63E9
T: FTMS + p ESI Full ms [85.40-1281.00]

HRMS Compound -7
Calculated Mass: 313.1750
Observed Mass (M+H): 314.1817
              (M+Na):336.1635

¹H NMR Compound 8

¹³C NMR Compound 8

DEPT Compound 8

98.85
66.40

60.01
52.82
36.39
34.13
23.33

190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
Chemical Shift (ppm)
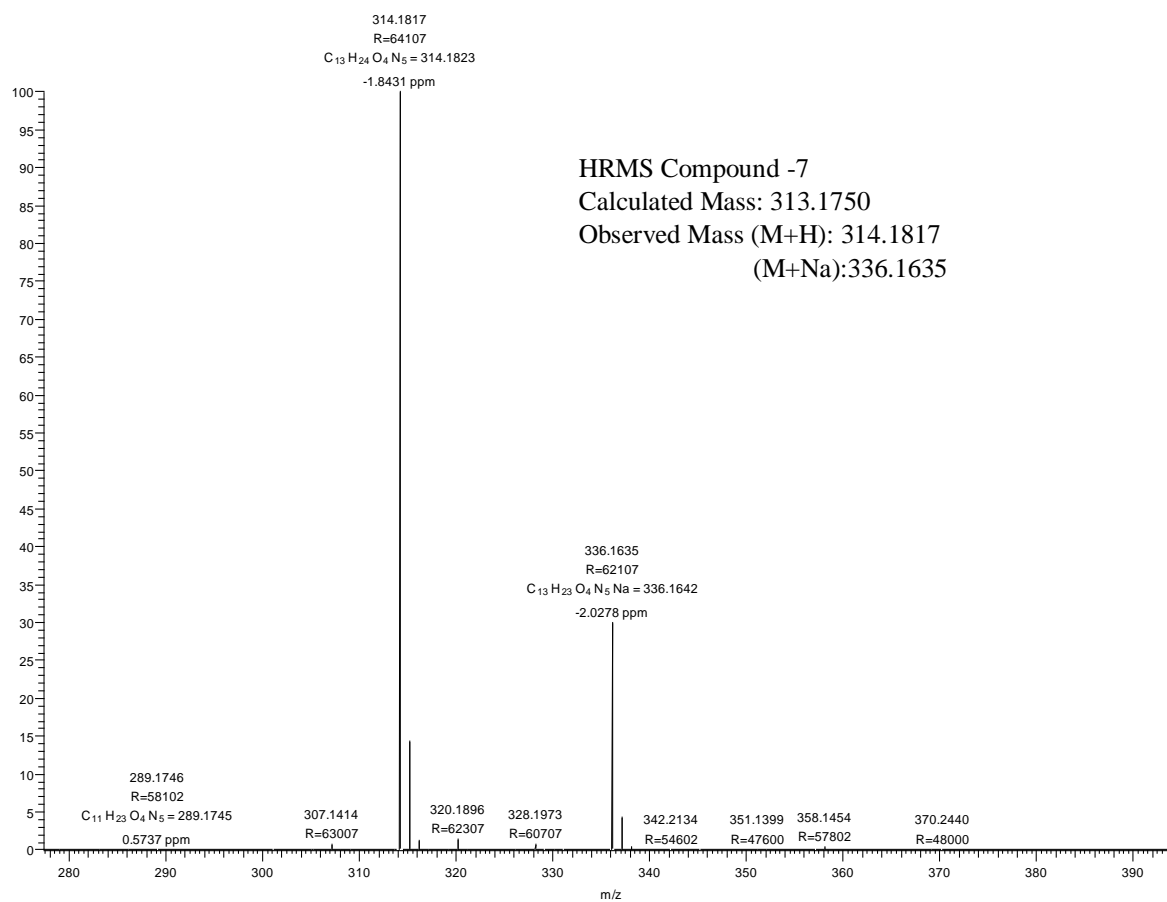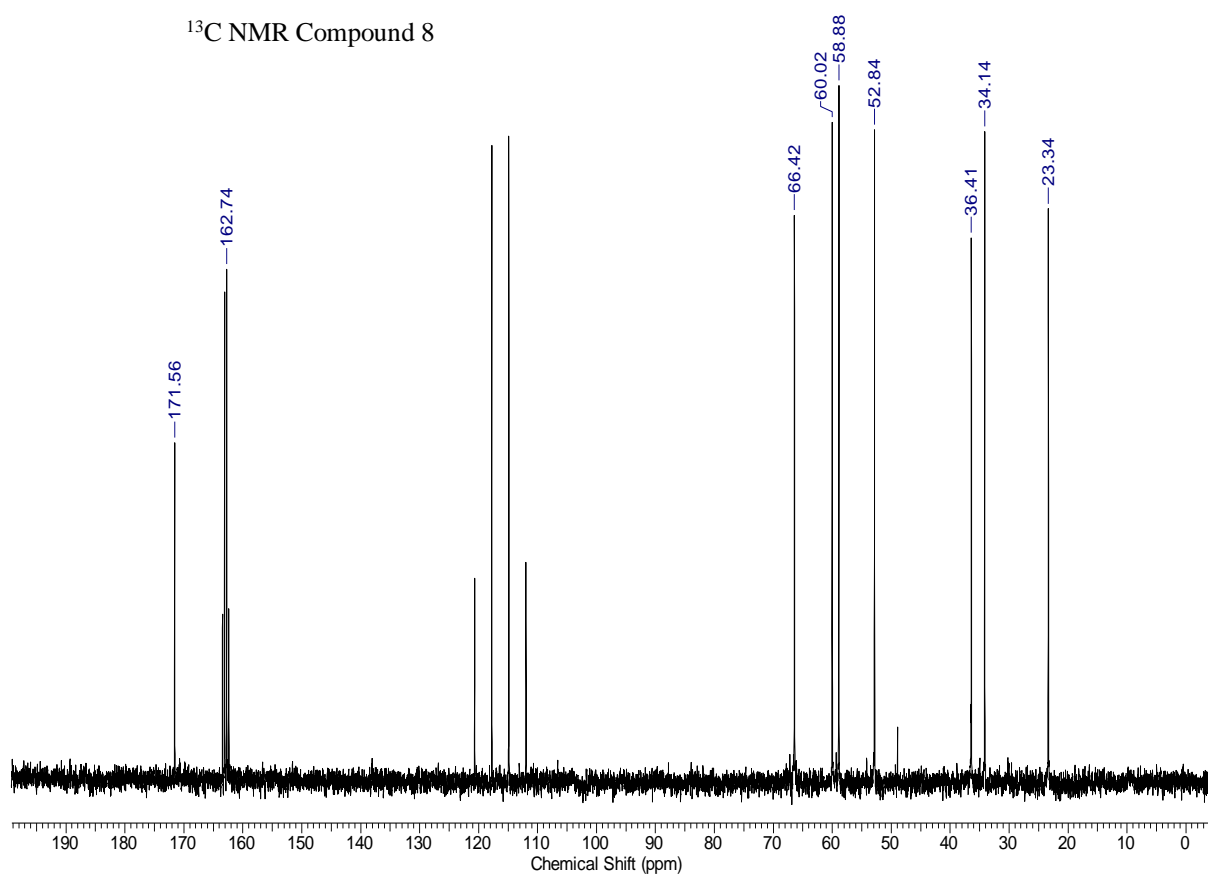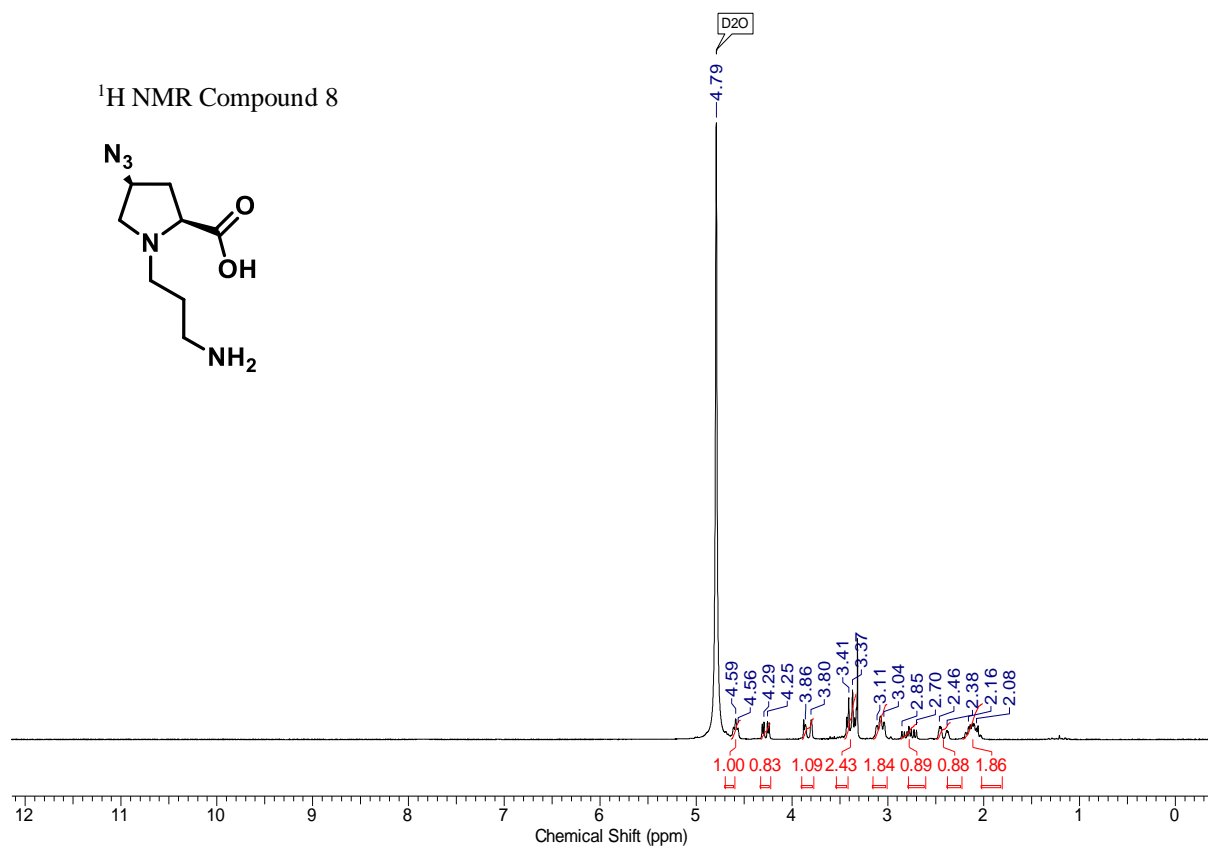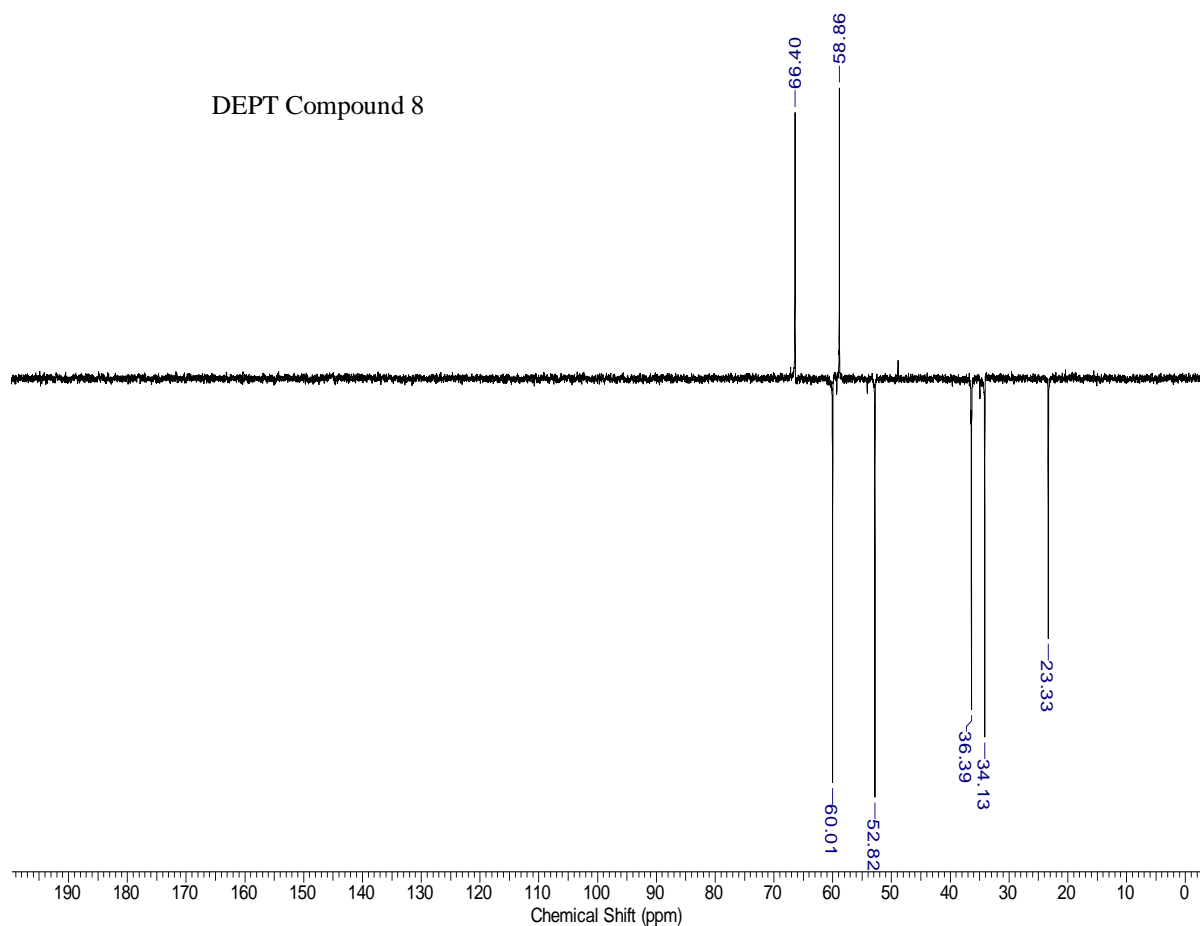
G-H11 #534 RT: 2.38 AV: 1 NL: 2.46E9
T: FTMS + p ESI Full ms [100.00-1500.00]

214.1299
R=77907
$C_8 H_{16} O_2 N_5$ = 214.1299
0.2898 ppm

HRMS Compound -8
Calculated Mass: 213.1226
Observed Mass (M+H): 214.1299

339.5582
R=47600

439.2688
R=53307

573.7489
R=31400

658.3998
R=43307

758.3241
R=24400

883.5390
R=37607

974.1414
R=25800

1118.6439
R=32907

1226.5938
R=19500

1337.7748
R=26906

100 200 300 400 500 600 700 800 900 1000 1100 1200 1300 1400 1500
m/z

¹H NMR Compound 9

¹³C NMR Compound 9

DEPT Compound 9

67.69

47.24

36.23

23.22

55.82
52.75

32.77

Chemical Shift (ppm)

H-13_150326135800 #79 RT: 0.35 AV: 1 NL: 2.09E8
T: FTMS + p ESI Full ms [85.40-1281.00]

188.1392
R=82507
$C_8 H_{18} O_2 N_3 = 188.1394$
-0.9457 ppm

HRMS Compound -9
Calculated Mass: 187.1321
Observed Mass (M+H): 188.1392

202.1548
R=80207

142.1338
R=95507
$C_7 H_{16} N_3 = 142.1339$
-0.5982 ppm

216.1292
R=76607

244.2016
R=71807

285.1668
R=66007

327.0074
R=59902

353.2660
R=54107

387.2872
R=53907

413.2176
R=48400

451.2547
R=36100

475.4142
R=47507

511.3108
R=34200

m/z
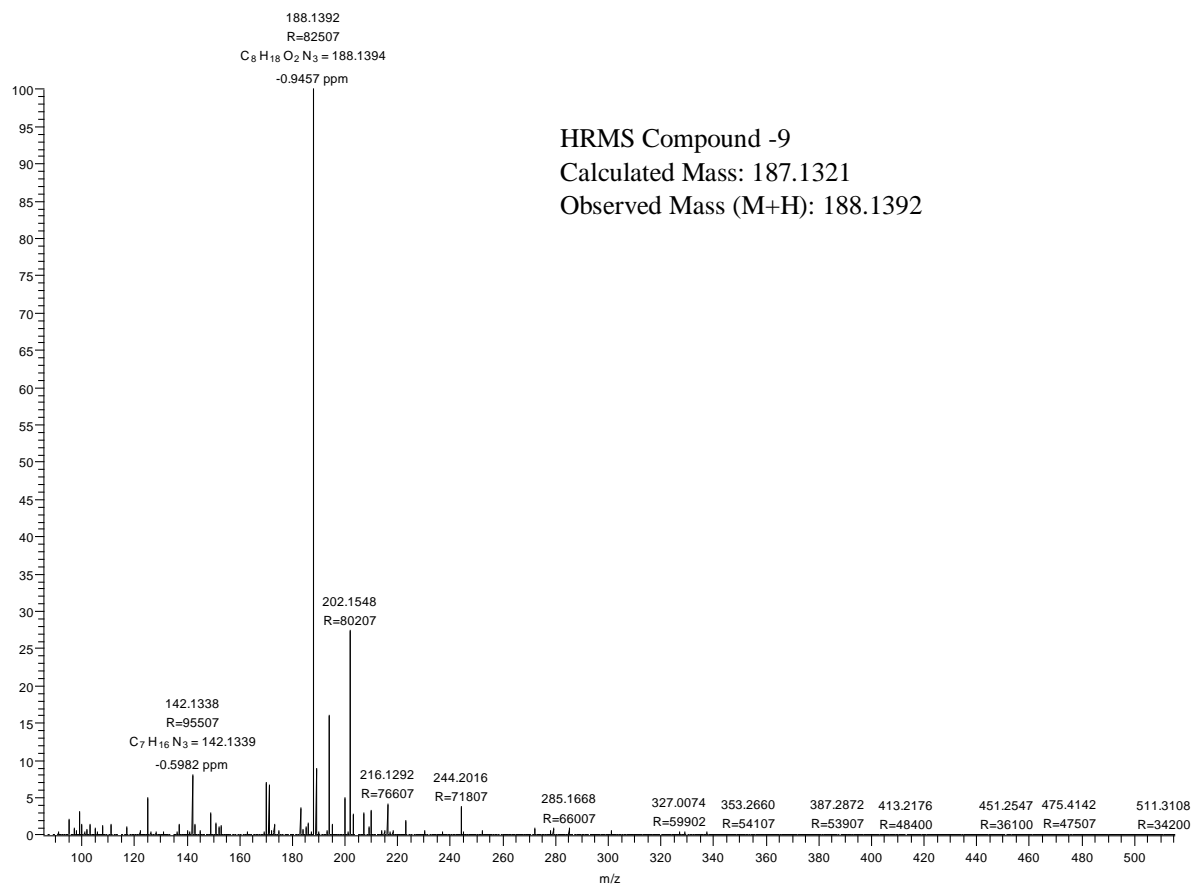
Compound 2 – glucose adduct, *m/z* 365.19

(B)

HARSHA-8 #74 RT: 0.33 AV: 1 NL: 1.10E6
T: FTMS + p ESI Full ms [86.00-1290.00]

365.1920
R=58102
$C_{15} H_{29} O_8 N_2 = 365.1918$
0.4774 ppm
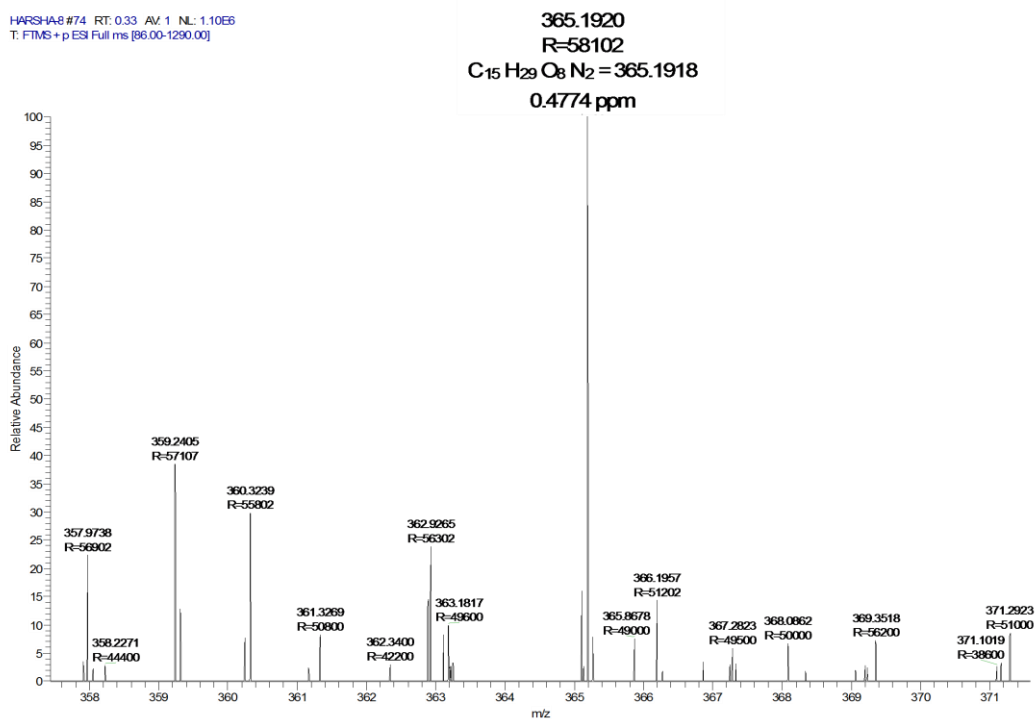
Compound 4 – glucose adduct, *m/z* 351.1765

HARSHA-9 #80 RT: 0.36 AV: 1 NL: 8.67E6
T: FTMS + p ESI Full ms [86.00-1290.00]

351.1765
R=60702
$C_{14} H_{27} O_8 N_2 = 351.1762$
0.8285 ppm

373.1585
R=57807
$C_{14} H_{26} O_8 N_2 Na = 373.1581$
0.9953 ppm
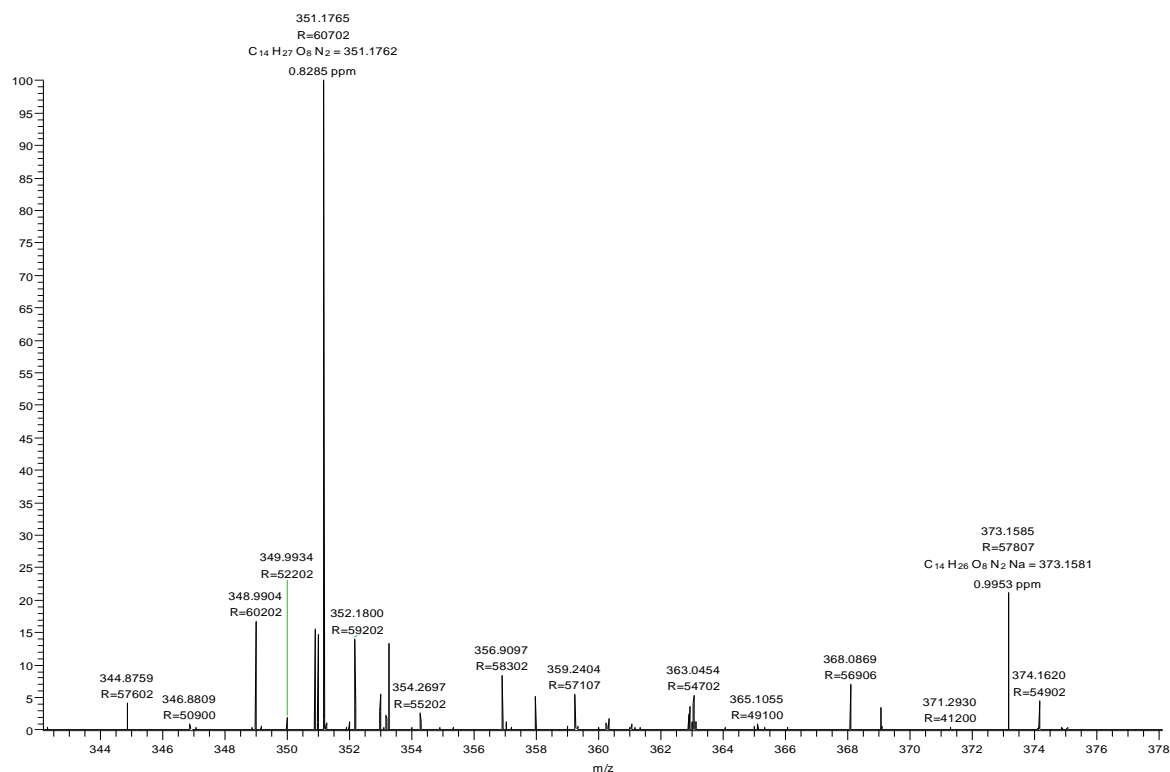
Compound 6- glucose adduct, *m/z* 390.198
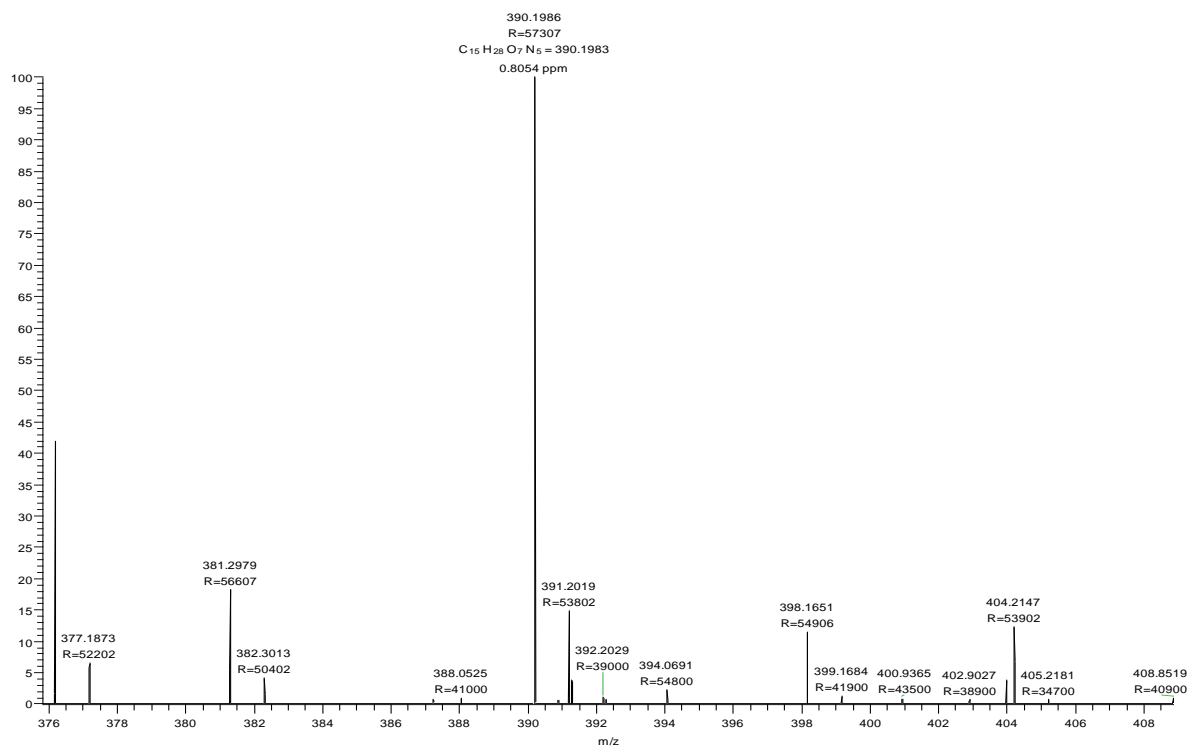
Compound 8 – glucose adduct, *m/z* 378.1832
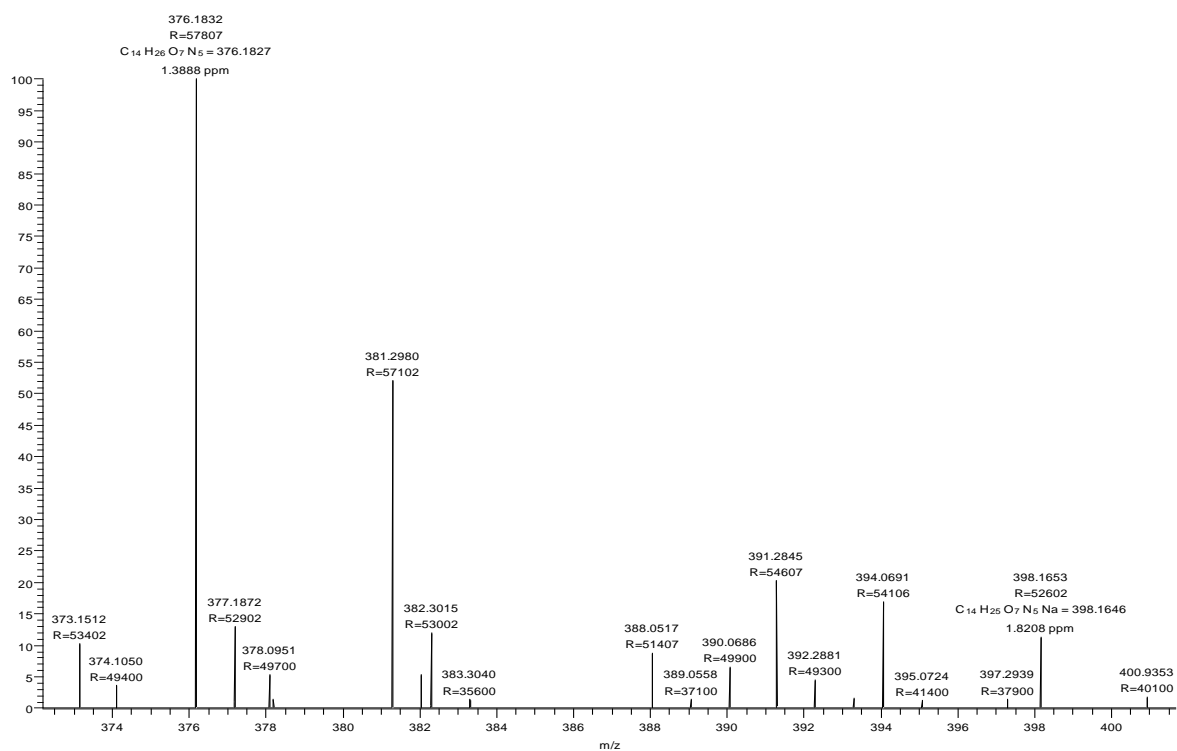
Compound 9- glucose adduct, *m/z* **350.1925**

HARSHA-13 #80 RT: 0.36 AV: 1 NL: 2.15E6
T: FTMS + p ESI Full ms [86.00-1290.00]

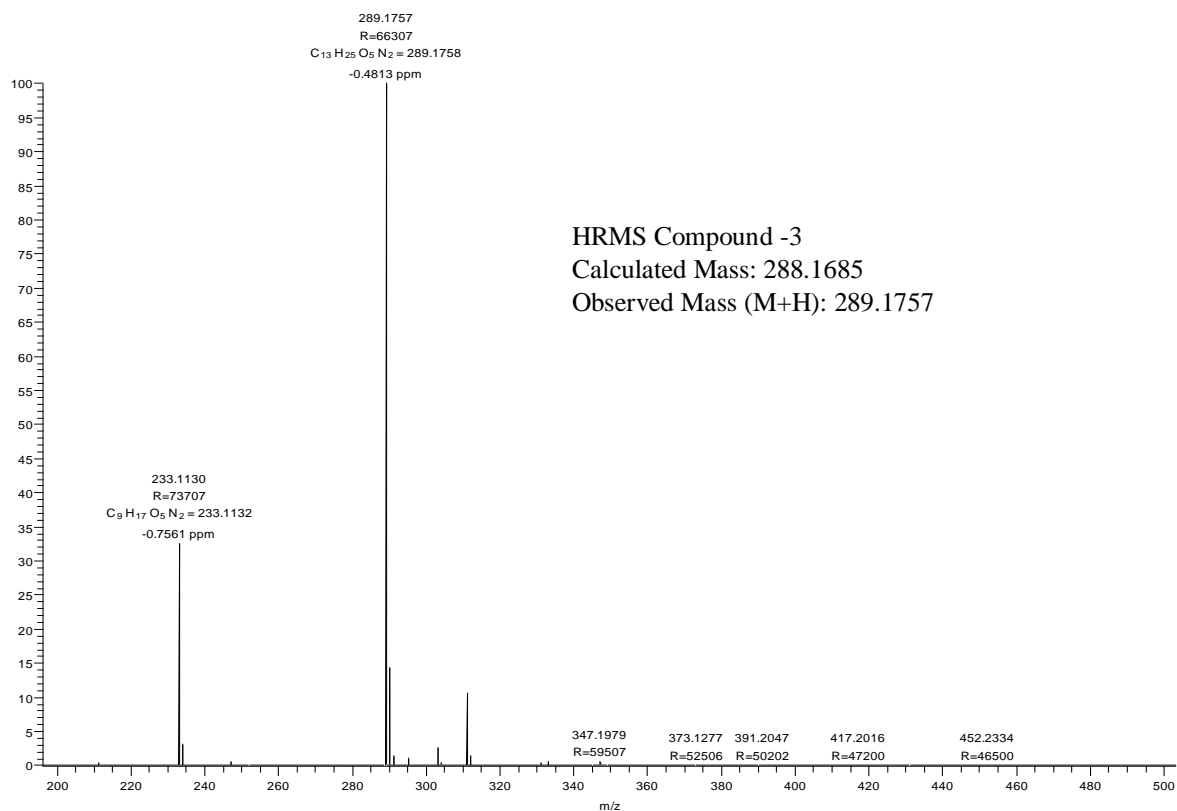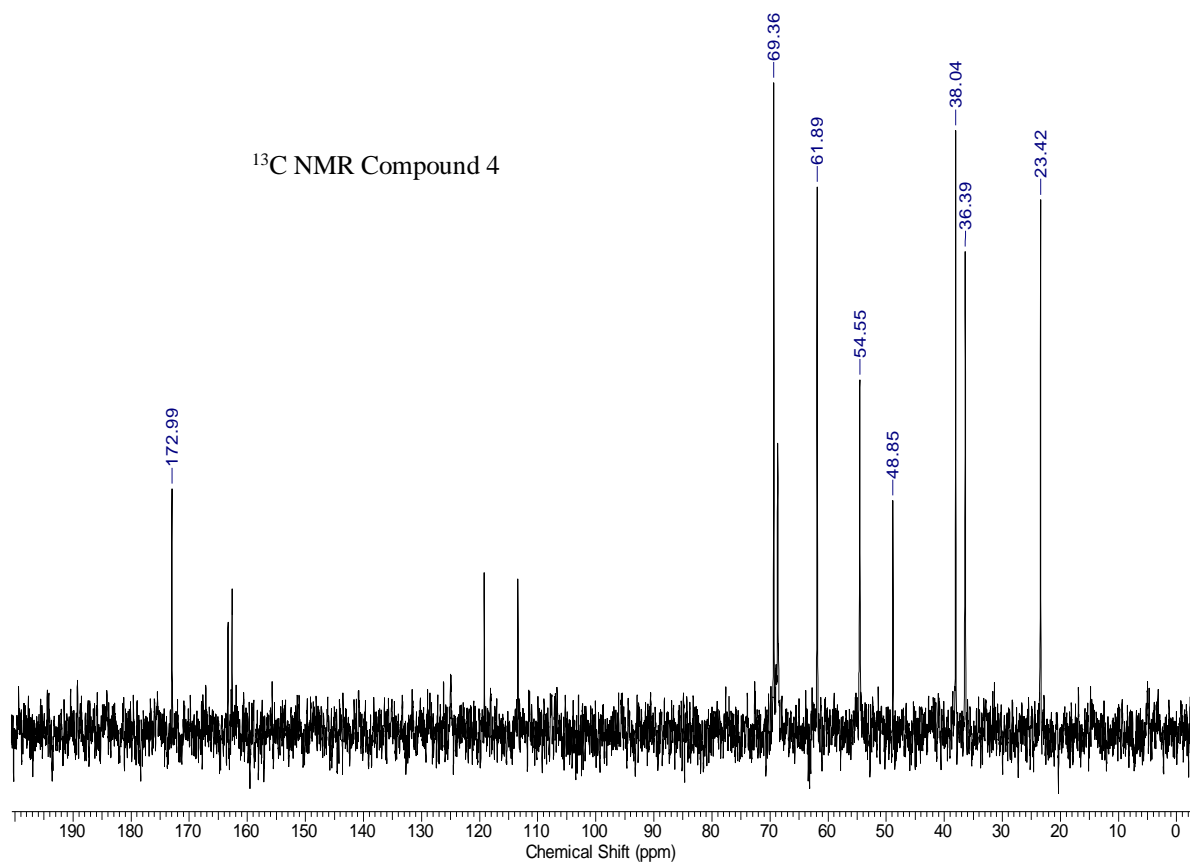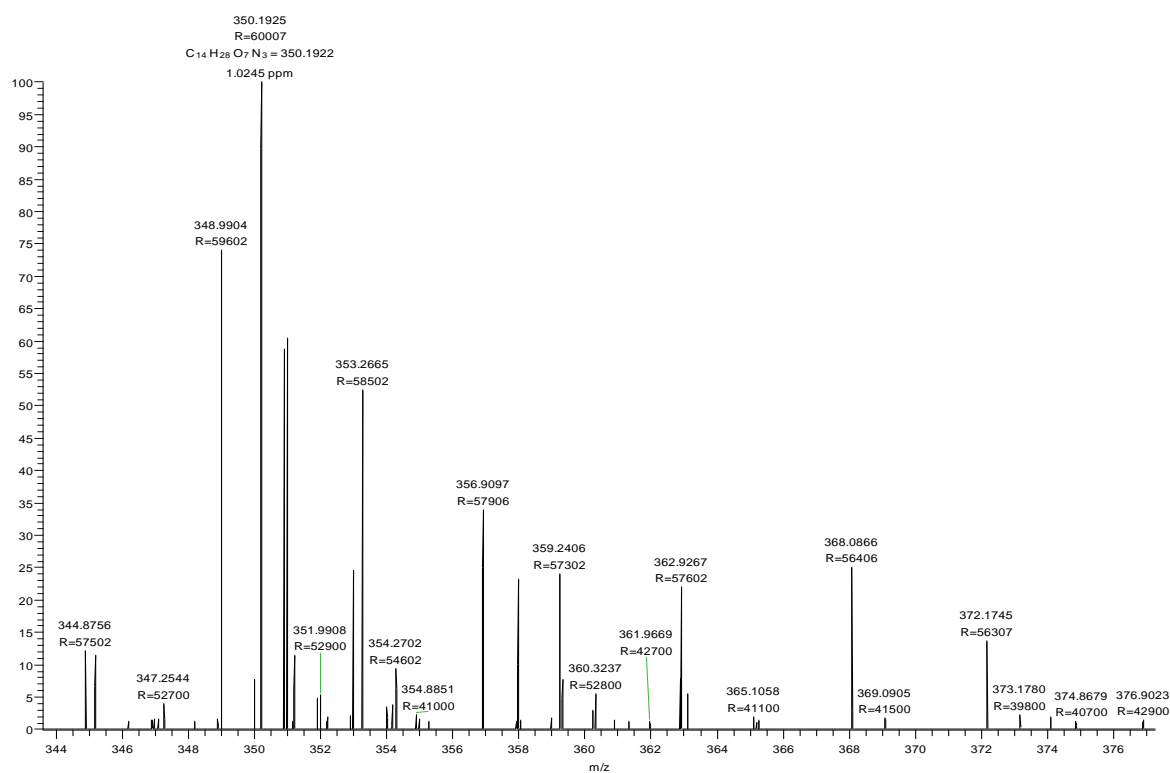# Chapter 4

Design, synthesis and evaluation of antifolates derived from Guanine

4.0 Introduction

Folic acid (FA) is a completely oxidised molecule composed of a pteridine ring, a *para*-aminobenzoic acid and an *L*-glutamic acid.[150,151] Compounds sharing a common chemical backbone are known as folates.[150] Vitamin B9, a folate is thought to be one of the 13 essential vitamins that helps in DNA replication and is a substrate for various enzymatic reactions involving amino acid synthesis and vitamin metabolism. This vitamin is not synthesised in the body and therefore must be obtained from diet or supplements.[152] Although folic acid is involved in important functions of the body, it is not metabolically active and must be reduced to participate in cell functions.[152]
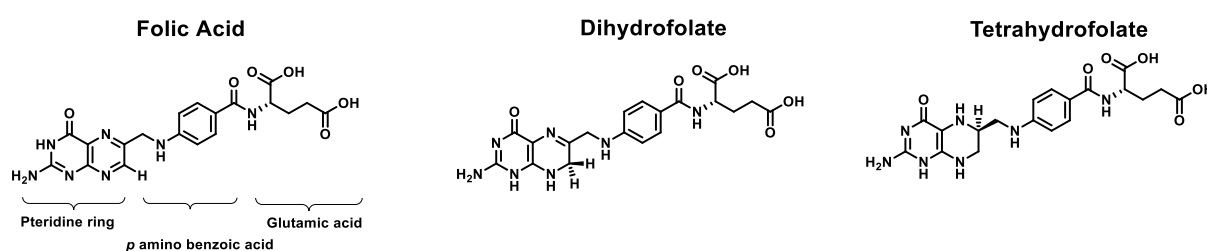
Figure 40: Chemical structures of folic acid, dihydrofolate and tetrahydrofolate

Whereas natural folates are present in the reduced form such as di-hydro- or tetrahydrofolate and may carry a carbon moiety (i.e., methyl, methylene, methenyl or formyl).[151] The most reduced and stable form, tetrahydrofolate is the only enzymatically active form which is generated in the folate pathway and acts as a C1 donor to generate deoxythymidine monophosphate, purine nucleotides, and the amino acids methionine and histidine.[53]Thus DNA replication would be halted if the folate pathway is inhibited. Therefore antifolate drugs act as inhibitors to the folate pathway and have been targeted to rapidly growing cancer cells, proliferating bacterial and protozoal pathogens.[61] 3 major pathways are involved in the folate transport namely folate receptor alpha (FRα), reduced folate carrier (RFC) and proton-coupled folate transporter (PCFT). FRα is present on the cell membrane and binds to and transports folates with high affinity *via* receptor mediated endocytosis, whereas PCFT has a lower affinity for folates compared to FRα. RFC has a comparable affinity as PCFT for reduced folate uptake.[153]

4.1 Antifolates

Structurally similar to natural folates, antifolates are a class of antimetabolites that compete with folates and bind to enzymes of the folate cycle thus halting nucleotide biosynthesis.[154] The search of new antifolates as antimetabolite agents continues to be an important area of research.

[60] One of the primary reasons that targeting the folate pathway has been successful for such a wide range of applications is the conservation and essentiality of the pathway across all forms of life. [155]Although there involves 50 years of study in this field, developing better antifolates still remains a challenge[156] directly inhibiting nucleotide synthesis involves binding to enzymes while indirect inhibition takes place by blocking folate cycle.[154] Although halting folate biosynthesis by antifolates is a potential option, inhibiting conserved enzymes has led to most successful drug candidates.[157] Additionally, incorporation into nucleic acids such as DNA and RNA is also a prevalent mechanism of antimetabolites.[158] Molecules with similarity in structure to folic acid are known as "classical antifolates" which compete with enzymes in the folate cycle thereby inhibiting folate metabolism while structurally different molecules known as "non classical antifolates" which act by inhibiting folic acid biosynthesis.[157]

(a) Methotrexate   **(b) Sulfacetamide**   **(c) Pemetrexed**
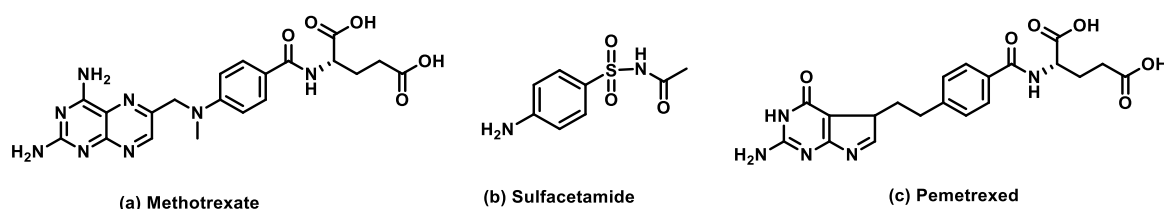
Figure 41: Representative chemical structures of antifolates: Methotrexate (a), Sulfacetamide (b) and Pemetrexed (c)

Since numerous enzymes are involved in the folate metabolism, the mechanism of antifolates depends on the enzyme they act upon. [157]Targeting important enzymes such as serine hydroxymethyltransferase, dihydrofolate reductase and thymidylate synthase have been exploited to treat cancer, malaria and various bacterial infections.[159]

4.2 Rationale

Many folate conjugated drugs and toxin, anti- FRα antibodies, antifolates and folate based imaging agents have been widely administered due to the dependency of many tumours.[160] Even though there have been decades of research devoted to the discovery of antifolates, there are only approximately eight compounds that are clinically used across all indications.[60] There are basically four categories of antimetabolites: antifolates, purine analogues, pyrimidine analogues and sugar-modified analogues.[158] Crystal structure studies reveal that the pteroate moiety of folic acid is buried inside the receptor, whereas its glutamate moiety sticks out of the pocket entrance, allowing it to be conjugated to desired molecules without adversely affecting FRα binding.[160] As seen in figure 42, *N*5 atom of folic acid forms one hydrogen bond with the H135 side chain of the receptor with a distance of 2.9Å. It can be anticipated that the presence

of a hydrogen bond between the side chain of H135 with an oxygen atom instead of the original *N*5 nitrogen atom would lead to a stronger hydrogen bonding between the ligand and the receptor owing to reduced hydrogen bonding distance and the possibility of an oxygen atom to form an additional hydrogen bond with arginine, R103 side chain. Thus, design of new folate receptor ligands was carried out by employing simple chemical transformations on commercially available guanine, as it structurally resembles pterin moiety and contains sites for necessary modifications required for the receptor binding while the glutamic acid part of the native folic acid was kept undisturbed.
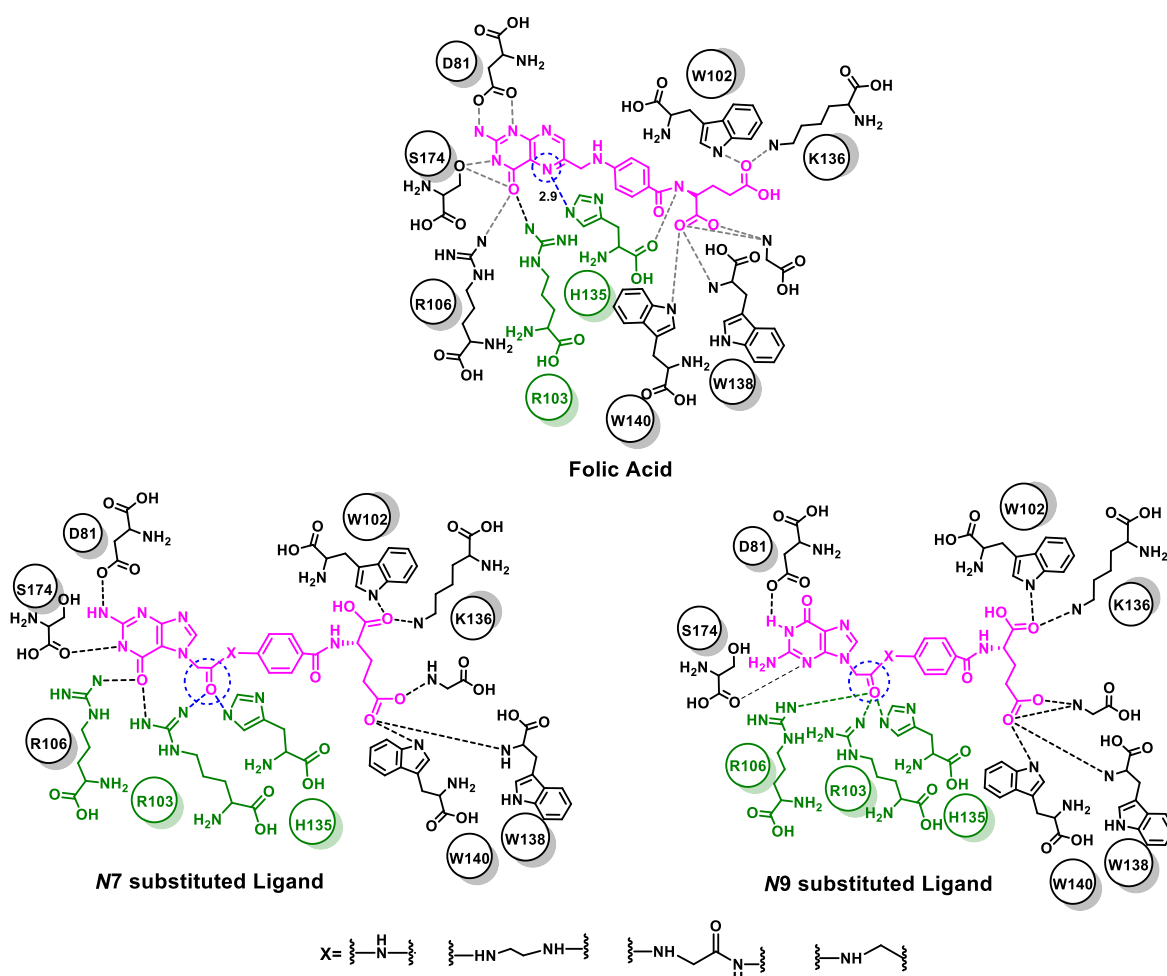
Figure 42: Folic acid with ligand binding pocket residues[160] **and designed *N*7 and *N*9 substituted ligands**

4.3 Results and Discussion

4.3.1 Synthesis of designed antifolates

A series of antifolate ligands were designed (Figure 43) keeping in mind the rationale described in Section 4.3
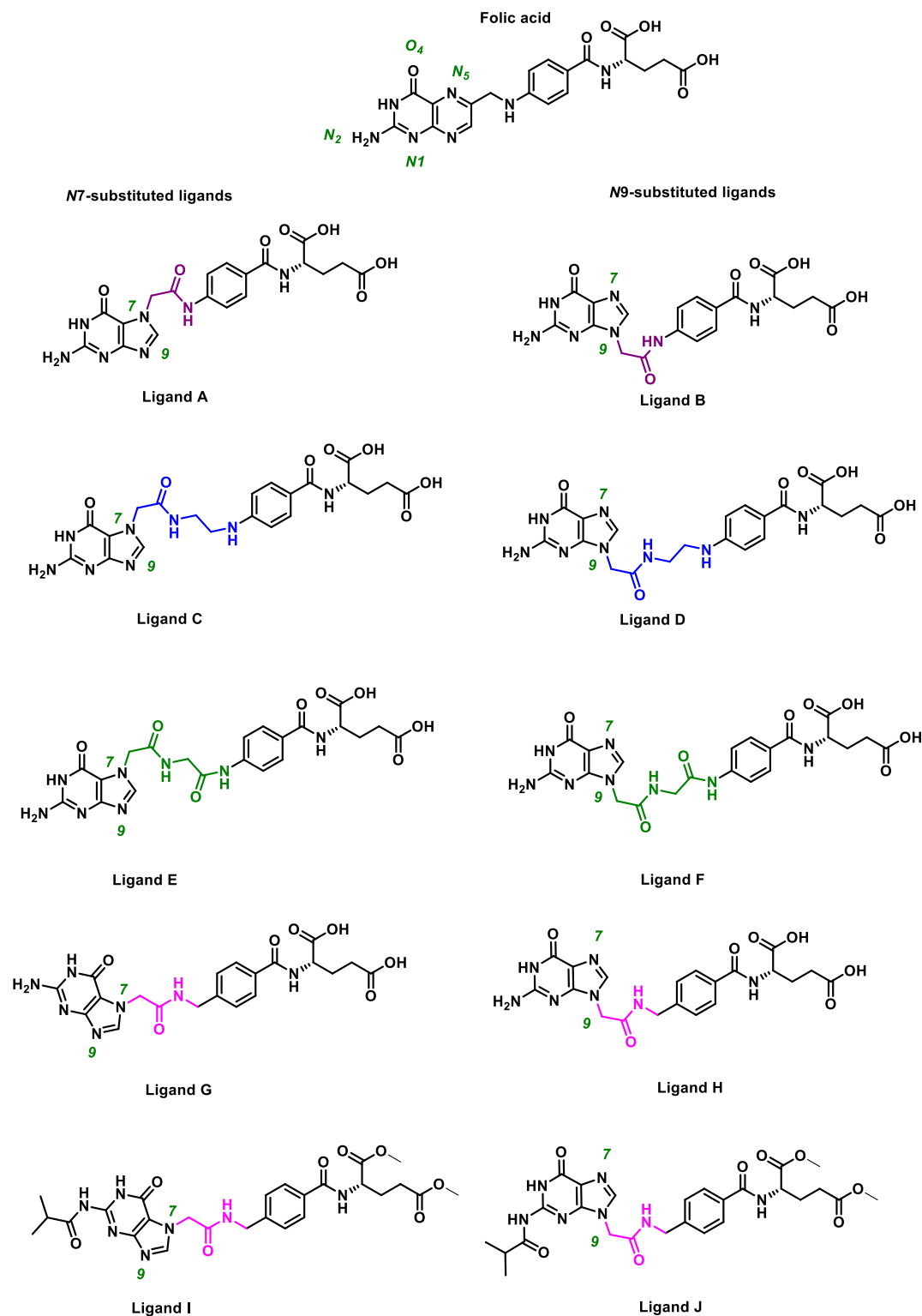
Figure 43: Chemical structures of designed antifolate ligands

The syntheses of ligands A and B were attempted starting from commercially available guanine. Guanine was reacted with isobutyryl chloride in dry DMF in presence of triethylamine to yield compound **1** in 64% yield, which was further subjected to *N* alkylation with ethylbromoacetate in presence of triethylamine to afford *N*7-substituted **2** and *N*9-substituted **3** in 27% and 30% yields respectively (Scheme 4) after column chromatography. The *N*7 substituted guanine (compound 2) was confirmed by X-ray analysis (figure 44). Further compound **2** and **3** were separately subjected to base mediated ester hydrolysis in presence of 1N LiOH to yield compounds **4** and **5** respectively in 61% and 66% yield respectively. Further, commercially available *para* amino benzoic acid was reacted with Cbz-Cl in presence of NaHCO₃ to give compound **6** in 80% yield. Also commercially available *L* glutamic acid was converted to its di-ester by reacting it with SOCl₂ in MeOH as a solvent to give compound **9** in 67% yield. Compound **6** was then reacted with compound **9** over 2 steps to give compound **7** in 67% yield. The Cbz protection was removed by hydrogenation from compound **7** in presence of palladium on carbon to give compound **8** in 82% yield. Further compound **5** and compound **8** were reacted under different conditions (table 6) but did not yield the desired ligand (Scheme 4).

Scheme 4: Attempted Synthesis of ligand B
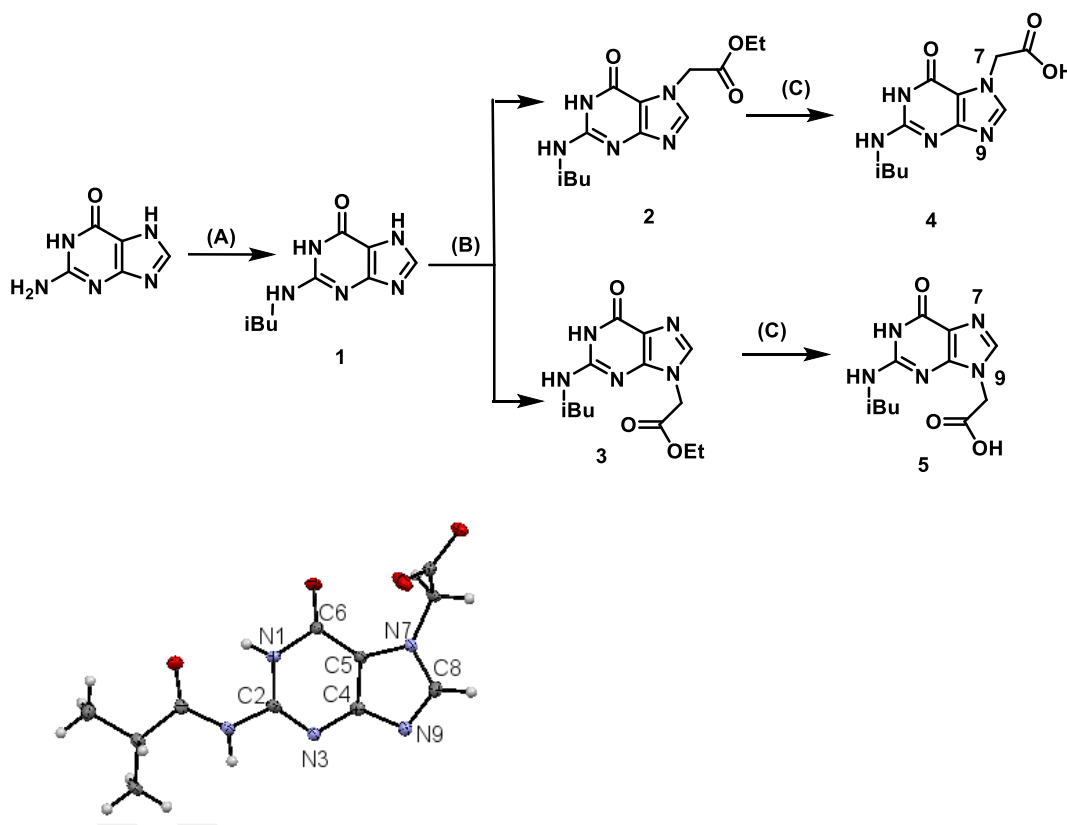
Figure 44: ORTEP structure of compound 2

Scheme 4 contd.
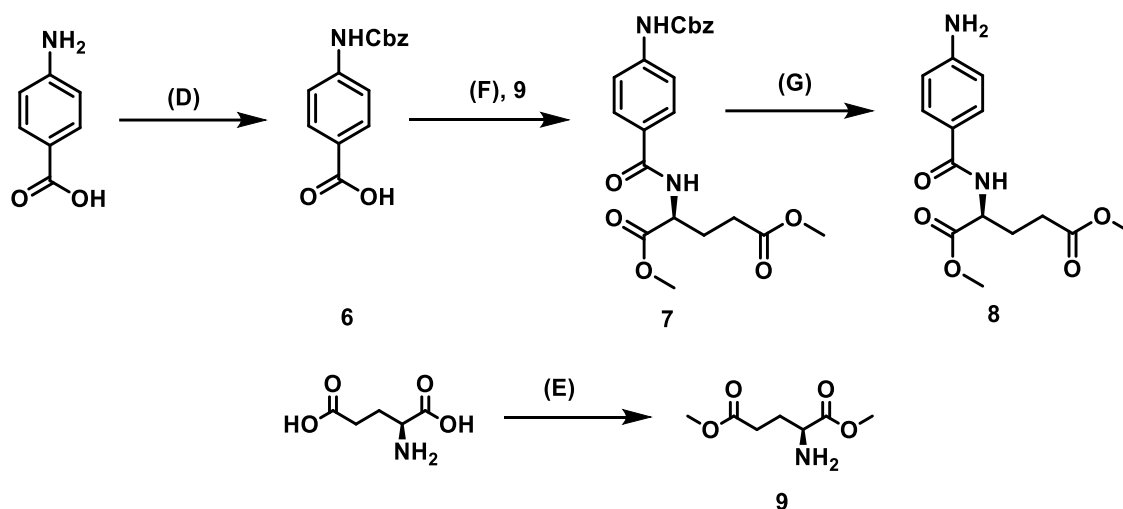
Reagents and conditions: (A) iBu-Cl, DMF, 150 °C, 64% **(B)** Ethylbromoacetate, TEA, DMF, 0° C- rt, 27% (*N7*) + 30% (*N9*) respectively **(C)** 1N LiOH, MeOH, rt, 61% (*N7*) + 66% (*N9*) **(D)** CBZ-Cl, NaHCO₃, 80% **(E)** SOCl₂, MeOH, 67% **(F) (i)** SOCl₂, DCM, DMF, reflux, **(ii)** TEA, DMF, rt, 67% **(G)** H₂, Pd/C, MeOH, rt, 82%
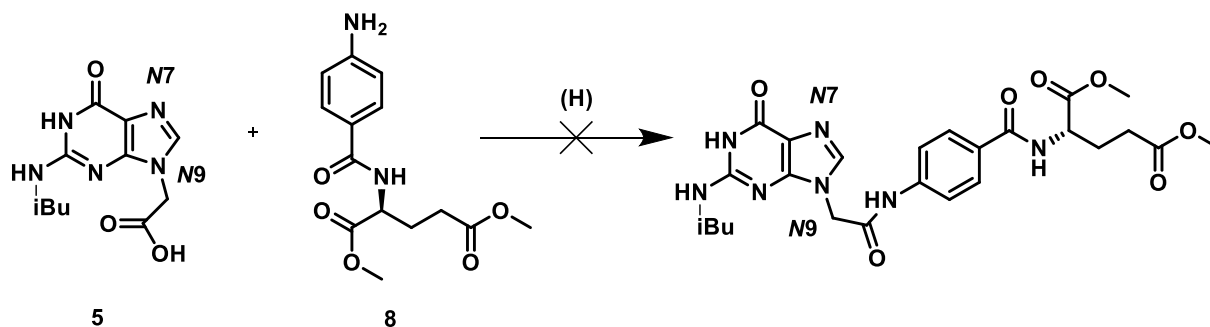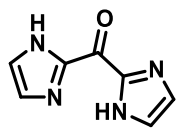
Table 6: Optimisation of different reaction conditions for reaction (H)

Entry No.	Reagent/ Base	Solvent	Time/ Temperature	Comment
1	HOBt (3 eq.) EDC (3 eq.) DIPEA (3 eq.)	DMF	20h, rt	No reaction, starting material recovered
2	HBTU (3 eq.) DIPEA (3 eq.)	DMF	20h, rt	No reaction, starting material

				recovered
3	PyBOP (3 eq.) TEA (3 eq.)	DMF	20h, rt	No reaction, starting material recovered
4	(i) SOCl₂ (5 eq.) (ii) TEA (3 eq.)	DCM DMF	5h, reflux	No reaction, starting material could not be recovered
5	(3 eq.) TEA (3 eq.)	THF	20h, rt	No reaction, starting material recovered

For the synthesis of ligands C and D (Scheme 5), commercially available glycine was reacted with Boc anhydride in presence of TEA to give compound **10** in 75% yield. Further compound **10** was reacted with *N,O*-dimethylhydroxylamine in presence of EDC and TEA to give compound **11** in 64% yield. Compound **11** was further reacted with LiAlH₄ at -78 ℃ to give compound **12** in 30% yield. Compound **12** was reacted under reductive amination conditions with methyl-4-aminobenzoate in presence of AcOH and NaBH₃CN to give compound **13**. Unfortunately, subjecting compound **13** to conditions of ester hydrolysis in presence of LiOH did not yield in the desired ligand, the starting material remained unreacted after 10h of reaction.

Scheme 5: Attempted Synthesis of ligand D
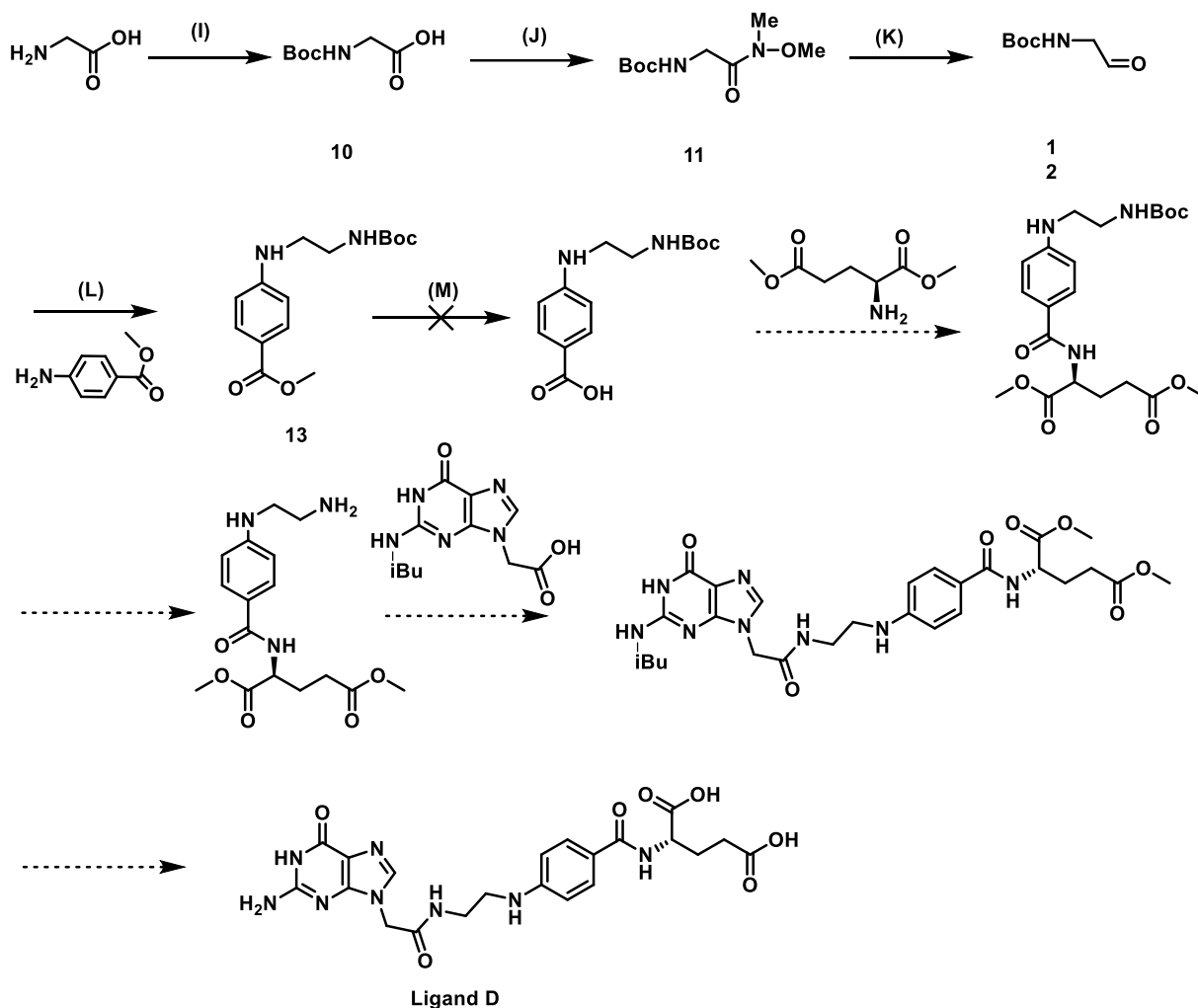
Reagents and conditions: (I) Boc anhydride, TEA, 75% **(J)** *N,O*-Dimethylhydroxylamine, EDC, TEA, 64% **(K)** LiAlH$_4$, -78 °C, 30% **(L)** methyl 4 aminobenzoate, AcOH, NaBH$_3$CN, MeOH, rt, 36% **(M)** 2N LiOH, THF:H$_2$O(1:1), reflux

The synthesis of ligand E (Scheme 6) was attempted from commercially available *p*-amino benzoic acid, which was reacted with chloroacetyl chloride and NaOH to give compound **14**. Compound **14** on further reacting with compound **9** in presence of PyBOP and TEA gave compound **15** in 50% yield. Compound **15** reacting with NaN$_3$ gave compound **16** in 40% yield. Further hydrogenation of compound **16** in presence of Pd/C gave a crude mixture which could not be purified. However, this mixture on reacting with compound **4** in presence of coupling reagents HOBt and EDC did not give the desired ligand.

Scheme 6: Attempted Synthesis of ligand E
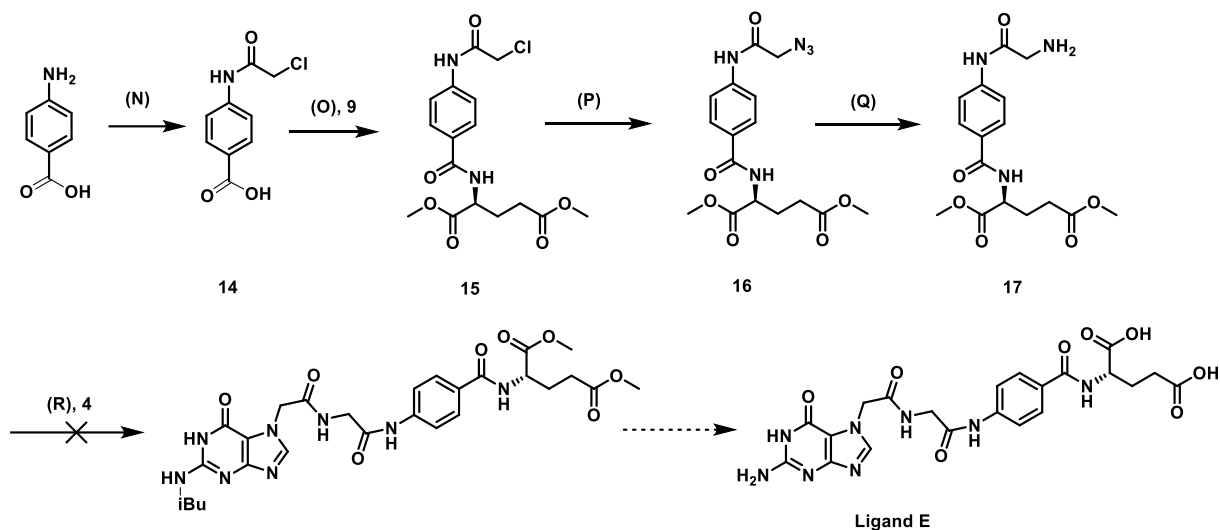
Reagents and conditions: **(N)** Chloroacetyl chloride, NaOH, 70% **(O)** PyBOP, TEA, 50% **(P)** NaN$_3$, DMF, 40% **(Q)** H$_2$, Pd/C **(R)** HOBt, EDC

Finally, a new short synthetic route was designed with potential antifolate ligands G, H, I and J keeping the rationale in mind. The title compounds were synthesized using a simple strategy outlined in schemes 4, 7 and 8. Commercially available 4-aminomethylbenzoic acid was reacted with Boc anhydride in H$_2$O: dioxane (1:1) with NaOH as the base to give compound **18** in 75% yield (Scheme 7). Compound **18** and **9** were reacted together in presence of PyBOP and triethylamine in dry DMF as solvent at room temperature to yield compound **19** with 67% yield. Further treatment of compound **19** with 50% TFA in DCM at room temperature gave compound **20** in 60% yield. Compound **5** was reacted with compound **20** (Scheme 8) in presence of HATU and triethylamine in DMF to yield compound **21** (ligand J) in 22 % yield. All compounds were characterised by appropriate spectroscopic techniques and confirmed by HRMS analysis.

Scheme 7: Synthesis of dimethyl (4-(aminomethyl)benzoyl)-*L*-glutamate
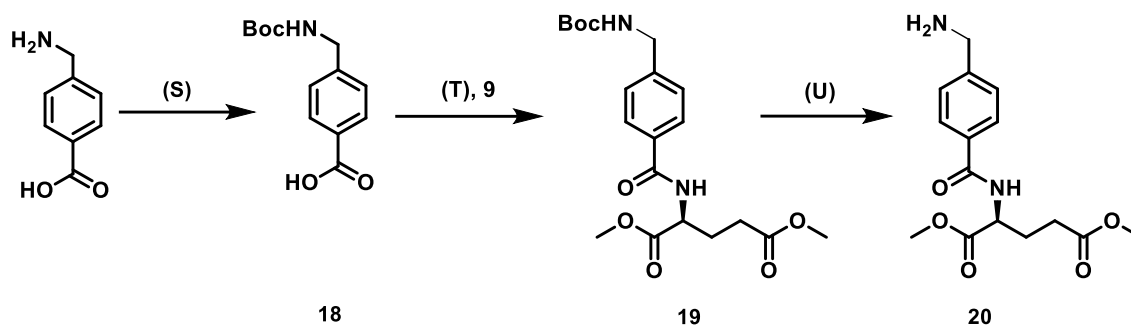
Reagents and conditions: (S) (Boc)$_2$O, NaOH, H$_2$O: Dioxane(1:1), rt, 75% **(T)** dimethyl *L*-glutamate, PyBOP, TEA, DMF, rt, 67% **(U)** 50% TFA/DCM, DCM, rt, 60%

Scheme 8: Synthesis of ligand J
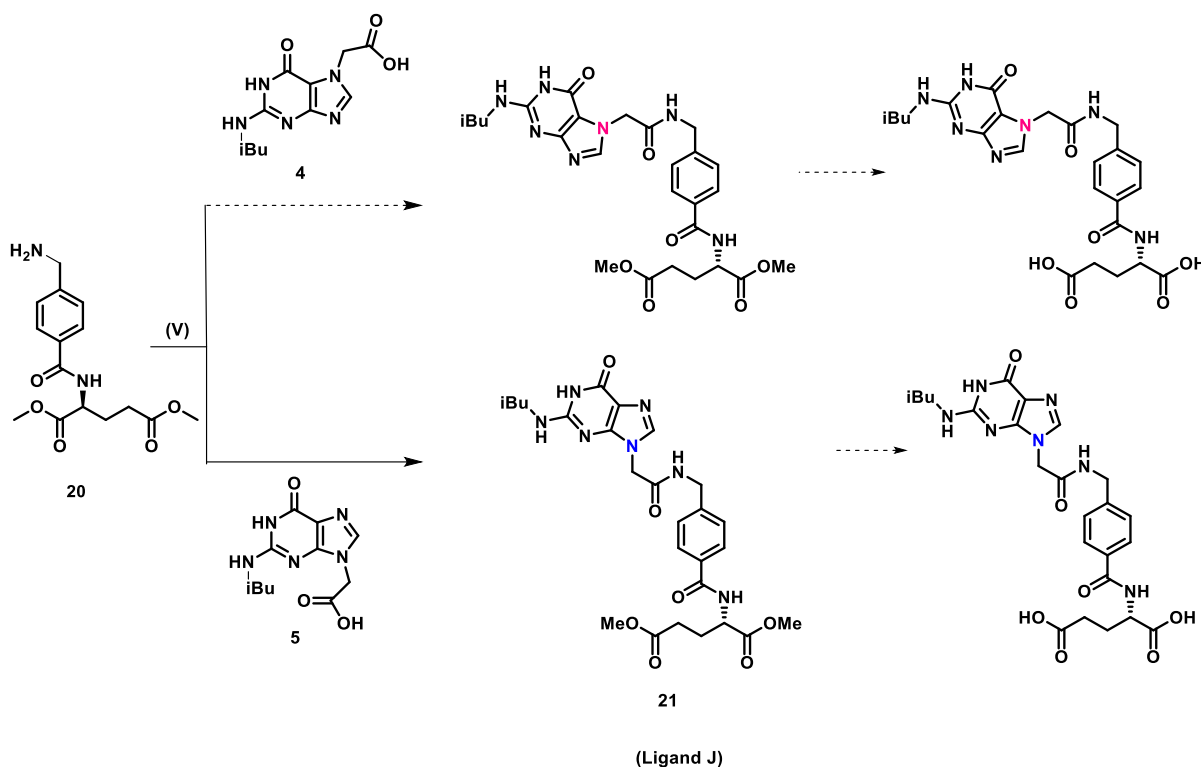
Reagents and conditions: (V) TEA, HATU, DMF, rt, 22%

4.3.2 Molecular docking study

Understanding small molecule interaction with ligands through molecular docking plays a very important role in the development of research and ultimately, development of drugs.[161] There exist different types of docking wherein protein-ligand docking is of special interest to researchers.[162] **For the sake of convenience, ligands G, H, I and J will be called as Ligands 1, 2, 3 and 4 respectively from now on in this chapter**. Hence in connection with efforts in searching for novel antifolates, studies were performed for the title ligands, folic acid and a marketed drug 'Pemetrexed' against the protein Serine hydroxymethyltransferase (SHMT). Pemetrexed is reported to act as a cancer chemotherapeutic against SHMT and the designed ligands of this study are structurally similar to this drug. SHMT is a ubiquitous enzyme found in all prokaryotes and is the only enzyme yet to be exploited as a target for cancer chemotherapy. Examples such as inhibition of the enzyme in cultured myeloma cells caused a dose dependent reduction of cell growth, a correlation between SHMT activity and tumour growth was also seen in human, mouse and rat leukemia models signify the importance of SHMT as a therapeutic target for cancer treatment [163]signify the importance of SHMT as a therapeutic target for cancer treatment. Additionally SHMT is an attractive target for cancer chemotherapy since it plays a pivotal role in nucleotide biosynthesis.

Two isoforms of SHMT are often found in humans and other higher organisms namely a cytosolic isoform (cSHMT) and a mitochondrial isoform (mSHMT).[84] All the computational procedures for this molecular docking study were carried out with the Molecular Operating Environment (MOE). Initially these protein receptors were prepared with the default 3D protonation procedure in MOE. Proteins in the study were the two isoforms of SHMT found in humans *viz* cSHMT (PDB Id: 1BJ4) and mSHMT (PDB Id: 6DK3) and another isoform found in *Mus musculus*; *mm*SHMT (PDB Id: 1EJI). Higher interaction energy of the ligands to the above proteins would suggest a better binding affinity and vice versa which would make them potential antifolates. Additionally all the ligands were docked against whole folate receptor FRα (PDB Id: 4LRH) and the known binding pocket separately to corroborate the rationale.

Table 7: Docking result for ligands in binding site of folate receptor

Entry no.	Ligand	Binding Energy (Kcal/mol)
1	1	-16.032
2	2	-14.839
3	3	-12.767
4	4	-12.969
5	Pemetrexed	-14.822
6	Folic Acid	-13.661

All the ligands in their ionic states were docked against binding pocket of folate receptor to determine the binding energies. As seen in Table 7 above, binding energy of ligand 1 and Pemetrexed (entry no. 1 and 5 respectively) is lower than the binding energy of folic acid (entry no. 6). Whereas ligands 2, 3 and 4 (table 7, entry 2, 3 and 4 respectively) also show binding energies similar to native folic acid. Among all the ligands studied, ligand 1 demonstrates the lowest binding energy with the binding pocket of the receptor. This result confirms that the newly designed antifolate analogues have affinity towards the receptor allowing for competitive binding with folic acid. The main interactions between the folic acid and ligands 1, 2, 3, 4 with folate receptor in the binding pocket and the predicted docked conformation are shown in the Figure 45. The high binding predicted for ligand 1 can be attributed to π-interaction with Trp A102, Trp A171 and Tyr A85 and hydrogen bonding with Ser A174 and Asp A81 in the binding pocket, among others (Figure 45 (A)). Similarly, ligand 2 shows π interaction with Tyr A85 and Trp A171 and hydrogen bonding with Thr A82 (Figure 45 (B)). Ligand 3 shows one π interaction with Trp A171 and hydrogen bonding with Thr A82, Asp A81, His A135 and Ser A174 (Figure 45 (C)). In case of ligand 4, no π interaction is seen with residues in the binding pocket while one hydrogen bonding interaction is seen with Ser A174 (Figure 45 (D). Commercial drug Pemetrexed shows π interaction with Trp A171, Tyr A85 while hydrogen bonding is seen with Asp A81, Trp A102, Trp A140 (Figure 45 (E)). In our study, folic acid shows π interaction with Trp A140 and hydrogen bonding interaction can be seen with Tyr A85 and His A135 (Figure 45 (F)). As discussed in the rationale, interactions between native folic acid and the folate receptor show $N1$ and $N2$ atoms (Figure 42) form strong hydrogen bonds with the side-chain carboxyl group of D81, the $N3$ and $O4$ atoms (Figure 42) with the S174 hydroxyl group. The $O4$ atom additionally forms two hydrogen

bonds with the guanidinium groups of R103 and R106, and the *N5* atom forms one hydrogen bond with the H135 side chain. Extensive interactions are also seen for the glutamate part of the molecule, which forms six hydrogen bonds with the reactive site residue W102, K136 and W140 side chains, as well as by backbone interactions with H135, G137 and W138. Although different softwares were used for docking in our case, in comparison to the reported literature,[160] similar interactions are seen in both the cases.

(A)

Ligand 1- 4LRH interaction

(B)

Ligand 2- 4LRH interaction

(C)

Ligand 3 - 4LRH interaction

(D)

Ligand 4 - 4LRH interaction

Pemetrexed - 4LRH interaction

Folic Acid - 4LRH interaction
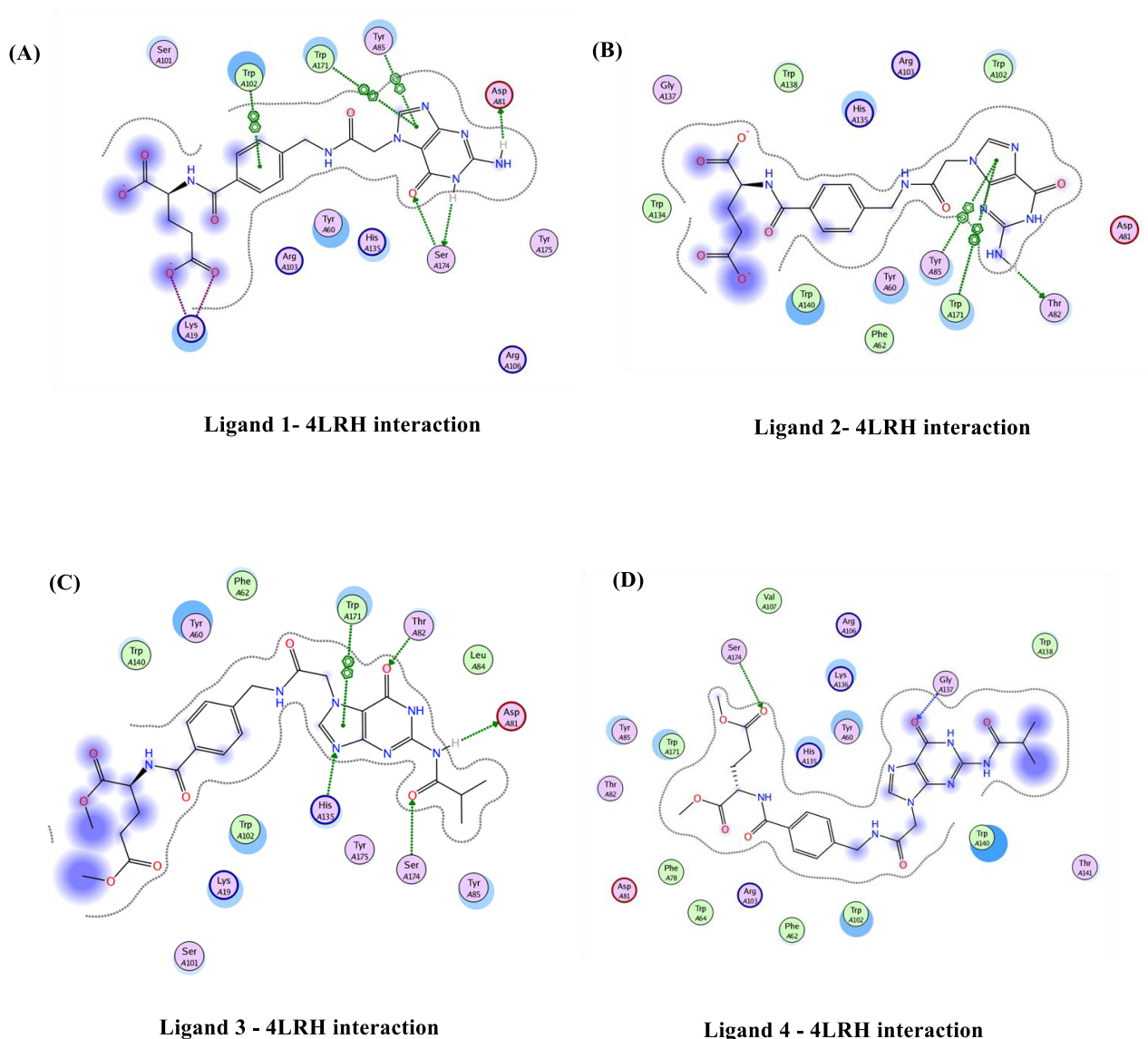
Figure 45: Two-dimensional depiction of the observed interactions between Ligands 1 (A), 2 (B), 3 (C), 4 (D), Pemetrexed (E), Folic acid (F) and active site of 4LRH obtained by MOE. Key interactions are depicted as follows: ⬭ polar , ⬭ acidic , ⬭ basic , ⬭ greasy , proximity contour , ┅▶ sidechain acceptor , ◀┅ sidechain donor , ┄▶ backbone acceptor , ◀┄ backbone donor , ● ligand exposure , ○ solvent residue , ◯ metal complex , ┄ solvent contact , ┄ metal contact , ◯ receptor contact , ◎◎ arene-arene , ◎ H arene-H , ◎+ arene-cation , **three letter code denotes amino acid in the protein, Single alphabet denotes protein chain, number denotes amino acid residue.**

Table 8: Docking energy of ligands and whole proteins (folate receptor; 4LRH, cytosolic SHMT; 1BJ4, mitochondrial SHMT; 6DK3, *mm*SHMT; 1EJI)

Entry No.	Name	Hydrogen Bonding	π interaction	Target	Binding Energy (Kcal/mol)
1	Ligand 1	Lys F 30, Arg F 36	Arg F 36, Arg F 39	(a) 4LRH	-12.195
		His 504, Arg 66, Thr 406	-	(b) 1BJ4	-13.163
		Asn 55, Arg 43	-	(c) 6DK3	-11.500
		Tyr A 73, Asp B 228, Thr B 388	-	(d) 1EJI	-14.724

2	Ligand 2	Pro D 27, Arg F 36	-	(a) 4LRH	-12.099
		Asn 78, Glu 77, Glu 401	Arg 66	(b) 1BJ4	-13.422
		Asn 42, Glu 54, Arg 465, Glu 469	Arg 43	(c) 6DK3	-11.605
		Asn B 385, Leu B 396, Tyr A 73	Arg B 402	(d) 1EJI	-17.754
3	Ligand 3	Thr D 46	Arg F 36, Lys F 30	(a) 4LRH	-11.022
		Lys 64, Gln 67		(b) 1BJ4	-11.530
		Arg 43, Ser 381	Arg 465	(c) 6DK3	-11.615
		Tyr A 83, Gly A 303, Thr B 254, His B 256, Ser B 203	Lys B 257	(d) 1EJI	-12.197
4	Ligand 4	Lys D 30	-	(a) 4LRH	-10.042
		Glu 72, Arg 69	Arg 66	(b) 1BJ4	-10.592
		Arg 263, Glu 378	Arg 43	(c) 6DK3	-10.086
		Lys B 257, Ser B 203, Asn B 124, Thr B 388	-	(d) 1EJI	-12.584
5	Pemetrexed	Arg F 36, Arg F 39, Lys f 30, Glu D 38	-	(a) 4LRH	-12.234
		Arg 69, Arg 66, Thr 406	Arg 66	(b) 1BJ4	-11.795
		Asn 42, Lys 39, Ser 381, Glu 378	-	(c) 6DK3	-11.365
		Glu A 75, Ser B 53, Arg b 402	-	(d) 1EJI	-15.350
6	Folic Acid	Arg F 39, Arg F 36, Pro D 27, Glu D 28, Lys F 30	Arg F 39, Arg F 36	(a) 4LRH	-11.344
		Glu 77, Arg 66, Arg 69	Arg 66	(b) 1BJ4	-13.806
		Arg 43, Ser 381, Arg 263	Arg 43	(c) 6DK3	-12.573
		Arg B 402, Tyr A 82, Glu A 75, Tyr A 83, Ser B 119	Lys B 257	(d) 1EJI	-17.149

Further, when the ligands were docked against whole FRα protein, (4LRH) interaction energy of ligands compared to folic acid was also found to be slightly low (table 8, entry 1, 2, 3 and 6 subentry (a)) when compared to native folic acid. This result was again in line with the hypothesis of this study since the designed ligands were envisaged to have a competitive binding with folic acid. Furthermore, when the ligands were docked against cytosolic SHMT (1BJ4), among all the ligands docked, ligand 1 (table 8, entry 1, subentry (b)) and ligand 2 (table 8, entry 2, subentry (b)) had a better interaction with the protein when compared to commercial drug Pemetrexed (table 8, entry 5, subentry (b)). When the ligands were docked against mitochondrial SHMT (6DK3), all the ligands showed similar binding interaction with the protein. Docking ligands against *mm*SHMT (1EJI) revealed low interaction energy for ligand 1(table 8, entry 1, subentry (d)), ligand 2 (table 8, entry 2, subentry (d)), Pemetrexed (table 8 entry 5, subentry (d)) and folic acid (table 8, entry 6, subentry (d)). Among all the designed ligands, ligand 2 (table 8, entry 2, subentry (d)) showed lowest interaction energy with *mm*SHMT conferring this ligand with highest binding with the protein. On the other hand, ligand 3 (table 8, entry 3, subentry (a), (b), (c), (d)) and ligand 4 (table 8, entry 4, subentry (a), (b), (c), (d)) overall showed high interaction energy revealing an important structural aspect which is the presence of a free amine at position 2 and the absence of the methyl ester on the glutamate moiety. Specifically the main interactions between the ligands 1, 2, 3, 4, Pemetrexed and folic acid with the 4 proteins and the predicted docked conformation are shown in the Figures 46, 47, 48 and 49.

Figure 46 shows interactions of all the studied ligands with folate receptor, 4LRH. To summarise the interactions, only ligands 1, 3 and folic acid exhibit π interactions, specifically with Arg F39, Arg F36 and Lys F30. The hydrogen bonding interactions are seen in all the ligands with Arg F36, Lys F30, NAG K303, Pro D27, Thr D46, Lys D30, Glu D28 and Arg F39.
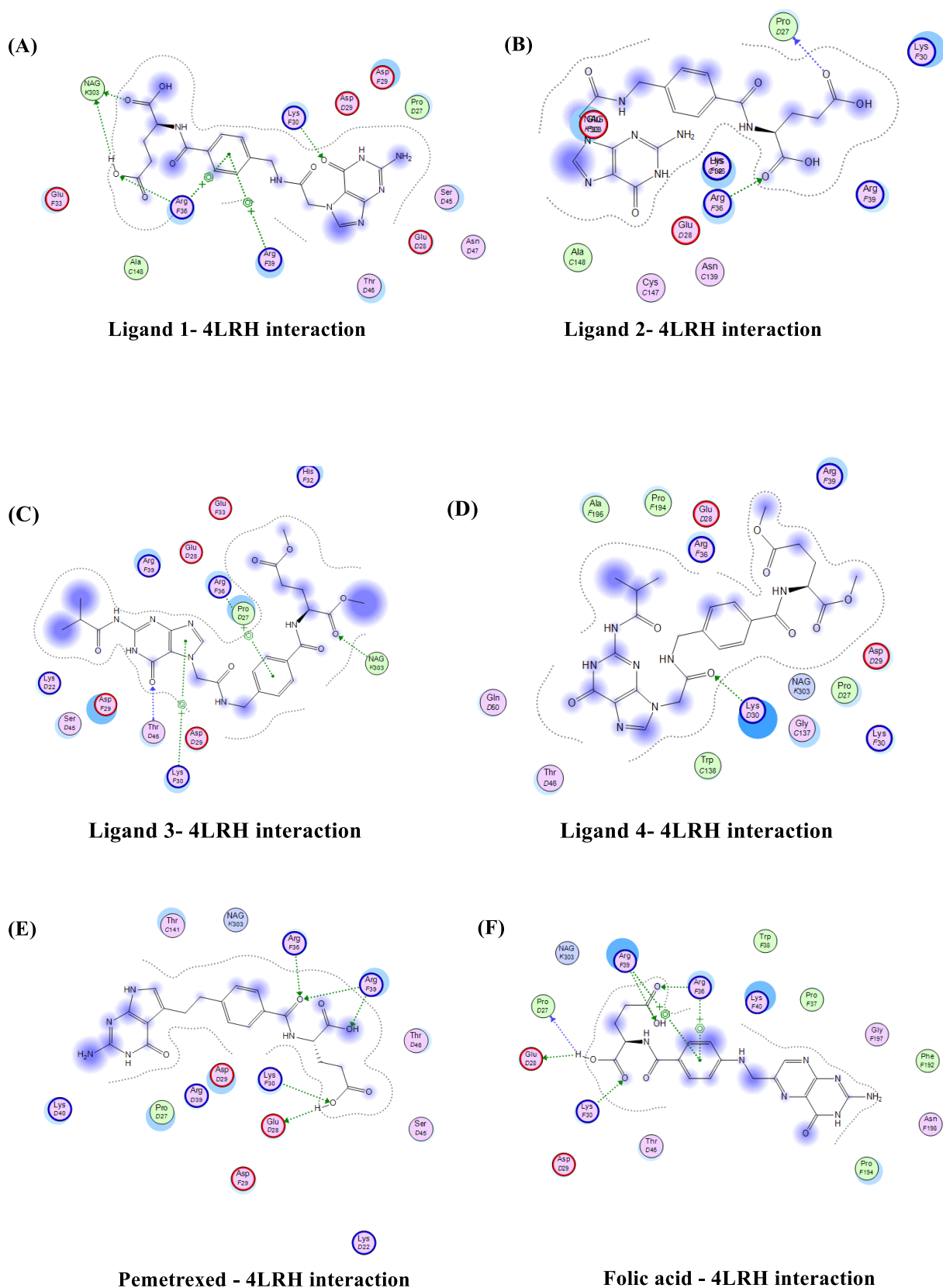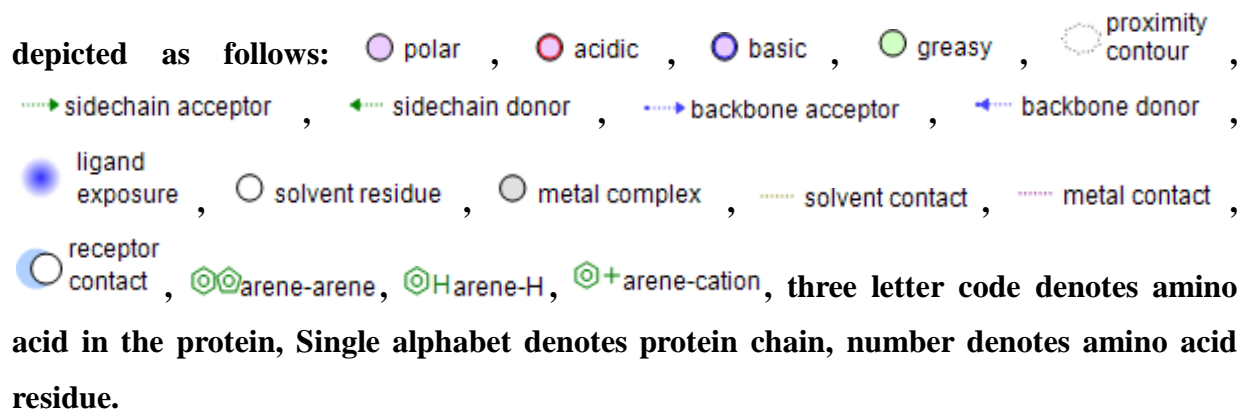
Figure 46: Two-dimensional depiction of the observed interactions between Ligands 1 (A), 2 (B), 3 (C), 4 (D), Pemetrexed (E), Folic acid (F) and 4LRH. Key interactions are

depicted as follows: ○ polar , ○ acidic , ○ basic , ○ greasy , ⬡ proximity contour ,
┈▶ sidechain acceptor , ◀ sidechain donor , ┈▶ backbone acceptor , ◀ backbone donor ,
🔵 ligand exposure , ○ solvent residue , ○ metal complex , ┈ solvent contact , ┈ metal contact ,
○ receptor contact , ◎◎ arene-arene, ◎ Harene-H, ◎+ arene-cation , **three letter code denotes amino acid in the protein, Single alphabet denotes protein chain, number denotes amino acid residue.**

Interactions of ligands with IBJ4 in Figure 47 can be summarised as π interactions with ligands 2, 3, 4 and folic acid specifically with Arg 43 and Arg 455 and hydrogen bonding interactions specifically with Arg 43, Asn 55, Asn 42, Arg, 465, Glu 54, Ser 381, Glu 378 and Arg 263.

(A)
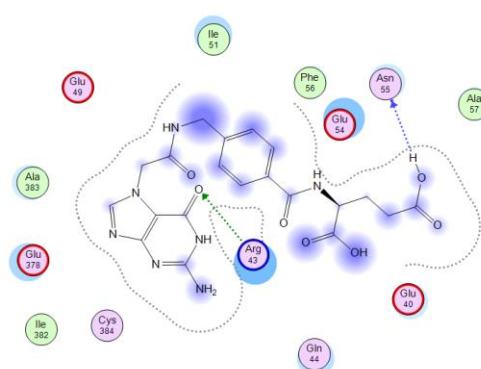
Ligand 1-1BJ4 interaction

(B)
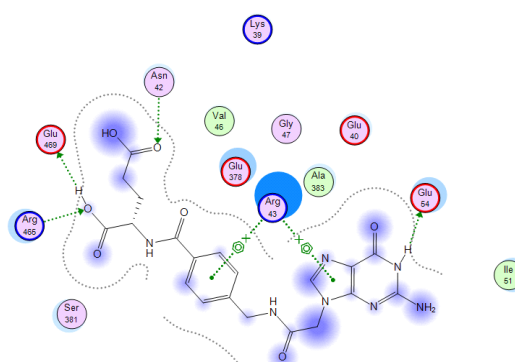
Ligand 2-1BJ4 interaction

(C)

Ligand 3- 1BJ4 interaction

(D)

Ligand 4- 1BJ4 interaction
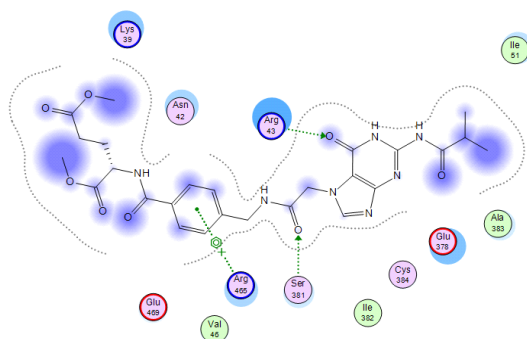
(E) Pemetrexed - 1BJ4 interaction
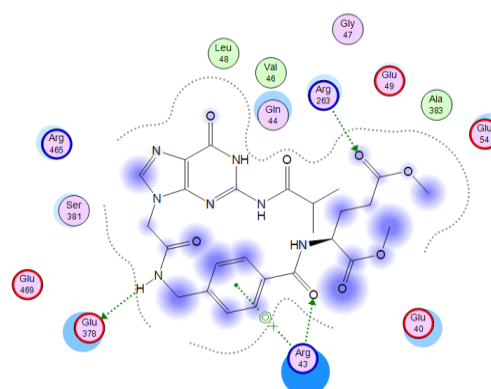
(F) Folic acid - 1BJ4 interaction
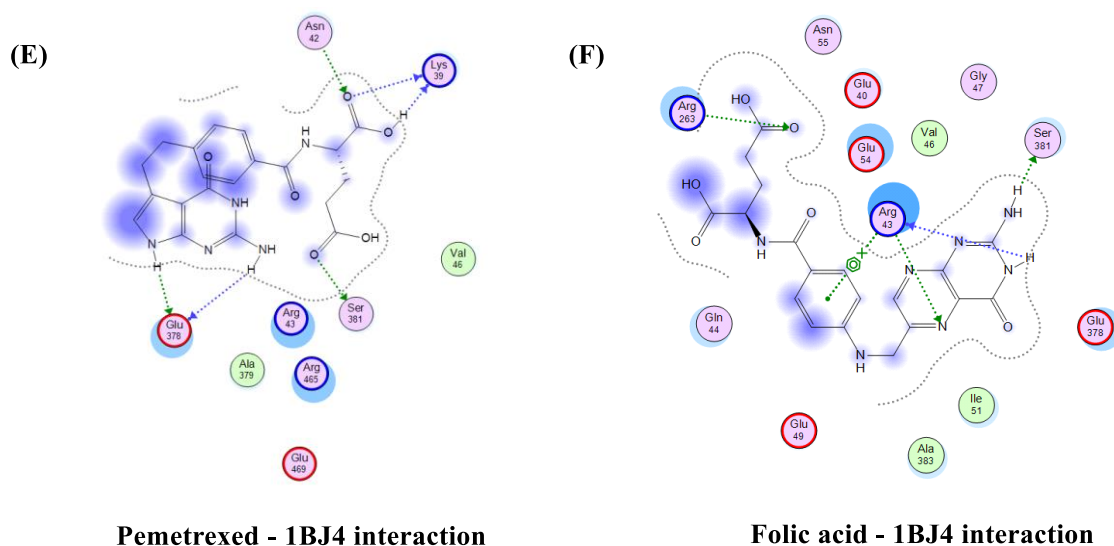
Figure 47: Two-dimensional depiction of the observed interactions between Ligands 1 (A), 2 (B), 3 (C), 4 (D), Pemetrexed (E), Folic acid (F) and 1BJ4. Key interactions are depicted as follows: ○ polar , ○ acidic , ○ basic , ○ greasy , ⬡ proximity contour , ┈▶ sidechain acceptor , ◀┈ sidechain donor , ┈▶ backbone acceptor , ◀┈ backbone donor , ● ligand exposure , ○ solvent residue , ○ metal complex , ┈ solvent contact , ┈ metal contact , ○ receptor contact , ⬡⬡ arene-arene , ⬡ Harene-H , ⬡+ arene-cation , **three letter code denotes amino acid in the protein, As there is only a single protein chain, the single alphabet denoting protein chain is not included, number denotes amino acid residue.**

Figure 48 shows π and hydrogen bonding interactions between all the ligands and the protein 6DK3. π interactions can be seen with ligands 2, 4, Pemetrexed and folic acid with Arg 66 only, whereas hydrogen bonding can be seen with His 504, Arg 66, Thr 406, Asn 78, Glu 77, Glu 401, Lys 64, Gln 67, Glu 72, Arg 69 and Thr 406.
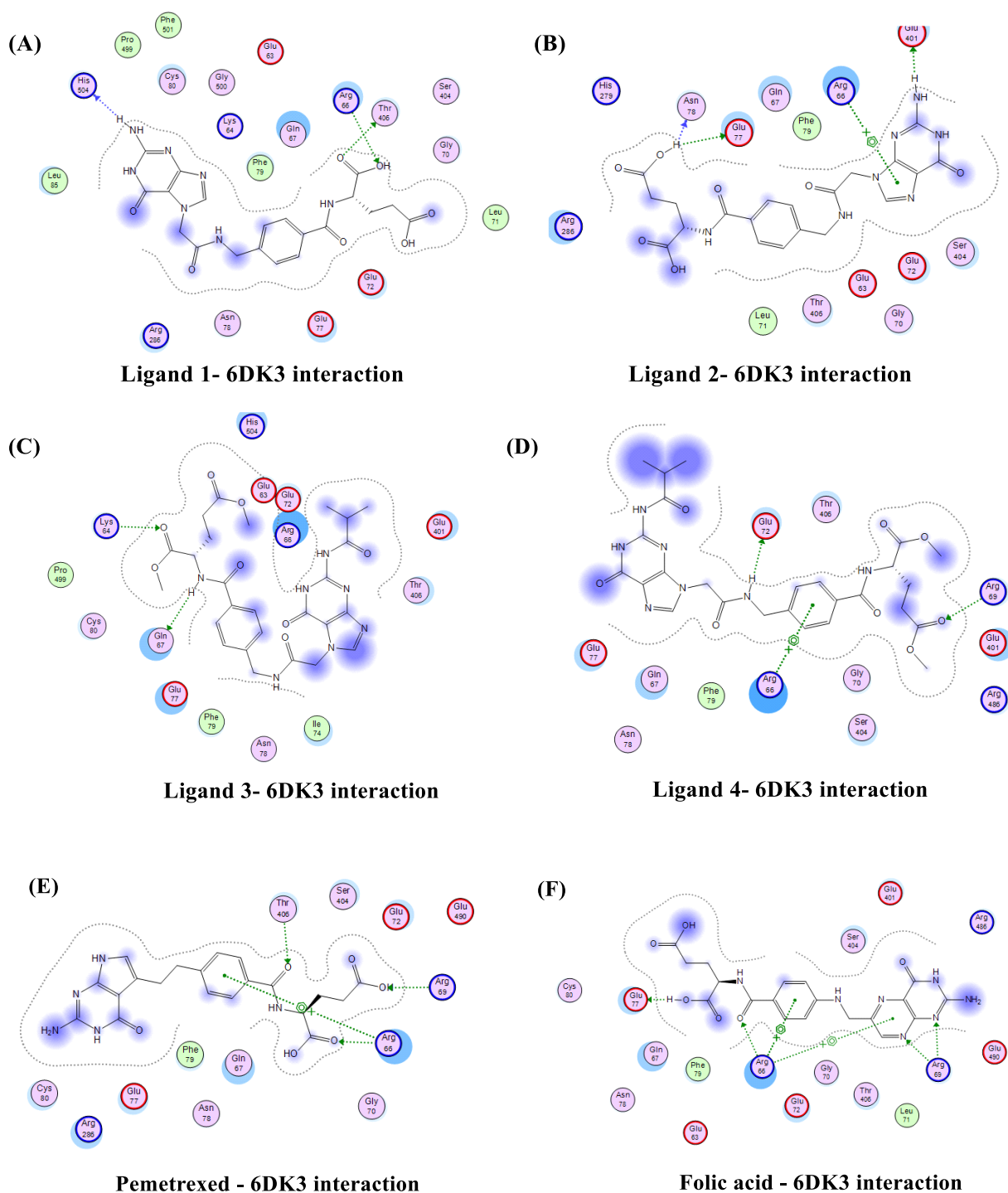
Figure 48: Two-dimensional depiction of the observed interactions between Ligands 1 (A), 2 (B), 3 (C), 4 (D), Pemetrexed (E), Folic acid (F) and 6DK3. Key interactions are depicted as follows: ○ polar , ○ acidic , ○ basic , ○ greasy , proximity contour , ····▶ sidechain acceptor , ◀···· sidechain donor , ···▶ backbone acceptor , ◀··· backbone donor , ligand exposure , ○ solvent residue , ○ metal complex , ···· solvent contact , ···· metal contact , ○ receptor contact , arene-arene, arene-H, arene-cation, **three letter code denotes amino**

acid in the protein, As only one protein chain is present, the single alphabet denoting protein chain is omitted, number denotes amino acid residue.

π and hydrogen bonding interactions between all the ligands and the protein 1EJI can be seen in Figure 49. The amino acids Arg B402, and Lys B257 are involved in π interactions between ligands 2, 3 and folic acid with the protein. Whereas amino acids Thr B288, Tyr A73, Asp B228, Asn B385, Leu B396, Tyr A83, Gly A303, Thr B254, Ser B203, Lys B257, Asn B124, Glu A75, Ser B53, Arg B402, Tyr A82, Ser B119 are involved in hydrogen bonding interactions with the protein.

(A)

Ligand 1- 1EJI interaction

(B)

Ligand 2- 1EJI interaction

(C)

Ligand 3- 1EJI interaction

(D)

Ligand 4- 1EJI interaction

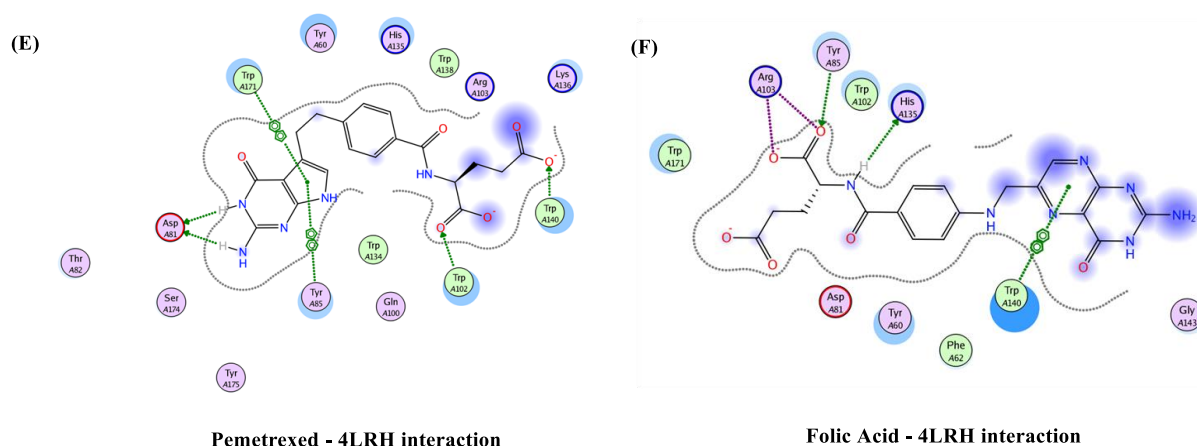**(E)** Pemetrexed – 1EJI interaction

(F) Folic acid – 1EJI interaction

Figure 49: Two-dimensional depiction of the observed interactions between Ligands 1 (A), 2 (B), 3 (C), 4 (D), Pemetrexed (E), Folic acid (F) and 1EJI. Key interactions are depicted as follows: ⬡ polar , ⬡ acidic , ⬡ basic , ⬡ greasy , proximity contour , ⋯▶ sidechain acceptor , ◀⋯ sidechain donor , ⋯▶ backbone acceptor , ◀⋯ backbone donor , ligand exposure , ⬡ solvent residue , ⬡ metal complex , ⋯ solvent contact , ⋯ metal contact , ⬡ receptor contact , ⬡⬡ arene-arene, ⬡ Harene-H, ⬡+ arene-cation, **three letter code denotes amino acid in the protein, Single alphabet denotes protein chain, number denotes amino acid residue.**

4.4 Conclusion

Antifolate analogues were designed and their synthesis from commercially available guanine was attempted. Guanine was chosen for the design and synthesis as it is structurally similar to the pteridine ring and has sites suitable for further chemical modifications and construction of the antifolates. All the analogues were structurally similar to commercial antifolate Pemetrexed. Molecular docking study was carried out for all the analogues against different proteins such as native folate receptor (4LRH), and different isoforms of Serine hydroxymethyltransferase (cSHMT, mSHMT, *mm*SHMT). The analogues were docked against whole protein as well as binding pocket of 4LRH. Docking study revealed that all the designed antifolates showed appreciable binding to the target proteins. From a structural aspect, *N*7 analogues showed better interaction with the target proteins than the *N*9 analogues. Additionally analogues bearing free amine at position 2 and bearing no ester group in glutamic acid were better in interaction with the target proteins.

4.5 Methods

4.5.1General Information

All reagents were purchased from Sigma-Aldrich. DMF, DCM, were dried following standard procedures. DMF, DCM were stored in 4 Ắ molecular sieves. Column chromatography was performed for purification of compounds on silica gel (100- 200 mesh or 60-120 mesh, Merck). TLCs were performed on Merck 5554 silica 60 pre-coated aluminium sheets. Compounds were visualized under UV light. ^{1}H and ^{13}C NMR spectra were obtained using Bruker AC-200, AC-400 or AC-500 NMR spectrometers. The chemical shifts are reported in delta (δ) values and referred to internal standard TMS for 1H. ^{1}H NMR data are reported in the order of chemical shift, multiplicity (s, singlet; d, doublet; t, triplet; q, quartet; br, broad; br s, broad singlet; m, multiplet and/ or multiple resonance), number of protons.

4.5.2MOE Methodology used for Docking Study

The drug compounds were also employed to energy minimization using MOE. Docking study were performed using all default parameters with Triangle Matcher, Rigid Receptor, initial scoring method London dG retaining 10 poses and final scoring method used was GBVI/WSA with 5 poses. Protein structures were prepared in MOE after removing water molecules, and all hydrogen atoms were added to the structure with their standard geometry followed by their energy minimization using default parameters with Forcefield value Amber 10:EHT and RMS gradient of 0.1 kcal/mol. The binding site of each receptor was identified through the MOE Site Finder program, which uses a geometric approach to calculate putative binding sites in a protein, starting from its tridimensional structure. Active sites were identified and dummy atoms were created around the resulting alpha spheres centers. The backbone and residues were kept fixed and the energy minimization was performed. Additionally, ligands were prepared by converting into 3D structures and energy were minimised.

4.5.3 Synthesis of designed antifolates
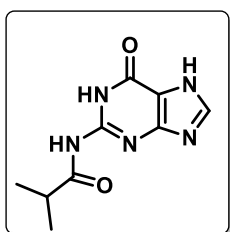
N-(6-oxo-6,7-dihydro-1*H*-purin-2-yl)isobutyramide (**1**) Guanine(1 g, 6.6mmol) was suspended in dry DMF; to it dry TEA (0.6 mL, 6.6 mmol) was added at room temperature. To this reaction mixture, isobutyric chloride (0.6mL, 6.6mmol) was added with continuous stirring. The reaction mixture was refluxed for 5 hr at 150ºC cooled to room

temperature and the solvent was evaporated under reduced pressure. The crude product was recrystallised in boiling ethanol/water (1:1) to give compound 1 as white crystals (0.9g 64%)

^{1}H NMR (400 MHz, DMSO-d6) δ: 1.11-1.13 (d, 6H), 2.73-2.80 (m, 1H), 8.03 (s, 1H) ppm

^{13}C NMR (100 MHz, DMSO-d6) δ: 18.97, 34.79, 180.10 ppm

^{13}C DEPT (100 MHz, DMSO-d6) δ: 46.07 (CH$_2$), 19.41, 35.23 ppm

HRMS: Calculated Mass for C$_9$H$_{11}$N$_5$O$_2$: 221.09 Observed Mass: [M+H] 222.09
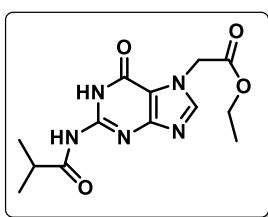
ethyl 2-(2-isobutyramido-6-oxo-1,6-dihydro-7*H*-purin-9 yl)acetate(2) To a suspension of compound 1 in dry DMF, (1g, 4.5mmol) at 0ºC was added TEA (0.75mL, 5.4mmol) and ethyl bromoacetate(0.55mL, 4.9mmol). The reaction mixtrure was stirred at room temperature for 10h. The reaction was monitored by TLC and after completion of the reaction, DMF was removed under vacuum. The and the residue was chromatographed using 0–5% (v/v) methanol in dichloromethane Eluted first was the less polar compound 2 (*N***7**-isomer) as a yellow solid (0.38g, 27%)

^{1}H NMR (400 MHz, DMSO-d6) δ: 1.10-1.14 (d, 6H), 1.21-1.25 (t, 3H), 2.67-2.81 (m, 1H), 4.12-4.22 (q, 2H), 5.21 (s, 2H), 8.15 (s, 1H), 11.60 (s, 1H), 12.16 (s, 1H) ppm

^{13}C NMR (100 MHz, DMSO-d6) δ: 14.00, 18.87, 34.73, 47.33, 61.35, 111.70, 144.93, 147.27, 152.63, 153.87, 167.87, 179.99 ppm

^{13}C DEPT (100 MHz, DMSO-d6) δ: 47.08 (CH$_2$), 61.10 (CH$_2$), 13.74, 18.62, 34.48, 144.67 ppm

HRMS Calculated Mass for C$_{13}$H$_{17}$N$_5$O$_4$: 307.13 Observed mass: [M+H] 308.13
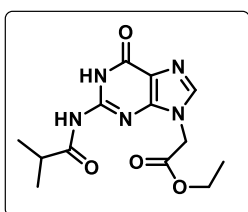
ethyl 2-(2-isobutyramido-6-oxo-1,6-dihydro-9*H*-purin-7-yl)acetate(3) To a suspension of compound 1 in dry DMF, (1g, 4.5mmol) at 0ºC was added TEA (0.75mL, 5.4mmol) and ethyl bromoacetate (0.55mL, 4.9mmol). The reaction mixtrure was stirred at room temperature for 10h. The reaction was monitored by TLC and after completion of the reaction, DMF was removed under vacuum. The and the residue was chromatographed using

0–5% (v/v) methanol in dichloromethane Eluted after the less polar *N7*-isomer was the *N9* isomer, compound 3 (0.42g, 30%)

¹H NMR (400 MHz, DMSO-d6) δ: 1.01-1.11 (d, 6H), 1.12-1.23 (t, 3H), 2.75-2.80 (m, 1H), 4.13-4.19 (q, 2H), 5.01 (s, 2H), 7.96 (s, 1H), 11.67(s, 1H), 12.10 (s, 1H) ppm

¹³C NMR (100 MHz, DMSO-d6) δ: 14.00, 18.85, 34.69, 44.30, 61.52, 119.71, 140.26, 148.17, 149.00, 154.84, 167.68, 180.23 ppm

¹³C DEPT (100 MHz, DMSO-d6) δ: 44.80 (CH₂), 62.02 (CH₂), 14.50, 19.35, 35.19, 140.76 ppm

HRMS Calculated Mass for $C_{13}H_{17}N_5O_4$: 307.13 Observed mass: [M+H] 308.13
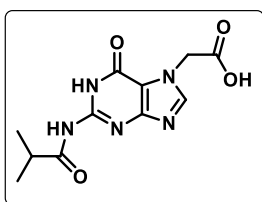
2-(2-isobutyramido-6-oxo-1,6-dihydro-7*H*-purin-7-yl)acetic acid(4) Compound 2 (0.1g, 0.32mmol)was dissolved in 2mL MeOH. To it 2N NaOH (2mL) was added dropwise. The reaction was monitored by TLC and after completion of reaction, the reaction mixture was neutralised using Dowex H⁺ resin; the resin was subsequently filtered off. The filtrate was concentrated under vacuum which afforded compound 4(0.055g, 61%)

¹H NMR (400 MHz, DMSO-d6) δ: 1.09-1.11 (d, 6H), 2.71-2.76 (m, 1H), 5.10 (s, 2H), 8.12 (s, 1H), 11.56 (s, 1H), 12.13 (s, 1H) ppm

¹³C NMR (100 MHz, DMSO-d6) δ: 18.96, 34.82, 47.45, 111.86, 145.03, 147.26, 152.75, 156.86, 169.39, 180.07 ppm

¹³C DEPT (100 MHz, DMSO-d6) δ: 47.19 (CH₂), 18.71, 34.56, 144.77 ppm

HRMS Calculated Mass for $C_{11}H_{13}N_5O_4$: 279.26 Observed mass: [M+H] 280.10
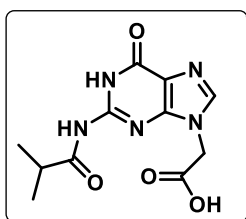
2-(2-isobutyramido-6-oxo-1,6-dihydro-9H-purin-9-yl)acetic acid(5) Compound 3 (0.1g, 0.32mmol)was dissolved in 2mL MeOH. To it 2N NaOH(2mL) was added dropwise. The reaction was monitored by TLC and after completion of reaction, the reaction mixture was neutralised using Dowex H⁺ resin; the resin was subsequently filtered off. The filtrate was concentrated under vacuum which afforded compound 4(0.06g, 66%)

¹H NMR (400 MHz, DMSO-d6) δ: 1.09-1.10 (d, 6H), 2.74-2.79 (m, 1H), 4.92 (s, 2H), 8.03 (s, 1H), 11.70 (s, 1H), 12.10 (s, 1H) ppm

¹³C NMR (100 MHz, DMSO-d6) δ: 18.89, 34.73, 45.68, 140.41, 148.18, 154.78, 169.04, 180.25 ppm

¹³C DEPT (100 MHz, DMSO-d6) δ: 44.29 (CH₂), 45.42 (CH₂), 18.85, 34.48, 140.15
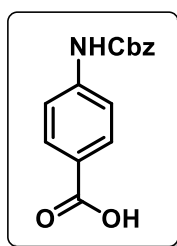
4-(((benzyloxy)carbonyl)amino)benzoic acid (6) *para* aminobenzoic acid (1g, 0.00729 mmol) was suspended in H₂O:THF (1:1) and NaHCO₃ (6.1g, 0.0729mmol) was added. This reaction mixture was cooled to 0 ℃ and Cbz-Cl (1 mL, 0.00729 mmol) was added dropwise and the reaction mixture was allowed to come to room temperature and stirred for 4 hr. After the completion of reaction indicated by TLC, the reaction mixture was acidified with 1N HCl giving a brown sticky solid. This solid was dissolved in ethyl acetate wherein a white solid precipitate, compound **6** was obtained which was filtered and used for further reactions without purification. (1.52g, 80%)

¹H NMR (400 MHz, DMSO-d6) δ: 5.21 (s, 2H), 7.36-7.45(m, 4 H), 7.69-7.72(d, *J*= 9 Hz, 2H), 8.05-8.07 (d, *J*=8 Hz, 2H), 10.38 (s, 1H) ppm
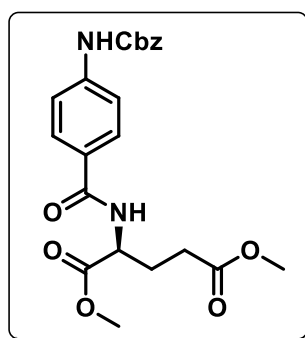
dimethyl (4-(((benzyloxy)carbonyl)amino)benzoyl)-*L*-glutamate (7) compound **6** (0.05 g, 0.184mmol) was dissolved in DCM and to this SOCl₂ (0.015 mL, 0.129 mmol) and catalytic amount of DMF was added and the reaction was refluxed for 4 hr. After the completion of 4 hr, the DCM was evaporated under vacuum, DMF and TEA (0.077 mL, 0.552 mmol) and compound 9 (0.097g, 0.552 mmol) was added and the reaction was stirred at room temperature for 10 hr. DMF was removed under vacuum and the crude reaction mixture was purified by column chromatography to yield compound **7** (0.053g, 67%)

¹H NMR (200 MHz, CHCl₃) δ: 1.80-2.06 (m, 3H), 3.48 (s, 3H), 3.54 (s, 3H), 4.31-4.36 (m, 1H), 5.07 (s, 2H), 7.29-7.32 (m, 5H), 7.43-7.47 (m, 2H), 7.69-7.75 (m, 2H), 8.48-8.52 (d, 1H), 9.97 (s, 1H) ppm
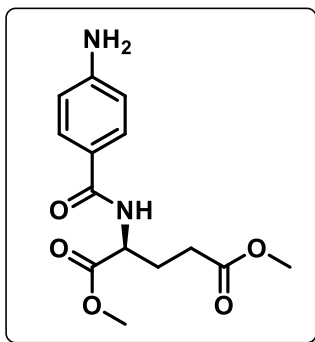
dimethyl (4-aminobenzoyl)-*L*-glutamate (8) compound **7**(0.1g, 0.233 mmol) was dissolved in MeOH (2 mL) followed by the addition of 10% Pd/C (10% w/w, 0.01g) and subjected to hydrogenation for 6 hr. After the completion of raction indicated by TLC, the reaction mixture was filtered over celite and the solvent was evaporated under vacuum giving compound **8** in high purity and was used for subsequent reactions without further purification. (0.057 g, 82%)

¹H NMR (200 MHz, CHCl₃) δ: 2.07-2.35 (m, 2H), 2.43-2.52 (m, 2H), 4.77-4.85 (m, 1H), 6.64-6.68 (d, *J*= 8 Hz, 2H), 7.62-7.67 (d, *J*= 8Hz, 2H) ppm
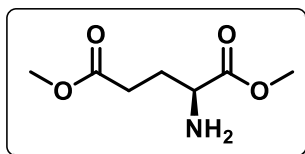
Dimethyl *L*-glutamate (9) Commercially available *L*-glutamic acid(1g, 6.9 mmol) was dissolved in MeOH (5mL). This reaction mixture was maintained at 0ºC and thionyl chloride was added dropwise (1mL, 1.39mmol). After completion of the reaction solvent was evaporated *in vacuo* to give compound **7** (0.8g, 67%).

¹H NMR (400 MHz, CHCl₃) δ: 2.36-2.41 (m, 2H), 2.63-2.74 (m, 2H), 3.66 (s, 3H), 3.81(s, 3H), 4.30 (br.s, 1H), 8.65 (s, 2H) ppm

¹³C NMR (100 MHz, CHCl₃) δ: 28.2, 43.2, 77.9, 126.8, 129.3, 145.3, 155.8, 167.2 ppm

¹³C DEPT (100 MHz, CHCl₃) δ: 52.1 (CH₂), 53.5 (CH₂), 25.4, 29.2 ppm

HRMS: Calculated Mass for $C_7H_{13}NO_4$: 175.08 Observed mass: [M+H] 176.09
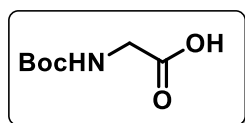
(tert-butoxycarbonyl)glycine (10) *L*-glycine (5g, 0.0666 mmol) was suspended in H_2O:1,4-dioxane (1:1) mixture to which NaOH (3.9g, 0.0999 mmol) followed by Boc anhydride (21.7g, 0.0999 mmol) was added and this reaction mixture was stirred at room temperature for 4 hr. after completion of the reaction indicated by TLC, partial reaction mixture was evaporated under vacuum and the remaining reaction mixture was acidified with 1N HCl and the crude mixture was extracted with DCM. A white coloured compound **10** was obtained on removing all the DCM under vacuum. This product was used for further reactions without purification. (8.7g, 75%)

¹H NMR (200 MHz, CHCl₃) δ: 1.46 (s, 9H), 3.95-3.97 (m, 2H) ppm

CHCl₃) **δ:** 1.46 (s, 9H), 3.95-3.97 (m, 2H) ppm
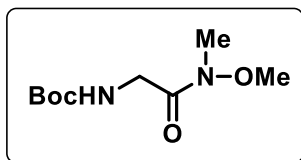
tert-butyl (2-(methoxy(methyl)amino)-2-oxoethyl)carbamate (11) compound **10** (0.2g, 1.14 mmol) was suspended in DCM and cooled to 0 ºC. To this, *N,O*-Dimethylhydroxylamine (0.174g, 1.52 mmol), EDC (0.212g, 1.36 mmol), TEA (0.266 mL, 1.9 mmol) and catalytic amount of DMAP was added and the reaction mixture was stirred at 0 ºC for 1hr after which the ice bath was removed and the reaction was stirred at room temperature for 10 hr. After the completion of reaction indicated by TLC, the reaction mixture was washed with 1N HCl and saturated NaHCO$_3$ and the DCM was removed under vacuum to give compound **11** as a white solid. (0.16g, 64%)

^{1}H NMR (200 MHz, CHCl$_3$) δ: 1.46 (s, 9H), 3.21 (s, 3H), 3.72 (s, 3H), 4.10 (br s, 2H) ppm
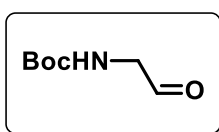
tert-butyl (2-oxoethyl)carbamate (12) LiAlH$_4$ (0.021g, 0.5733 mmol) was suspended in dry THF and cooled to -78 ºC. To this mixture compound **11** (0.1g, 0.4587 mmol) which was separately dissolved in dry THF was added *via* syringe in a drop wise manner. After 10 min, the reaction was quenched with 10% KHSO$_4$ and the reaction mixture was extracted with ether which gave compound **12** as colourless oil. (0.022g, 30%)
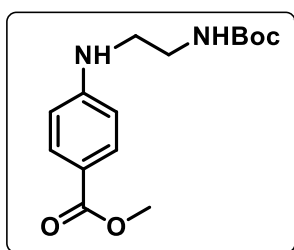
methyl 4-((2-((tert-butoxycarbonyl)amino)ethyl)amino)benzoate (13) to a solution of methyl 4-aminobenzoate (0.057g, 0.377 mmol) and acetic acid (0.022 mL, 0.377 mmol) in dry MeOH containing molecular sieves was added compound **12** (0.06 g, 0.377 mmol) under inert condition and the reaction was stirred at room temperature for 1 hr after which NaBH$_3$CN (0.023 g, 0.377 mmol) was added and the reation was stirred at room temperature for 10 hr. After completion of reaction, the solvent was evaporated under pressure and the residue was dissolved in ethyl acetate and washed with 1N HCl. Ethyl acetate layer was evaporated under pressure to give the crude compound **13** which was purified by column chromatography. (0.04 g, 36%)

^{1}H NMR (200 MHz, CHCl$_3$) δ: 1.43 (s, 9H), 3.28-3.46 (m, 3H), 3.83 (s, 3H), 6.51-6.55 (d, *J*=8Hz, 2H), 7.81-7.85 (d, *J*= 8Hz, 2H) ppm
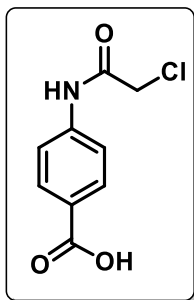
4-(2-chloroacetamido)benzoic acid (14) *para* amino benzoic acid (1g, 0.00729 mmol) was suspended in H_2O and cooled to 0 ℃. To this, NaOH (0.87g, 0.2189 mmol) was added followed by drop wise addition of chloroacetyl chloride (1.7mL, 0.2189 mmol). The reaction was allowed to come to room temperature and stirred for 5hr. After completion of reaction indicated by TLC, the reaction mixture was acidified wit 1N HCl and the white precipitate was filtered to give compound **14**. (1.1g, 70%)

¹H NMR (200 MHz, DMSO) δ: 4.30 (s, 2H), 7.68-7.73 (d, *J*=8 Hz, 2H), 7.90-7.94 (d, *J*= 8Hz, 2H), 10.63 (s, 1H) ppm
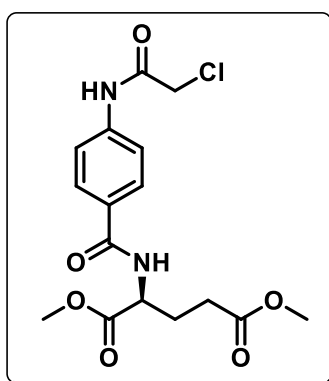
dimethyl (4-(2-chloroacetamido)benzoyl)-L-glutamate (15) to compound **14** (2g, 0.00862 mmol) in dry DMF was added compound **9** (4.5g, 0.0258 mmol), PyBOP (4.4g, 0.00862 mmol), and TEA (3.6 mL, 0.0258 mmol) and stirred at room temperature for 10hr. After the completion of reaction, the solvent was evaporated under vacuum and the crude mixture was purified by column chromatography to give compound **15** (1.7g, 50%)

¹H NMR (200 MHz, CHCl₃) δ: 2.14-2.37 (m, 2H), 2.46-2.55 (m, 2H), 3.67 (s, 3H), 3.79 (s, 3H), 4.22 (s, 2H), 7.65-7.69 (d, *J*=8 Hz, 2H), 7.82-7.87 (d, *J*=8 Hz, 2H), 8.41 (br s, 1H) ppm
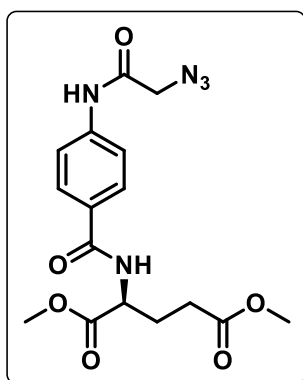
dimethyl (4-(2-azidoacetamido)benzoyl)-*L*-glutamate (16) to a solution of compound **15** (0.1 g, 0.270 mmol) in DMF, NaN_3 (0.175 g, 2.70 mmol) was added and the reaction was carried out at 65 ℃ with continuous stirring. After the completion of reaction, the solvent was evaporated under vacuum and the crude product was purified by column chromatography to yield the desired compound **16** (0.04 g, 40%)

¹H NMR (200 MHz, CHCl₃) δ: 2.11-2.36 (m, 2H), 2.44-2.53 (m, 2H), 3.65 (s, 3H), 3.76 (s, 3H), 4.11 (s, 2H), 4.72-4.82 (m, 1H), 7.59-7.64 (d, J=8 Hz, 2H), 7.76-7.80 (d, J= 8 Hz, 2H), 8.57 (br.s, 1H)
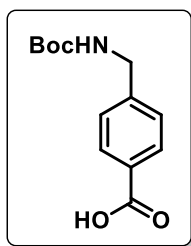
4-((((*tert*-butoxycarbonyl)amino)methyl)benzoic acid (18) Commercially available 4 (aminomethyl)benzoic acid (1g, 6.6mmol) was suspended in H$_2$O/1,2-dioxane (1:1). To this mixture, NaOH (0.396g, 9.9mmol) was added with vigorous stirring. Boc anhydride (1.7g, 7.9mmol) was added to this reaction mixture. The reaction was monitored by TLC. After completion of the reaction, the pH of the reaction mixture was adjusted to acidic with 2N HCl, and the precipitate was filtered under vaccum to yield a white coloured crude compound **6** (1.6g, 75%). This crude compound **18** was found to be pure by NMR and was used for subsequent reactions without further purification.

^{1}H NMR (400 MHz, DMSO-d6) δ: 1.39 (s, 9H), 4.18-4.20 (d, *J* = 6 Hz, 2H), 7.33-7.35 (d, *J* = 8 Hz, 2H), 7.88-7.90 (d, *J* = 8 Hz, 2H) ppm

^{13}C NMR (100 MHz, DMSO-d6) δ: 28.2, 43.2, 77.9, 126.89, 129.37, 145.37, 155.83, 167.23 ppm

^{13}C DEPT (100 MHz, DMSO-d6) δ: 42.96 (CH$_2$), 27.98, 126.64, 129.12 ppm

HRMS Calculated Mass for C$_{13}$H$_{17}$NO$_4$: 251.28 Observed mass: [M+Na] 274.10
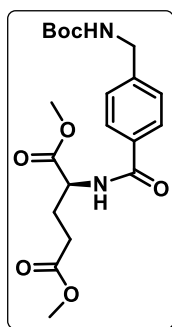
Dimethyl (4-((((*tert*-butoxycarbonyl)amino)methyl)benzoyl)-*L* glutamate (19) Compound **18** was dissolved DMF (0.5g,1.99mmol). To this reaction mixture TEA (0.83mL, 5.97mmol), PyBOP (1.03g, 1.99mmol) and compound **9** (1.04g, 5.97mmol) were added. The reaction mixture was stirred at room temperature for 10h. The solvent was evaporated *in vacuo* and the crude mixture was purified by column chromatography to give the product as a white solid (1.6g, 60%).

^{1}H NMR (400 MHz, CHCl$_3$) δ: 1.43 (s, 9H), 2.11-2.27 (m, 2H), 2.42-2.49 (m, 2H), 3.63 (s, 3H), 3.75 (s, 3H), 4.31 (s, 2H), 4.74-4.80 (m, 1H), 5.16 (br.s, 1H), 7.27-7.31 (d, *J* = 8 Hz, 2H), 7.73-7.75 (d, *J* = 8Hz, 2H) ppm

^{13}C NMR (100 MHz, CHCl$_3$) δ: 26.9, 28.2, 30.1, 44.0, 51.8, 127.3, 132.3, 155.8, 166.8, 172.3, 173.5 ppm

^{13}C DEPT (100 MHz, CHCl$_3$) δ: 27.0 (CH$_2$), 30.2 (CH$_2$), 44.2 (CH$_2$), 28.4, 51.9, 127.4 ppm

HRMS Calculated Mass for C$_{20}$H$_{28}$N$_2$O$_7$: 408.19 Observed mass: [M+H] 409.19
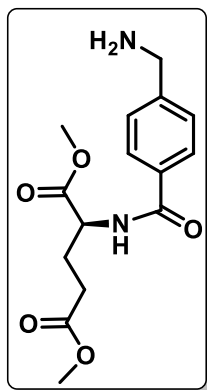
Dimethyl (4-(aminomethyl)benzoyl)-*L*-glutamate (20) To compound **19** (0.5g, 1.22 mmol), a 50% solution of TFA in DCM (5mL) was added with vigorous stirring at RT. After completion of the reaction, as monitored by TLC, the solvents were removed under vacuum and the crude compound was purified by column chromatography. The resultant TFA salt of the compound was neutralised with Et$_3$N and the crude compound was purified by column chromatography to give compound **20** as a brown oil (0.3g, 60%).

^{1}H NMR (400 MHz, MeOH-d4) δ: 2.22-2.43 (m, 2H), 2.58-2.62 (m, 2H), 3.74 (s, 3H), 3.84 (s, 3H), 4.31 (s, 2H), 4.76 (br.s, 1H), 7.67-7.69 (d, *J* = 8 Hz, 2H), 8.00-8.02 (d, *J* = 8 Hz, 2H) ppm

^{13}C NMR (100 MHz, MeOH-d4) δ: 31.3, 43.9, 53.0, 129.3, 130.2, 135.6, 138.4, 173.7, 175.0 ppm

^{13}C DEPT (100 MHz, MeOH-d4) δ: 29.0 (CH$_2$), 41.6 (CH$_2$) ppm

HRMS Calculated Mass for C$_{15}$H$_{20}$N$_2$O$_5$: 308.14 Observed mass: [M+H] 309.14
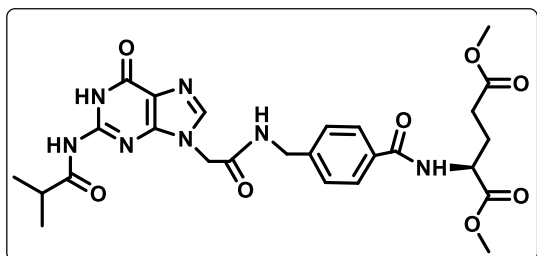
Dimethyl (4-((2-(2-(isobutylamino)-6-oxo-1,6-dihydro-9*H*-purin-9-yl)acetamido)methyl)benzoyl)-*L*-glutamate (21) To compound **5** (0.1g, 0.358mmol) in dry DMF was added HATU (0.136g, 0.3584mmol) and TEA (0.149mL, 0.752mmol) and the reaction was stirred at room temperature for 10 min. To this reaction mixture, compound **20** (0.099g, 0.322mmol) was added. The reaction mixture was further stirred and monitored by TLC at room temperature for 10h. After completion of the reaction, the solvent was removed under reduced pressure. The residue was taken up in DCM and washed with saturated NaHCO$_3$. The compound was extracted in DCM and the combined organic layers were dried over sodium sulphate and concentrated under vacuum The crude mixture was purified by column chromatography to yield pure **21** a white solid (0.045g, 22%).

^{1}H NMR (500 MHz, CDCl$_3$) δ: 1.26-1.28 (d, *J* = 6 Hz, 6H), 2.22-2.39 (m, 2H), 2.57-2.74 (m, 3H), 3.63 (s, 3H), 3.91 (s, 3H), 4.64-4.68 (m, 1H), 4.98-5.06 (m, 2H), 5.46-5.49 (m, 1H), 7.01-7.03 (d, *J* = 8 Hz, 2H), 7.67-7.68 (d, *J* = 8 Hz, 2H), 8.31 (s, 1H), 8.78 (s, 1H), 9.59 (s, 1H), 10.27 (s, 1H), 11.69 (s, 1H) ppm

¹³C NMR (125 MHz, CDCl₃) δ: 18.2, 19.5, 25.6, 30.6, 35.7, 51.7, 52.9, 113.8, 121.0, 127.7, 131.4, 140.8, 141.9, 147.6, 149.8, 156.9, 166.8, 180.2 ppm

¹³C DEPT (125 MHz, CDCl₃) δ: 25.6 (CH₂), 30.6 (CH₂), 42.8 (CH₂), 46.4 (CH₂), 68.6 (CH₂), 18.2, 19.6, 31.0, 35.7, 51.8, 52.8, 113.9, 127.6, 141.9 ppm

HRMS Calculated Mass for $C_{26}H_{31}N_7O_8$: 569.22 Observed mass: [M+H] 570.23

Appendix

Characterization	Page number
Compound 1- ^{1}H, ^{13}C NMR	122
Compound 1- DEPT, HRMS	123
Compound 2- ^{1}H, ^{13}C NMR	124
Compound 2- DEPT, HRMS	125
Compound 3- ^{1}H, ^{13}C NMR	126
Compound 3- DEPT, HRMS	127
Compound 4- ^{1}H, ^{13}C NMR	128
Compound 4- DEPT, HRMS	129
Compound 5- ^{1}H, ^{13}C NMR	130
Compound 5- DEPT NMR	131
Compound 6- ^{1}H NMR	131
Compound 7- ^{1}H NMR	132
Compound 8- ^{1}H NMR	132
Compound 9- ^{1}H, ^{13}C NMR	133
Compound 9- DEPT, HRMS	134
Compound 10- ^{1}H NMR	135
Compound 11- ^{1}H NMR	135
Compound 13- ^{1}H NMR	136
Compound 14- ^{1}H NMR	136
Compound 15- ^{1}H NMR	137
Compound 16- ^{1}H NMR	137
Compound 16- IR	138
Compound 18- ^{1}H NMR	138
Compound 18- ^{13}C, DEPT	139
Compound 18- HRMS	140
Compound 19- ^{1}H NMR	140
Compound 19- ^{13}C, DEPT	141
Compound 19- HRMS	142
Compound 20- ^{1}H NMR	142
Compound 20- ^{13}C, DEPT	143
Compound 20- HRMS	144
Compound 21- ^{1}H NMR	144
Compound 21- ^{13}C, DEPT	145
Compound 21- HRMS	146

iBu guanine latest.esp

¹H NMR Compound 1
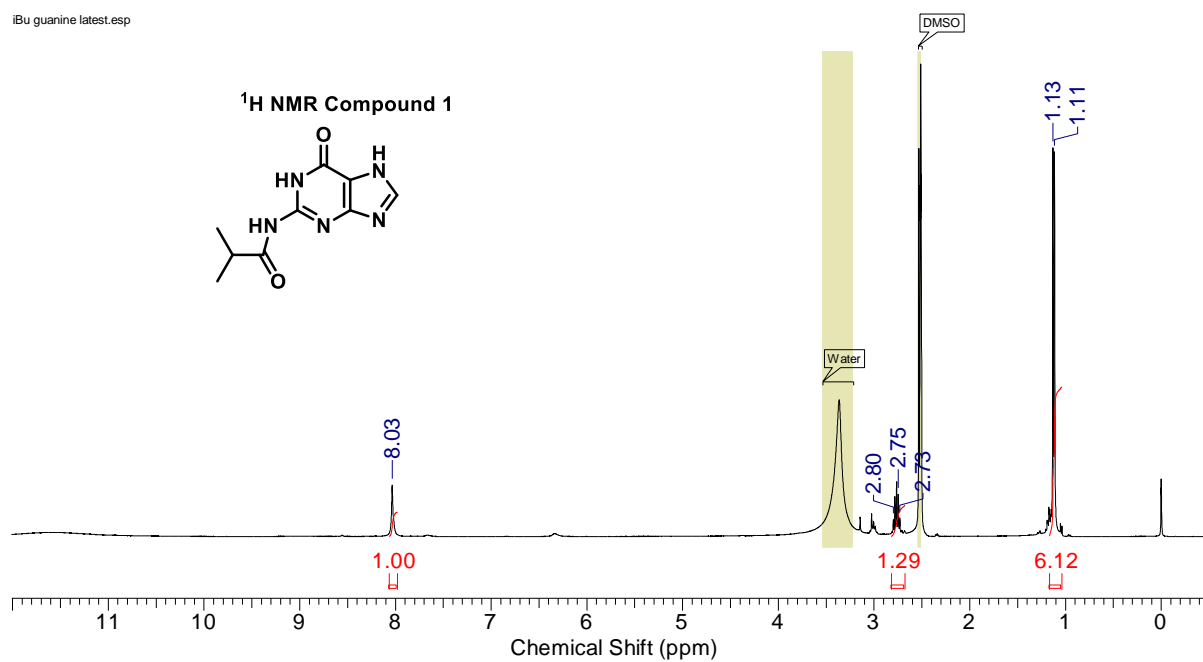
AV-500-20200819-155812-42827.003.001.1r.esp

¹³C NMR Compound 1
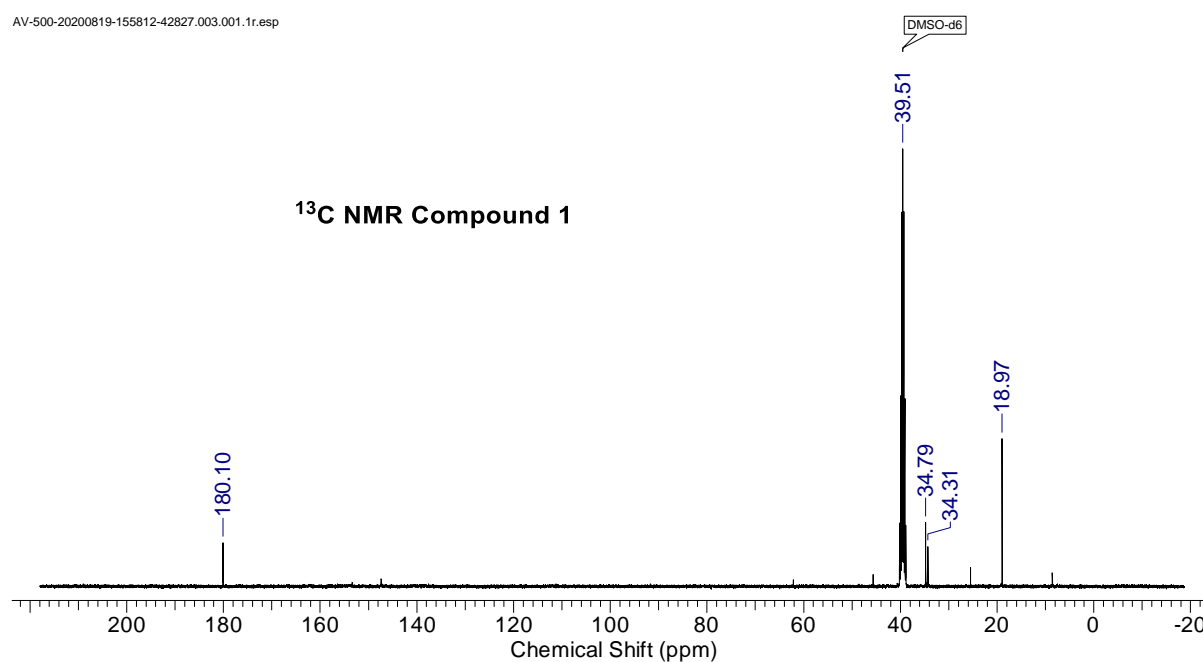
AV-500-20200819-155812-42827.002.001.1r.esp

DEPT Compound 1

19.41

35.23

46.07

Chemical Shift (ppm)
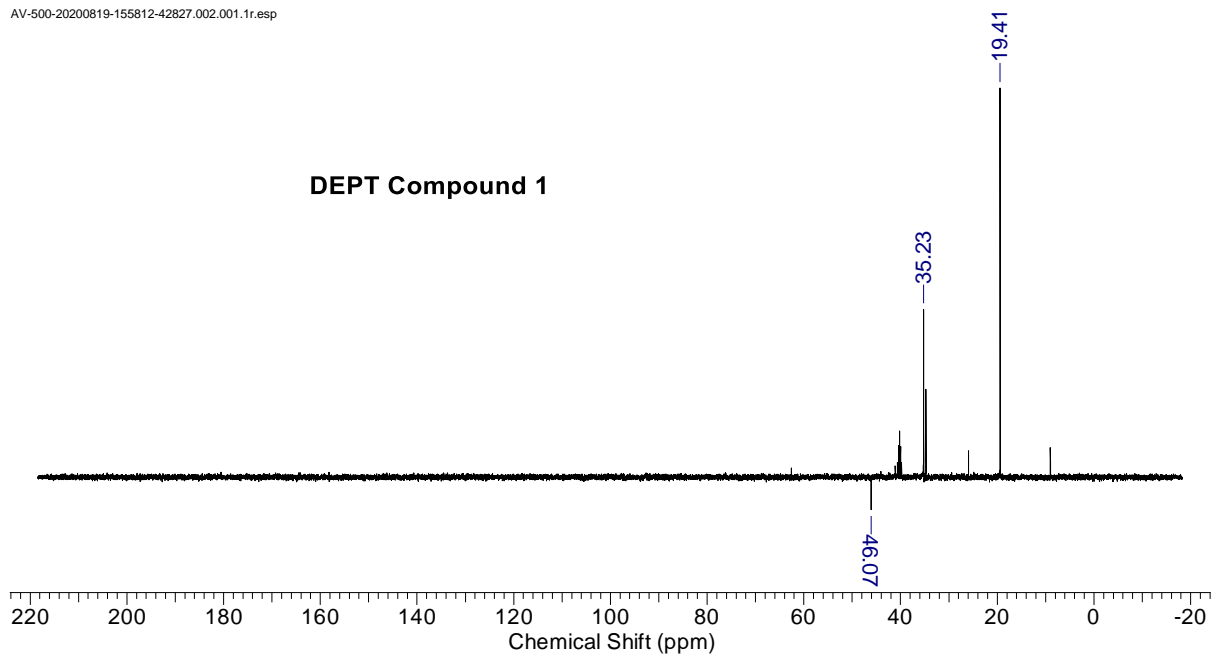
A-1 #252   RT: 1.42   AV: 1   NL: 2.34E8
T: FTMS + p ESI Full ms [100.0000-1500.0000]

222.0986
R=76407
$C_9 H_{12} O_2 N_5 = 222.0986$
0.0241 ppm

HRMS Compound -1
Calculated Mass: 221.09
Observed Mass (M+H): 222.09

152.0566
R=91507
$C_5 H_6 O N_5 = 152.0567$
-0.8027 ppm

260.0541
R=70007

292.1401
R=66707

327.0077
R=61606

365.1051
R=58207

395.3418
R=57007

465.1718
R=52507

516.4260
R=45807

572.4886
R=44607

m/z
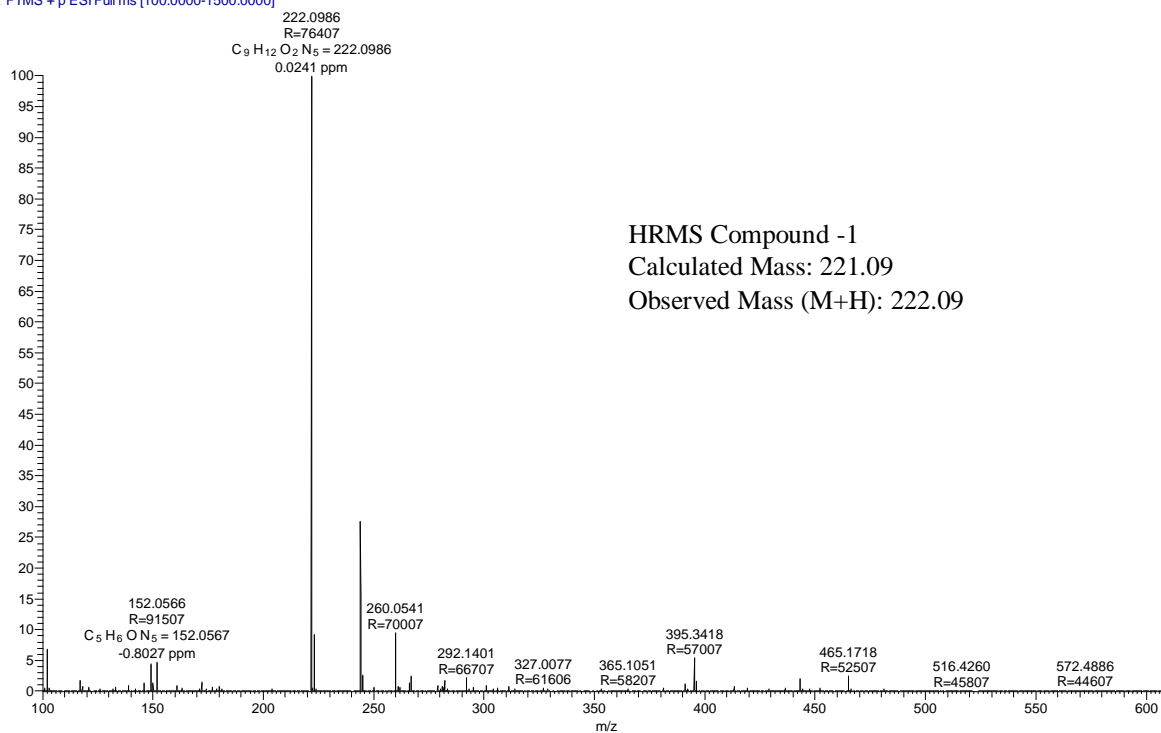
N7 ester.esp
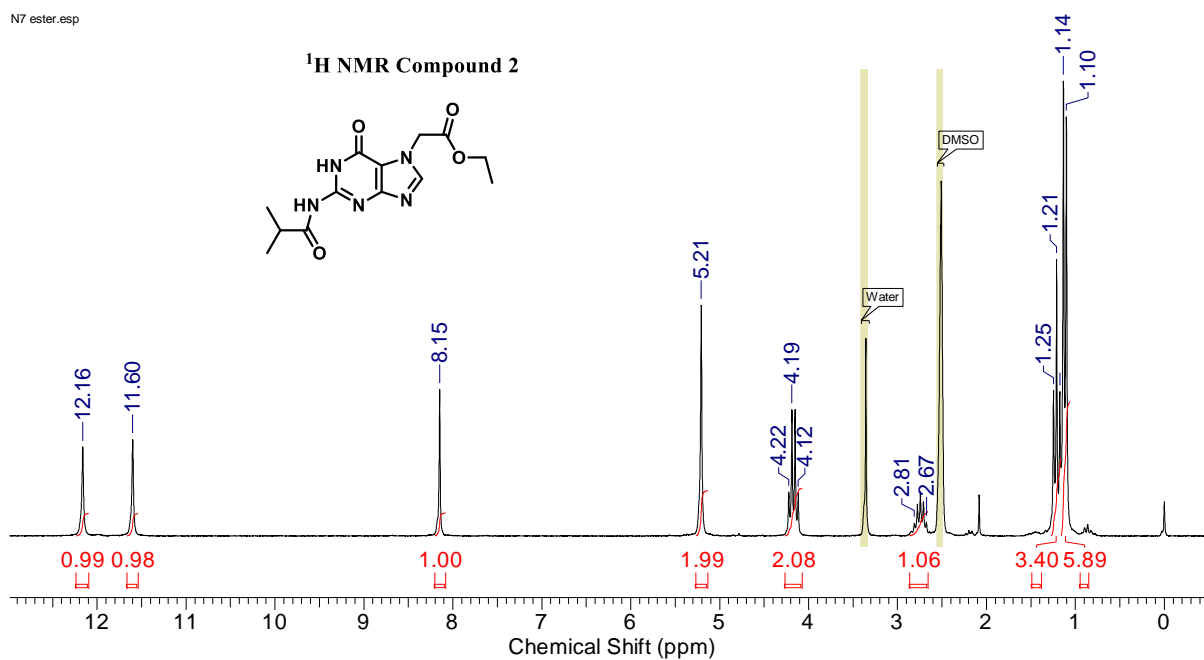
¹H NMR Compound 2

AV-400-20200824-115052-11826.003.001.1r.esp

¹³C NMR Compound 2
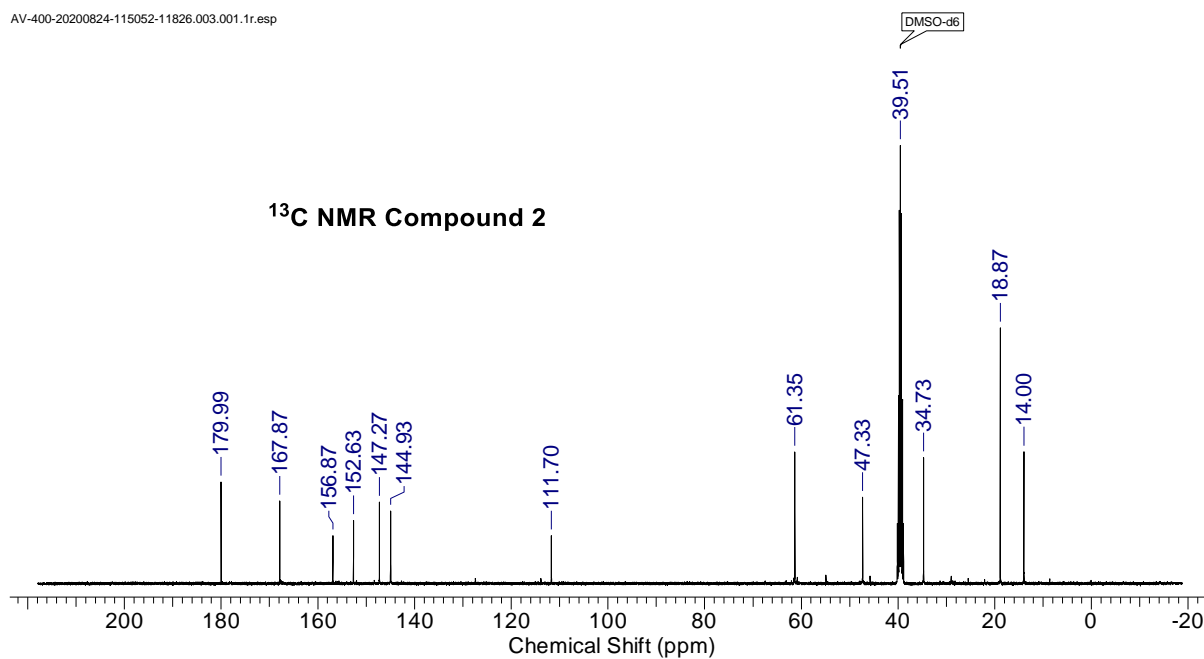
N7 ester DEPT.esp
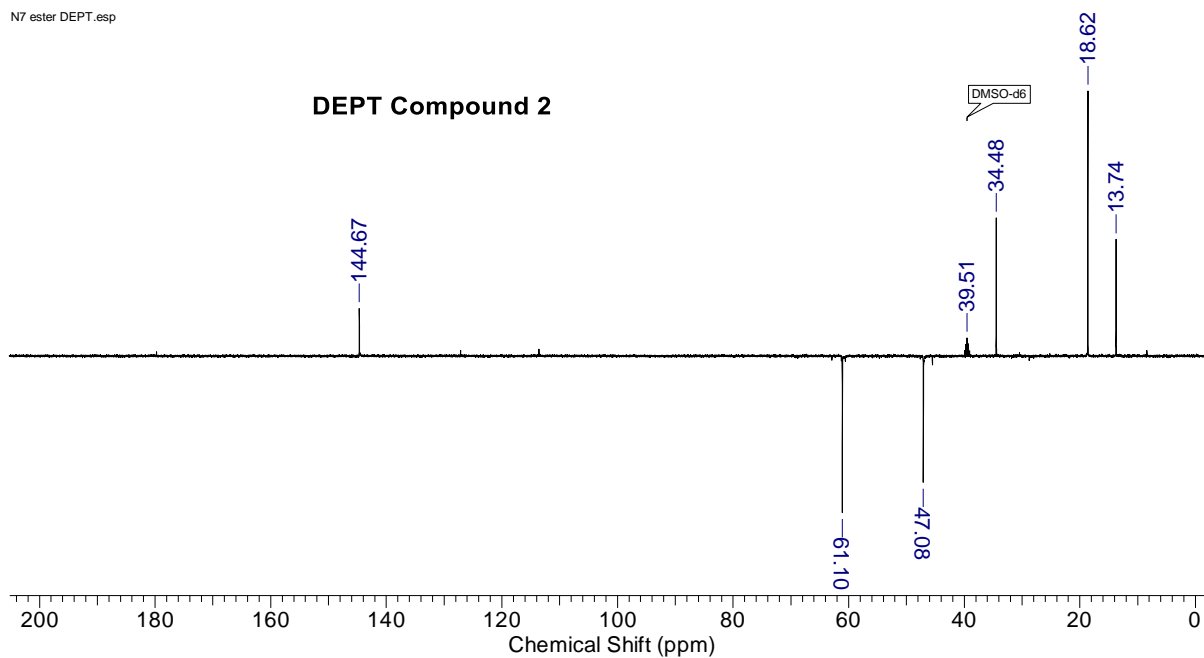
DEPT Compound 2

B-2_191218115032 #257 RT: 1.43 AV: 1 NL: 1.69E9
T: FTMS + p ESI Full ms [100.0000-1500.0000]

HRMS Compound -2
Calculated Mass: 307.13
Observed Mass (M+H): 308.13
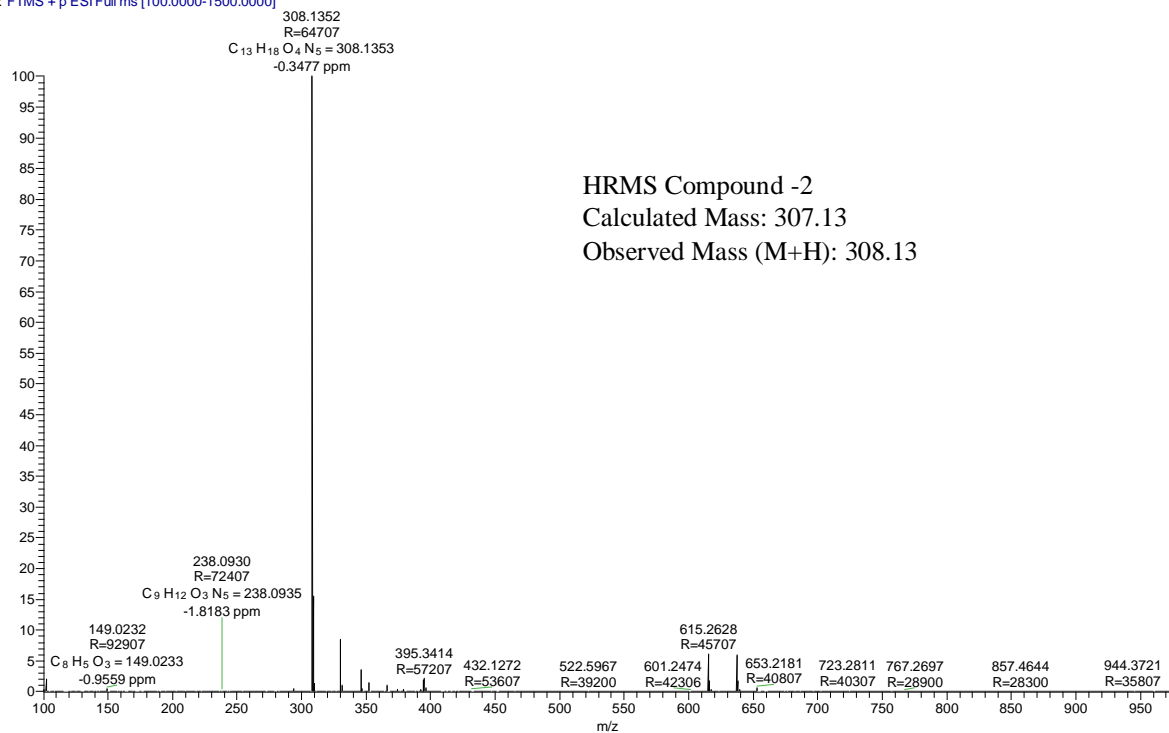
N9 ester.esp
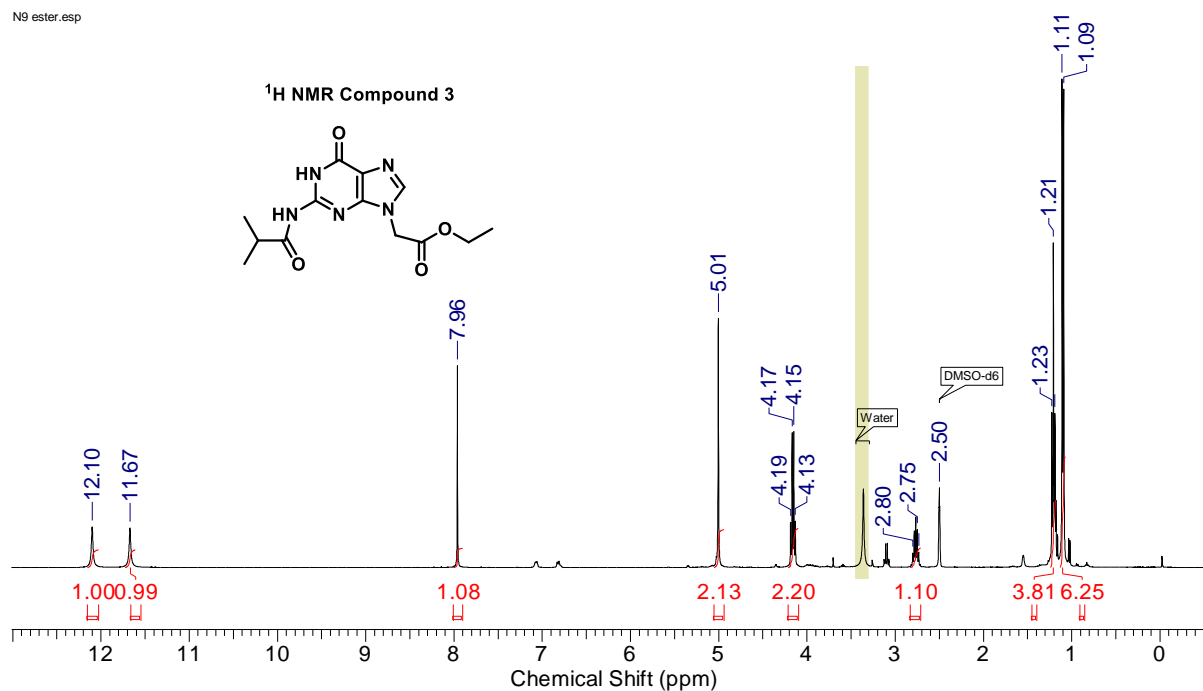
¹H NMR Compound 3

AV-500-20200824-115128-11826.003.001.1r.esp

¹³C NMR Compound 3

AV-500-20200824-115128-11826.002.001.1r.esp

DEPT Compound 3

—140.76

—62.02

—44.80

—35.19

—19.35

—14.50

Chemical Shift (ppm)

220 200 180 160 140 120 100 80 60 40 20 0 -20

C-3 #248 RT: 1.41 AV: 1 NL: 9.55E7
T: FTMS + p ESI Full ms [100.0000-1500.0000]

308.1353
R=65107
$C_{13} H_{18} O_4 N_5 = 308.1353$
-0.1497 ppm

149.0233
R=93307
$C_8 H_5 O_3 = 149.0233$
-0.4439 ppm

353.1929
R=60307

419.3153
R=54307

516.4261
R=47107

600.5203
R=45007

684.6137
R=42007

733.5743
R=39507

782.7219
R=33605

m/z

100 150 200 250 300 350 400 450 500 550 600 650 700 750 800

HRMS Compound -3
Calculated Mass: 307.13
Observed Mass (M+H): 308.13

N7 acid.esp

¹H NMR Compound 4

N7 acid.esp

¹³C NMR Compound 4
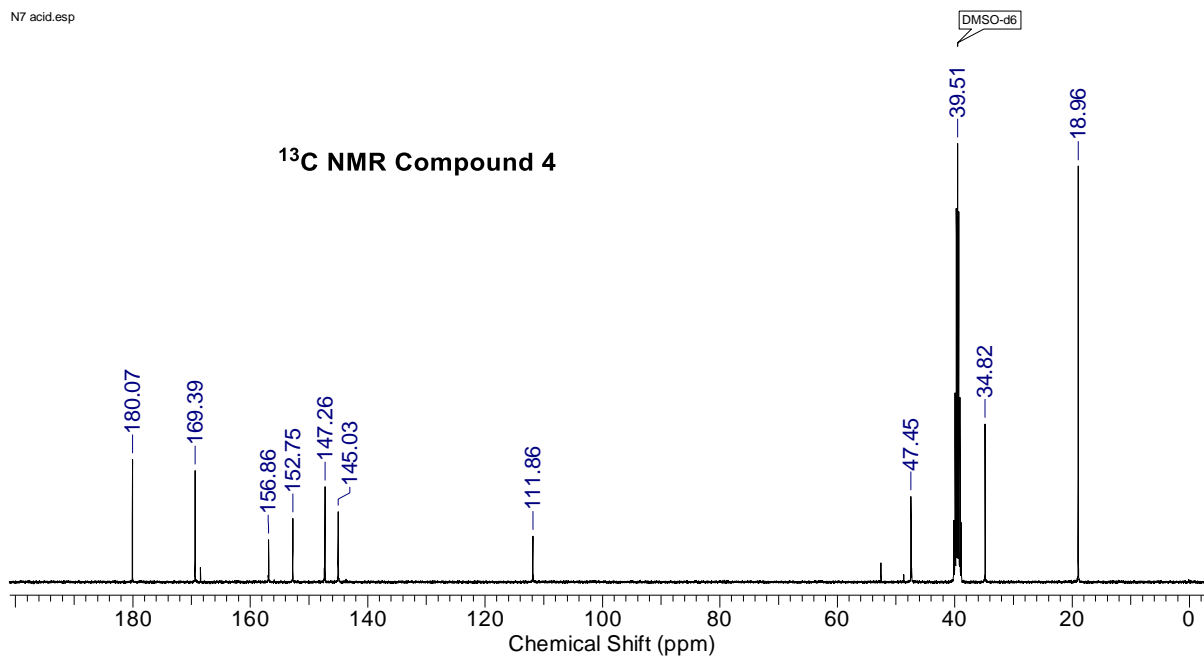
N7 acid DEPT.esp
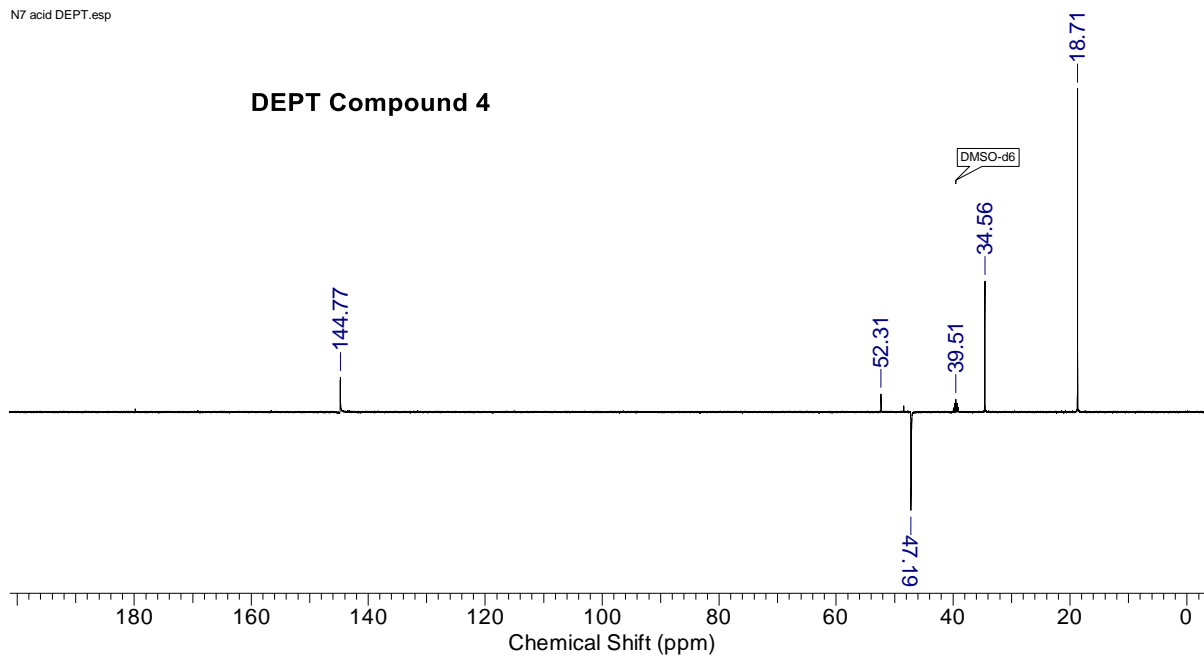
DEPT Compound 4

HC-C #256 RT: 1.90 AV: 1 NL: 3.24E8
T: FTMS + p ESI Full ms [100.0000-1500.0000]

280.1046
R=68407
$C_{11} H_{14} O_4 N_5 = 280.1040$
1.9754 ppm

210.0622
R=77907
$C_7 H_8 O_3 N_5 = 210.0622$
0.3568 ppm

116.9863
R=103102

143.0017
R=94907

336.0614
R=61407

376.0929
R=55702

406.1032
R=53807

457.5986
R=41104

HRMS Compound -4
Calculated Mass: 279.26
Observed Mass (M+H): 280.10

N9 acid.esp

¹H NMR Compound 5

N9 acid.esp

¹³C NMR Compound 5

N9 acid DEPT.esp

DEPT Compound 5

18.65

DMSO-d6

140.15

34.48

44.29
45.42

180   160   140   120   100   80   60   40   20   0
Chemical Shift (ppm)

compound 7.esp

Water

¹H NMR Compound 6

NHCbz

O   OH

DMSO

5.21

10.38

8.07
8.05
7.72
7.69
7.45
7.36

0.89   1.75 1.81 4.73   2.00

11   10   9   8   7   6   5   4   3   2
Chemical Shift (ppm)

Atish.008.esp

¹H NMR Compound 7

Atish.007.esp

¹H NMR Compound 8

Ediester.esp

¹H NMR Compound 9

E dister carbon.esp

¹³C NMR Compound 9

E dister dept.esp

DEPT Compound 9

—29.92
—25.44

53.57
—52.12

180   160   140   120   100   80   60   40   20   0
Chemical Shift (ppm)

H-9 #140 RT: 0.80 AV: 1 NL: 1.60E9
T: FTMS + p ESI Full ms [100.0000-1500.0000]
176.0915
R=86203
$C_7 H_{14} O_4 N$ = 176.0917
-1.1184 ppm

HRMS Compound -9
Calculated Mass: 175.08
Observed Mass (M+H): 176.09

100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0

144.0653
R=95007

117.0740
R=102102

160.0683
R=86602

198.0734
R=80307

235.2052
R=61800

282.2786
R=65007

318.2399
R=57602

359.3146
R=54400

391.2837
R=55407

457.2761
R=45102

502.3731
R=34900

533.4521
R=35600

566.3667
R=41304

600.5184
R=37900

100   150   200   250   300   350   400   450   500   550   600
m/z

Atish.010.esp

¹H NMR Compound 10

BocHN—CH₂—COOH (structure: BocHN connected to CH₂–C(=O)–OH)

1.46

CHLOROFORM-d

7.27

3.97
3.95

1.99   9.00

Chemical Shift (ppm)

11 10 9 8 7 6 5 4 3 2 1 0

Atish.011.esp

TMS

¹H NMR Compound 11

0.00

Structure: BocHN—CH₂—C(=O)—N(OMe)(Me)

1.46

3.72

3.21

CHLOROFORM-d

7.27

4.10

1.90 3.07 3.00   10.98

Chemical Shift (ppm)

11 10 9 8 7 6 5 4 3 2 1 0

Atish.013.esp

¹H NMR Compound 13

¹H NMR Compound 13

CHLOROFORM-d

7.85
7.81
7.27
6.55
6.51
4.16
4.05
3.83
3.46
3.28
1.43

1.71 1.30 2.99 3.54 8.59

Chemical Shift (ppm)

compound 2.esp

¹H NMR Compound 14

DMSO-d6

12.74
10.63
7.94
7.90
7.73
7.68
4.30
2.50

0.71 0.96 2.00 2.01 2.03

Chemical Shift (ppm)

Atish.015.esp

¹H NMR Compound 15

CHLOROFORM-d

8.41
7.87 7.82
7.69
7.65
7.27
4.85 4.82
4.76
4.22
3.79
3.67
2.55 2.52
2.46
2.37
2.25
2.14

0.88 2.00 1.97   0.97 1.98 3.06 2.97 1.94 2.22

Chemical Shift (ppm)

11 10 9 8 7 6 5 4 3 2 1 0

Atish.016.esp

¹H NMR Compound 16

CHLOROFORM-d

8.57
7.80
7.76
7.64
7.59
7.27
4.82
4.79
4.72
4.11
3.76
3.65
2.53
2.49
2.44
2.36
2.28
2.11

1.02 2.00 2.09   1.10 2.06 3.11 3.46 2.10 2.49

Chemical Shift (ppm)

11 10 9 8 7 6 5 4 3 2 1 0

IR Compound-16

AV-400-20200818-110716-42827.001.001.1r.esp

¹H NMR Compound 18
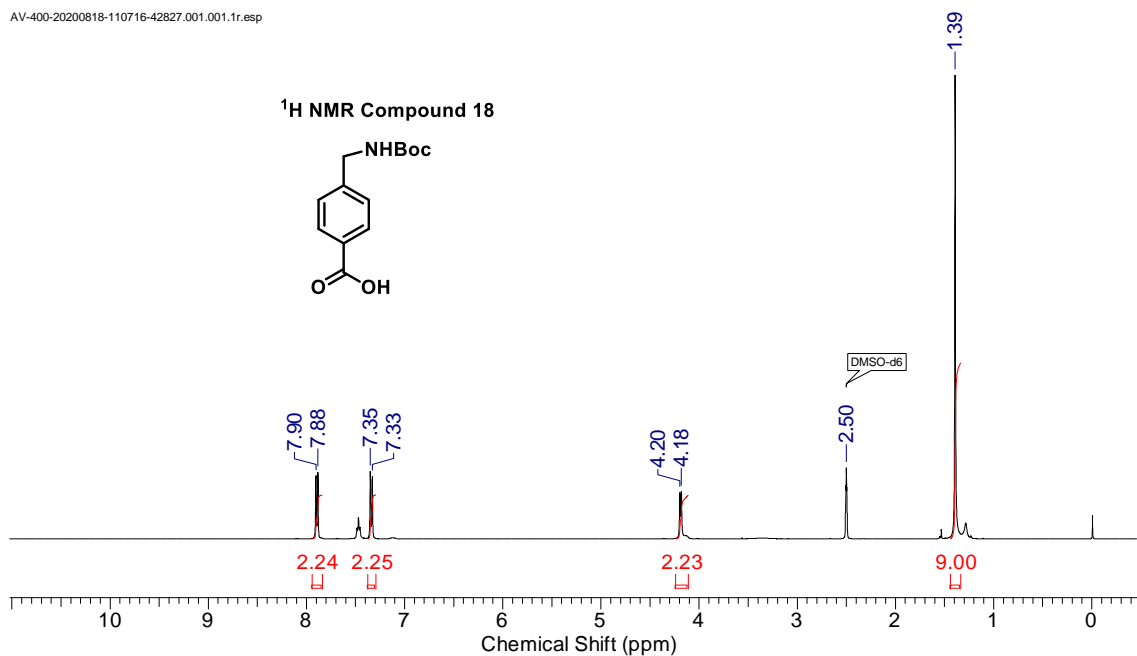
AV-400-20200818-110716-42827.003.001.1r.esp

DMSO-d6

¹³C NMR Compound 18

167.23
155.83
145.37
129.37
126.89
77.96
43.21
39.51
28.23

Chemical Shift (ppm)
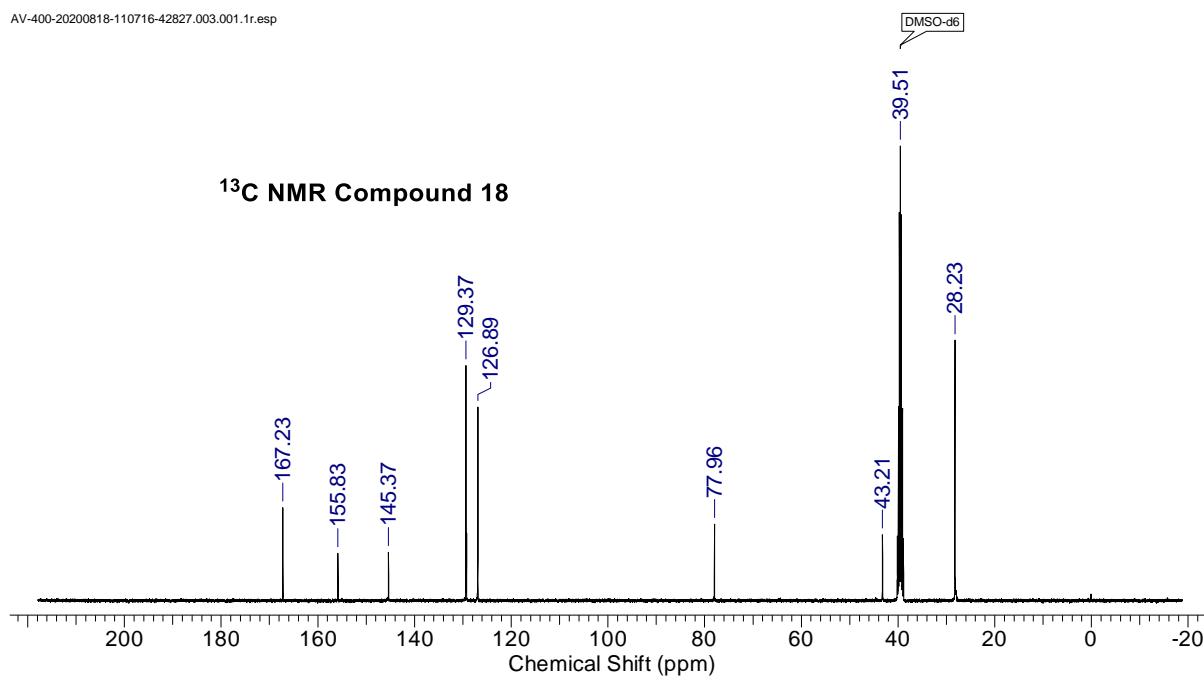
AV-400-20200818-110716-42827.002.001.1r.esp

DMSO-d6

DEPT Compound 18

129.12
126.64
27.98
39.51
42.96

Chemical Shift (ppm)

E-5 #257 RT: 1.47 AV: 1 NL: 1.99E7
T: FTMS + p ESI Full ms [100.0000-1500.0000]

274.1048
R=68107
C₁₃H₁₇O₄NNa = 274.1050
-0.5521 ppm

395.3421
R=57007

HRMS Compound -18
Calculated Mass: 251.12
Observed Mass (M+Na): 274.10

218.0423
R=76907
C₉H₉O₄NNa = 218.0424
-0.2343 ppm

296.0867
R=65807

336.0753
R=61507

367.3359
R=55904

413.2662
R=55507

441.2975
R=53007

469.3289
R=51607

511.4721
R=47307

541.1862
R=47207

569.1852
R=45507

m/z

Boc aminomethyl glutamate diester.esp

¹H NMR Compound 19

3.75
3.63
1.43

NHBoc

CHLOROFORM-d

7.75
7.73
7.31
7.27

5.16
4.80
4.74
4.31

2.49
2.42
2.27
2.14
2.11

2.01 2.25   0.98 1.01 1.96 3.00 2.99   2.12 2.73 9.13

Chemical Shift (ppm)

Boc amino methyl diester carbon.esp

¹³C NMR Compound 19

- 173.58
- 172.38
- 166.84
- 155.88
- 132.30
- 127.35

CHLOROFORM-d
- 77.00
- 79.58
- 51.82
- 44.09
- 30.14
- 28.27
- 26.95

Chemical Shift (ppm)

Boc amino methyl diester DEPT.esp

DEPT Compound 19

- 127.49
- 51.97
- 44.23
- 28.42
- 30.29
- 27.09

Chemical Shift (ppm)

F-6 #255 RT: 1.45 AV: 1 NL: 5.98E8
T: FTMS + p ESI Full ms [100.0000-1500.0000]

409.1967
R=55807
C$_{20}$ H$_{29}$ O$_7$ N$_2$ = 409.1969
-0.5911 ppm

HRMS Compound - 19
Calculated Mass: 408.19
Observed Mass (M+H): 409.19

178.0493
R=81502
C$_9$ H$_8$ O$_3$ N = 178.0499
-2.9358 ppm

282.2785
R=66307
C$_{20}$ H$_{28}$ N = 282.2216
201.4924 ppm

454.2537
R=53307

477.1707
R=49907

551.1268
R=46302

600.5185
R=43507

642.6010
R=35105

684.6124
R=39707

738.6580
R=34800

780.3499
R=29100

m/z

Aminomethyl glutamate diester.esp

^{1}H NMR Compound 20

8.02
8.00
7.69
7.67

Methanol

4.77
4.75
4.73

4.31

3.84
3.74

METHANOL-d4

3.31

2.62
2.60
2.58
2.43
2.39
2.38
2.22
2.20

2.49 2.49   1.16 2.32 3.00 2.42   2.34 2.50

Chemical Shift (ppm)

Aminomethyl glutamate diester.esp

¹³ C NMR Compound 20

METHANOL-d4

49.15

130.24
129.38

175.05
173.76

138.43
135.60

53.05
43.94

31.33

Chemical Shift (ppm)

180   160   140   120   100   80   60   40   20   0

Aminomethyl glutamate diester DEPT.esp

127.96
127.09

DEPT Compound 20

50.77

41.65

29.03
25.04

Chemical Shift (ppm)

180   160   140   120   100   80   60   40   20   0
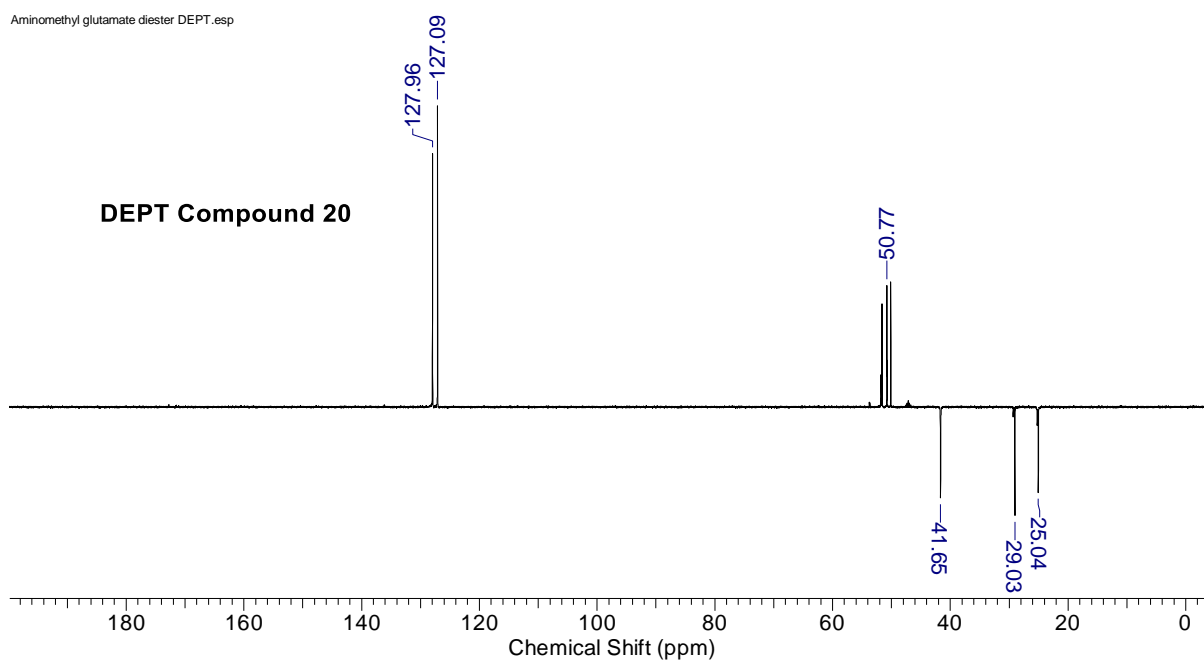
G-7 #140   RT: 0.80   AV: 1   NL: 3.06E8
T: FTMS + p ESI Full ms [100.0000-1500.0000]

309.1444
R=64607
C$_{15}$ H$_{21}$ O$_5$ N$_2$ = 309.1445
-0.3868 ppm

HRMS Compound - 20
Calculated Mass: 308.14
Observed Mass (M+H): 309.14

393.0964
R=56607
C$_{19}$ H$_{16}$ O$_7$ N Na = 393.0819
36.9046 ppm

639.2634
R=44507

221.1646
R=76007
C$_{13}$ H$_{21}$ O N$_2$ = 221.1648
-1.0310 ppm

436.3417
R=51707

516.4253
R=45602

572.4875
R=43207

617.2816
R=45307

684.6125
R=41107

727.4517
R=34400

777.4638
R=33600

m/z
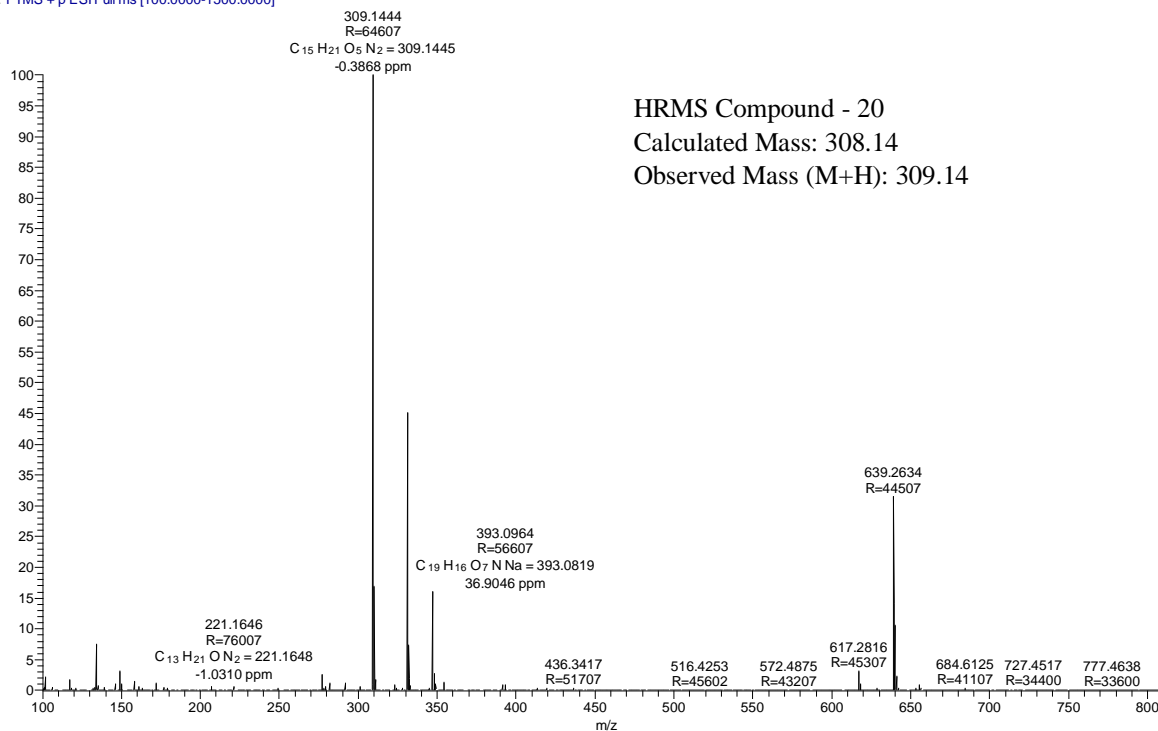
AV-400-20191218-123418-42827.001.001.1r.esp

^{1}H NMR Compound 21

CHLOROFORM-d

11.69
10.27
9.59
8.78
8.31
7.68
7.67
7.27
7.03
7.01
5.49
5.46
5.06
4.98
4.68
4.65
4.64
3.91
3.63
2.74
2.69
2.57
2.39
2.22
1.28
1.26

0.79   0.94   0.88   0.97 0.95   2.00   1.95   0.89 1.91 0.99 2.96 3.15   3.23 2.33   6.86

Chemical Shift (ppm)

AV-400-20191218-123418-42827.003.001.1r.esp

CHLOROFORM-d

77.00

^{13}C NMR Compound 21

180.24
166.81
156.93
149.83
147.67
141.95
140.88
131.40
127.76
121.01
113.89
52.98
51.76
35.71
30.65
25.63
19.59
18.25

Chemical Shift (ppm)
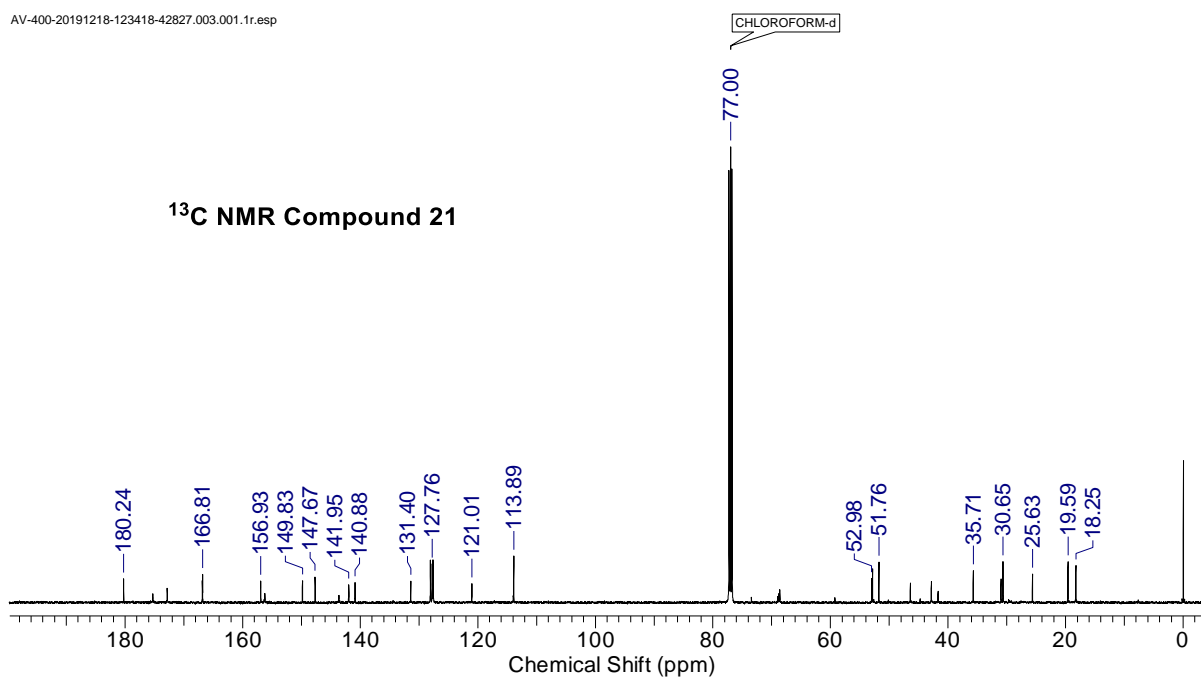
AV-400-20191218-123418-42827.002.001.1r.esp

128.10
127.63
113.91
53.02
52.89
51.81
35.74
31.02
19.63
18.29

TMS

DEPT Compound 21

141.99

68.63
46.45
42.87
30.68
25.66

Chemical Shift (ppm)
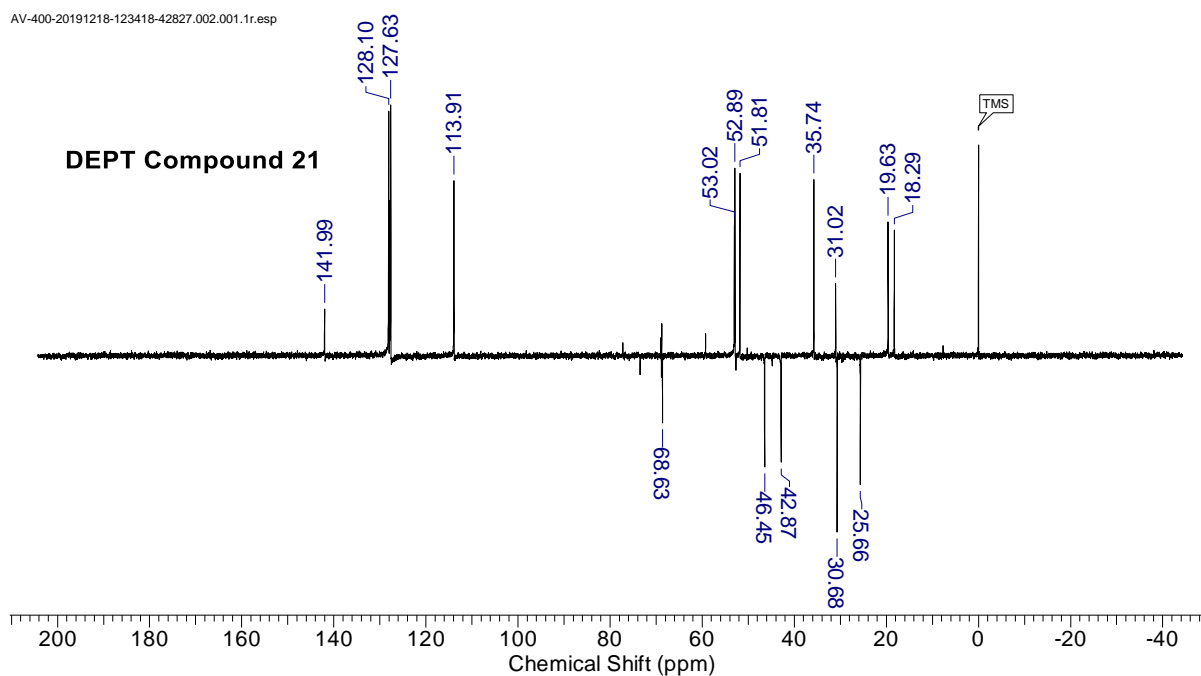
D-4_191218120317 #239 RT: 1.37 AV: 1 NL: 2.90E7
T: FTMS + p ESI Full ms [100.0000-1500.0000]

570.2313
R=47507
$C_{26}H_{32}O_8N_7$ = 570.2307
1.0081 ppm

HRMS Compound – 21
Calculated Mass: 569.22
Observed Mass (M+H): 570.23

593.2161
R=46102

282.2790
R=67707

391.2842
R=56807

436.3423
R=53007

516.4266
R=48007

615.2889
R=45007

684.6136
R=42107

738.6598
R=37507

808.7388
R=36005

898.7834
R=30907

m/z
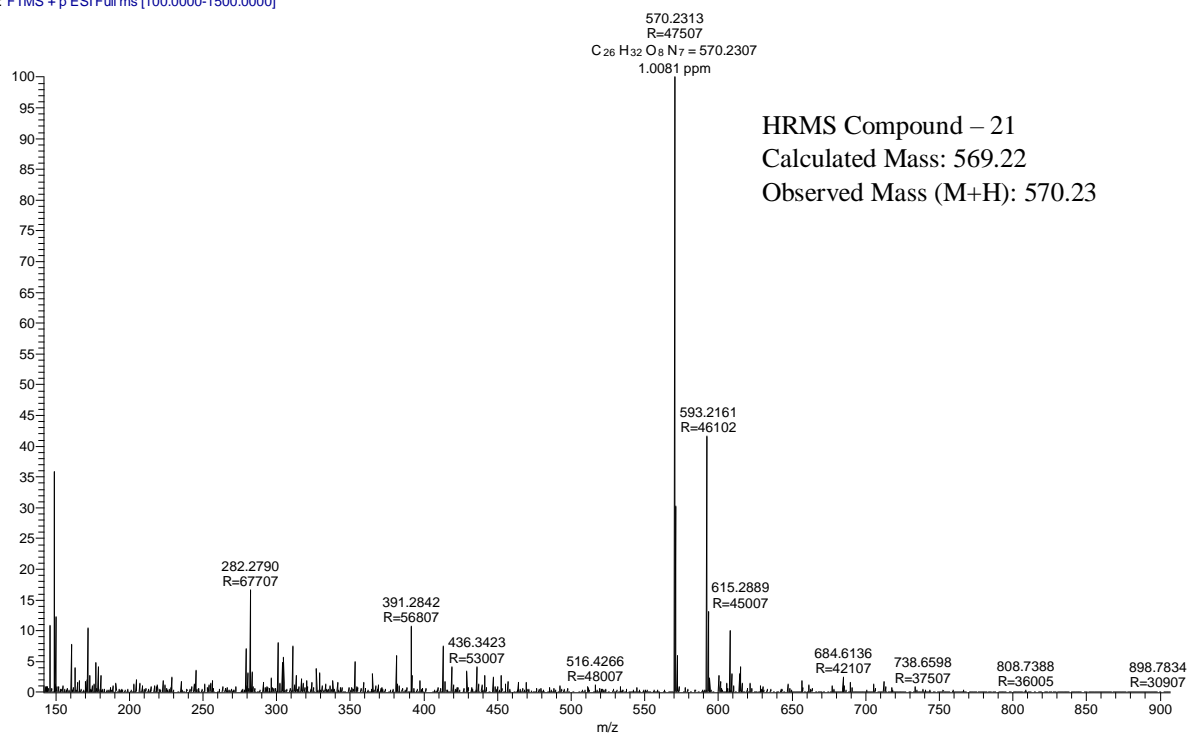
Appendix

(1)   Zhang, L.; Lu, Y.; Ye, Y.-h.; Yang, S.-h.; Tu, Z.-c.; Chen, J.; Wang, H.; Wang, H.-h.; Yuan, T., *J. Agric. Food Chem.* **2019**, *67*, 236-246

(2)   Xu, W.; Wu, J.; An, Y.; Xiao, C.; Hao, F.; Liu, H.; Wang, Y.; Tang, H. *J., Proteome Res.* **2012**, *11*, 3423-3435

(3)   Fu, Z.; Gilbert, E. R.; Liu, D., *Curr Diabetes Rev.* **2013**, *9,* 25-53

(4)   Khawandanah, J., *Nutr. Diabetes,* **2019**, *9*, 33

(5)   Skyler, J. S., *J. Med. Chem.,* **2004**, *47,* 4113-4117

(6)   Viljoen, A.; Sinclair, A. J., *Med. Clin. North Am.,* **2011**, *95,* 615-629

(7)   Aekplakorn, W.; Chariyalertsak, S.; Kessomboon, P.; Sangthong, R.; Inthawong, R.; Putwatana, P.; Taneepanichskul, S., *Diabetes Care.* **2011**, *34, 1980-1985*

(8)   Bharatam, P. V.; Patel, D. S.; Adane, L.; Mittal, A.; Sundriyal, S., *Curr. Pharm. Des.* **2007**, *13,* 3518-3530

(9)   Pàmies, L. G., Universitat Rovira I Virgili Thesis, 2011

(10)   Perera, N.; Ritchie, R. H.; Tate, M., *ACS Pharmacol. Transl. Sci.* **2020**, *3*, 11-20

(11)   DeCoster, V. A., *Health Soc. Work.,* **2001**, *26,* 26-37

(12)   Meneses, M. J.; Silva, B. M.; Sousa, M.; Sai, R.; Oliveira, P. F.; Alves, M. G. *Curr. Pharm. Des.* **2015**, *21,* 3606-3620

(13)   Kato, A.; Hayashi, E.; Miyauchi, S.; Adachi, I.; Imahori, T.; Natori, Y.; Yoshimura, Y.; Nash, R. J.; Shimaoka, H.; Nakagome, I.; Koseki, J.; Hirono, S.; Takahata, H., *J. Med. Chem.,* **2012**, *55*, 10347-10362

(14)   Modak, M.; Dixit, P.; Londhe, J.; Ghaskadbi, S.; Devasagayam, T. P., *J Clin Biochem Nutr.,* *2007, 40, 163-73*

(15)   *Classification of diabetes mellitus*; Geneva: World Health Organization, 2019

(16)   Wilcox, G., *The Clinical Biochemist Reviews,* **2005**, *26*, 19-39

(17)   Quianzon, C. C.; Cheikh, I., *J Community Hosp Intern Med Perspect.,* **2012**, *2*, 18701

(18)   Vinther, T. N.; Norrman, M.; Ribel, U.; Huus, K.; Schlein, M.; Steensgaard, D. B.; Pedersen, T.; Pettersson, I.; Ludvigsen, S.; Kjeldsen, T.; Jensen, K. J.; Hubalek, F., *Protein Sci.* **2013**, *22,* 296-305

(19)   Hua, Q. *Protein Cell,* **2010**, *1,* 537-51

(20) Simoni, R. D.; Hill, R. L.; Vaughan, M., *J. Biol. Chem.* **2002**, *277*, e15

(21) Mao, R.; Chen, Y.; Chi, Z.; Wang, Y., *Appl. Microbiol. Biotechnol.* **2019**, *103*, 8737-8751

(22) Huus, K.; Havelund, S.; Olsen, H. B.; van de Weert, M.; Frokjaer, S., *Biochemistry,* **2005**, *44*, 11171-11177

(23) Fujita-Yamaguchi, Y.; Choi, S.; Sakamoto, Y.; Itakura, K., *J. Biol. Chem.* **1983**, *258,* 5045-5049

(24) Jellinger, P. S., *Clin. Cornerstone,* **2007**, *8,* S30-42

(25) Kahn, C. R., *Annu. Rev. Med.,* **1985**, *36,* 429-51

(26) Zaykov, A. N.; Mayer, J. P.; Gelfanov, V. M.; DiMarchi, R. D., *ACS Chem. Biol.,* **2014**, *9,* 683-691

(27) Hartman, I., *Clin. Med. Res.* **2008**, *6*, 54-67

(28) Son, S.; Chae, S. Y.; Kim, C. W.; Choi, Y. G.; Jung, S. Y.; Lee, S.; Lee, K. C., *J. Med. Chem.,* **2009**, *52*, 6889-6896

(29) Chaudhury, A.; Duvoor, C.; Reddy Dendi, V. S.; Kraleti, S.; Chada, A.; Ravilla, R.; Marco, A.; Shekhawat, N. S.; Montales, M. T.; Kuriakose, K.; Sasapu, A.; Beebe, A.; Patil, N.; Musham, C. K.; Lohani, G. P.; Mirza, W., *Front. endocrinol.,* **2017**, *8*, 6-6.

(30) Washburn, W. N. *J., Med. Chem.* **2009**, *52*, 1785-1794

(31) Katherine E. Nori Janosz, K. C. Z., Wendy M. Miller, Peter A. McCullough, *Ther. Adv. Cardiovasc. Dis.,* **2009**, *3*, 387-395

(32) Glotfelty, E. J.; Delgado, T. E.; Tovar-y-Romo, L. B.; Luo, Y.; Hoffer, B. J.; Olson, L.; Karlsson, T. E.; Mattson, M. P.; Harvey, B. K.; Tweedie, D.; Li, Y.; Greig, N. H., *ACS Pharmacol. Transl. Sci.,* **2019**, *2*, 66-91

(33) Ndisang, J. F.; Vannacci, A.; Rastogi, S., *J. Diabetes Res.*, *2017*, 1478294

(34) Tsujihara, K.; Hongu, M.; Saito, K.; Kawanishi, H.; Kuriyama, K.; Matsumoto, M.; Oku, A.; Ueta, K.; Tsuda, M.; Saito, A., *J. Med. Chem.*, **1999**, *42*, 5311-5324

(35) Hsu, W.-H.; Lee, B.-H.; Hsu, Y.-W.; Pan, T.-M., *J. Agric. Food Chem.*, **2013**, *61*, 6873-6879

(36) Bierhaus, A.; Humpert, P. M.; Morcos, M.; Wendt, T.; Chavakis, T.; Arnold, B.; Stern, D. M.; Nawroth, P. P., *J. Mol. Med. (Berl).,* **2005**, *83,* 876-86

(37) Singh, V. P.; Bali, A.; Singh, N.; Jaggi, A. S., *Korean J. Physiol. Pharmacol.,*

2014, *18*, 1-14

(38) Vlassara, H.; Palace, M. R., *J. Intern. Med.* **2002**, *251,* 87-101

(39) Singh, R.; Barden, A.; Mori, T.; Beilin, L. *Diabetologia,* **2001**, *44,* 129-146

(40) Gkogkolou, P.; Bahm, M., *Dermatoendocrinol.,* **2012**, *4,* 259-70

(41) Lund, M. N.; Ray, C. A., *J. Agric. Food Chem.,* **2017**, *65*, 4537-4552

(42) Peng, X.; Ma, J.; Chen, F.; Wang, M., *Food Funct.,* **2011**, *2,* 289-301

(43) Abbas, G.; Al-Harrasi, A. S.; Hussain, H.; Hussain, J.; Rashid, R.; Choudhary, M. I., *Pharm Biol.,* **2016**, *54,* 198-206

(44) Vrhovac Madunic, I.; Breljak, D.; Sabolic, I., *Period. Biol.,* **2014**, *116*, 61-131

(45) Scheepers, A.; Joost, H. G.; Scharmann, A., *J. Parenter. Enteral. Nutr.,* **2004, *28, 364-371***

(46) Lin, Y.-S.; Tungpradit, R.; Sinchaikul, S.; An, F.-M.; Liu, D.-Z.; Phutrakul, S.; Chen, S.-T., *J. Med. Chem.* **2008**, *51*, 7428-7441

(47) Kim, S.; Go, G. W.; Imm, J. Y., *J. Agric. Food Chem.,* **2017**, *65,* 3819-3826

(48) Shapira, R.; Rudnick, S.; Daniel, B.; Viskind, O.; Aisha, V.; Richman, M.; Ayasolla, K. R.; Perelman, A.; Chill, J. H.; Gruzman, A.; Rahimipour, S., *J. Med. Chem. 2013, 56, 6709-6718*

(49) Burant, C. F.; Bell, G. I., *Biochemistry* **1992**, *31*, 10414-10420

(50) Bai, N.; He, K.; Roller, M.; Zheng, B.; Chen, X.; Shao, Z.; Peng, T.; Zheng, Q., *J. Agric. Food Chem.,* **2008**, *56*, 11668-11674

(51) Brown, G. K., *J. Inherit. Metab. Dis.,* **2000**, *23,* 237-246

(52) Tam, C.; O'Connor, D.; Koren, G., *Obstet. Gynecol. Int.*, **2012**, 485179

(53) Blancquaert, D.; Storozhenko, S.; Loizeau, K.; De Steur, H.; De Brouwer, V.; Viaene, J.; Ravanel, S.; Rabeill, F.; Lambert, W.; Van Der Straeten, D., *Crit. Rev. Plant Sci.,* **2010**, *29*, 14-35

(54) Dawadi, S.; Kordus, S. L.; Baughn, A. D.; Aldrich, C. C., *Org. Lett.* **2017**, *19,* 5220-5223

(55) Vora, A.; Riga, A.; Dollimore, D.; Alexander, K. S., *Thermochim. Acta*, **2002**, *392-393*, 209-220

(56) Rossi, M.; Amaretti, A.; Raimondi, S., *Nutrients*, **2011**, *3*, 118-134

(57) Purcell, W. T.; Ettinger, D. S., *Curr. Oncol. Rep.,* **2003**, *5*, 114-125

(58) Visentin, M.; Zhao, R.; Goldman, I. D., *Hematol. Oncol. Clin. North Am.,* **2012**, *26,* 629-648

(59) Santatiwongchai, J.; Gleeson, D.; Gleeson, M. P., *J. Phys. Chem. B,* **2019**, *123*,

407-418

(60)   Wright, D. L.; Anderson, A. C., *Expert Opin. Ther. Pat.,* **2011**, *21*, 1293-1308

(61)   Zhou, W.; Scocchera, E. W.; Wright, D. L.; Anderson, A. C., *Med. Chem. Comm.,* **2013**, *4*, 908-915

(62)   Hartman, P. G., *J. Chemother.,* **1993**, *5*, 369-376

(63)   Singh, P.; Kaur, M.; Sachdeva, S., *J. Med. Chem.,* **2012**, *55*, 6381-6390

(64)   Lele, A. C.; Raju, A.; Khambete, M. P.; Ray, M. K.; Rajan, M. G. R.; Arkile, M. A.; Jadhav, N. J.; Sarkar, D.; Degani, M. S., *ACS Med. Chem. Lett.* **2015**, *6*, 1140-1144

(65)   Kompis, I. M.; Islam, K.; Then, R. L., *Chem. Rev.,* **2005**, *105*, 593-620

(66)   Wang, Z.; Antoniou, D.; Schwartz, S. D.; Schramm, V. L., *Biochemistry,* **2016**, *55*, 157-166

(67)   Raimondi, M. V.; Randazzo, O.; La Franca, M.; Barone, G.; Vignoni, E.; Rossi, D.; Collina, S., *Molecules.,* **2019**, *24,* 1140

(68)   Reynolds, R. C.; Campbell, S. R.; Fairchild, R. G.; Kisliuk, R. L.; Micca, P. L.; Queener, S. F.; Riordan, J. M.; Sedwick, W. D.; Waud, W. R.; Leung; Dixon, R. W.; Suling, W. J.; Borhani, D. W. *J. Med. Chem.,* **2007**, *50*, 3283-3289

(69)   Scocchera, E.; Reeve, S. M.; Keshipeddy, S.; Lombardo, M. N.; Hajian, B.; Sochia, A. E.; Alverson, J. B.; Priestley, N. D.; Anderson, A. C.; Wright, D. L., *ACS Med. Chem. Lett.,* **2016**, *7*, 692-696

(70)   Koehn, E. M.; Kohen, A., *Arch. Biochem. Biophys.,* **2010**, *493*, 96-102

(71)   Stull, F. W.; Bernard, S. M.; Sapra, A.; Smith, J. L.; Zuiderweg, E. R. P.; Palfey, B. A., *Biochemistry,* **2016**, *55*, 3261-3269

(72)   Costi, M. P.; Gelain, A.; Barlocco, D.; Ghelli, S.; Soragni, F.; Reniero, F.; Rossi, T.; Ruberto, A.; Guillou, C.; Cavazzuti, A.; Casolari, C.; Ferrari, S. *J. Med. Chem.,* **2006**, *49*, 5958-5968

(73)   Stout, T. J.; Tondi, D.; Rinaldi, M.; Barlocco, D.; Pecorari, P.; Santi, D. V.; Kuntz, I. D.; Stroud, R. M.; Shoichet, B. K.; Costi, M. P., *Biochemistry,* **1999**, *38*, 1607-1617

(74)   Papamichael, D., *Oncologist,* **1999**, *4*, 478-487

(75)   Jarmula, A., *Mini. Rev. Med. Chem., 2010, 10, 1211-1222*

(76)   Ruszkowski, M.; Sekula, B.; Ruszkowska, A.; Dauter, Z., *Front Plant Sci.,* **2018***, 9,* 584

(77)   Zhang, Y.; Shang, X.; Lai, S.; Zhang, Y.; Hu, Q.; Chai, X.; Wang, B.; Liu, S.;

Wen, T., *ACS Synth. Biol.,* **2018**, *7*, 635-646

(78)   Ducker, G. S.; Ghergurovich, J. M.; Mainolfi, N.; Suri, V.; Jeong, S. K.; Hsin-Jung Li, S.; Friedman, A.; Manfredi, M. G.; Gitai, Z.; Kim, H.; Rabinowitz, J. D., *Proc. Natl. Acad. Sci. U. S. A.,* **2017**, *114*, 11404-11409

(79)   Tramonti, A.; Nardella, C.; di Salvo, M. L.; Barile, A.; Cutruzzola , F.; Contestabile, R., *Biochemistry,* **2018***, 57,* 6984-6996

(80)   Heil, S. G.; Van der Put, N. M.; Waas, E. T.; den Heijer, M.; Trijbels, F. J.; Blom, H. J., *Mol. Genet. Metab., 2001, 73,* 164-172

(81)   Vivoli, M.; Angelucci, F.; Ilari, A.; Morea, V.; Angelaccio, S.; di Salvo, M. L.; Contestabile, R., *Biochemistry,* **2009**, *48*, 12034-12046

(82)   Nonaka, H.; Nakanishi, Y.; Kuno, S.; Ota, T.; Mochidome, K.; Saito, Y.; Sugihara, F.; Takakusagi, Y.; Aoki, I.; Nagatoishi, S.; Tsumoto, K.; Sando, S., *Nat. Commun.,* **2019**, *10*, 876

(83)   Scaletti, E.; Jemth, A.-S.; Helleday, T.; Stenmark, P. l. *FEBS Lett.,* **2019**, *593*, 1863-1873

(84)   Daidone, F.; Florio, R.; Rinaldo, S.; Contestabile, R.; di Salvo, M. L.; Cutruzzola , F.; Bossa, F.; Paiardini, A., *Eur. J. Med. Chem.,* **2011***, 46,* 1616-1621

(85)   Witschel, M. C.; Rottmann, M.; Schwab, A.; Leartsakulpanich, U.; Chitnumsub, P.; Seet, M.; Tonazzi, S.; Schwertz, G.; Stelzer, F.; Mietzner, T.; McNamara, C.; Thater, F.; Freymond, C.; Jaruwat, A.; Pinthong, C.; Riangrungroj, P.; Oufir, M.; Hamburger, M.; Maoser, P.; Sanz-Alonso, L. M.; Charman, S.; Wittlin, S.; Yuthavong, Y.; Chaiyen, P.; Diederich, F., *J. Med. Chem.,* **2015***, 58,* 3117-3130

(86)   Schwertz, G.; Witschel, M. C.; Rottmann, M.; Bonnert, R.; Leartsakulpanich, U.; Chitnumsub, P.; Jaruwat, A.; Ittarat, W.; Schaofer, A.; Aponte, R. A.; Charman, S. A.; White, K. L.; Kundu, A.; Sadhukhan, S.; Lloyd, M.; Freiberg, G. M.; Srikumaran, M.; Siggel, M.; Zwyssig, A.; Chaiyen, P.; Diederich, F. *J. Med. Chem.,* **2017**, *60*, 4840-4860

(87)   Lee, A. C.; Harris, J. L.; Khanna, K. K.; Hong, J. H., *Int. J. Mol. Sci.,* **2019***, 20,* 2383

(88)   Sachdeva, S., *Int. J. Pept. Res. Ther.,* **2016**, *23*, 49-60

(89)   Lau, J. L.; Dunn, M. K., *Bioorg. Med. Chem.,* **2018***, 26,* 2700-2707

(90)   Gilbert, H. F. *Peptide Bonds, Disulfide Bonds and Properties of Small Peptides*;

Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Ltd, 2001

(91) Huhmann, S.; Stegemann, A.-K.; Folmert, K.; Klemczak, D.; Moschner, J.; Kube, M.; Koksch, B., *Beilstein J. Org. Chem.,* **2017**, *13*, 2869-2882

(92) Bruno, B. J.; Miller, G. D.; Lim, C. S., *Ther. Deliv.,* **2013**, *4,* 1443-1467

(93) De Marco, R.; Spinella, M.; De Lorenzo, A.; Leggio, A.; Liguori, A., *Org. Biomol. Chem.,* **2013**, *11*, 3786-3796

(94) Isidro-Llobet, A.; Kenworthy, M. N.; Mukherjee, S.; Kopach, M. E.; Wegner, K.; Gallou, F.; Smith, A. G.; Roschangar, F., *J. Org. Chem.*, **2019**, *84*, 4615-4628

(95) Amblard, M.; Fehrentz, J.-A.; Martinez, J.; Subra, G., *Mol. Biotechnol.,* **2006**, *33*, 239-254

(96) Merrifield, R. B., *J. Am. Chem. Soc.,* **1963**, *85*, 2149-2154

(97) Palomo, J. M., *RSC Adv.,* **2014**, *4*, 32658-32672

(98) Wu, J.; An, G.; Lin, S.; Xie, J.; Zhou, W.; Sun, H.; Pan, Y.; Li, G., *Chem. Commun.,* **2014**, *50*, 1259-1261

(99) Mitscher, L. A., *J. Nat. Prod.,* **2001**, *64*, 142-142

(100) Khattab, S. N.; Subirós-Funosas, R.; El-faham, A.; Albericio, F. *Eur. J. Org. Chem.*, **2010**, *2010*, 3275-3280

(101) Grant, G. A. Synthetic Peptides. A User's Guide; W.H. Freeman and Company: New York, 1992

(102) Isidro-Llobet, A.; A•lvarez, M.; Albericio, F., *Chem. Rev.,* **2009**, *109*, 2455-2504

(103) Diaz-Eufracio, B. I.; Palomino-Hernandez, O.; Houghten, R. A.; Medina-Franco, J. L. *Mol. Divers.,* **2018**, *22,* 259-267

(104) Gongora-Benitez, M.; Tulla-Puche, J.; Albericio, F., *Chem. Rev.,* **2014**, *114*, 901-926

(105) Barkan, D. T.; Cheng, X.-l.; Celino, H.; Tran, T. T.; Bhandari, A.; Craik, C. S.; Sali, A.; Smythe, M. L., *BMC Bioinf.,* **2016**, *17*, 481

(106) Conibear, A. C.; Daly, N. L.; Craik, D. J., *Biopolymers,* **2012**, *98,* 518-524

(107) Gowd, K. H.; Blais, K. D.; Elmslie, K. S.; Steiner, A. M.; Olivera, B. M.; Bulaj, G., *Biopolymers,* **2012**, *98*, 212-223

(108) Zheng, Y.; Li, Z.; Ren, J.; Liu, W.; Wu, Y.; Zhao, Y.; Wu, C. *Chem. Sci.,* **2017**, *8*, 2547-2552

(109) Han, T. S.; Zhang, M.-M.; Gowd, K. H.; Walewska, A.; Yoshikami, D.; Olivera, B. M.; Bulaj, G., *ACS Med. Chem. Lett.,* **2010**, *1*, 140-144

(110) Kourra, C. M. B. K.; Cramer, N., *Chem. Sci.,* **2016**, *7*, 7007-7012

(111) Fang, G.-M.; Chen, X.-X.; Yang, Q.-Q.; Zhu, L.-J.; Li, N.-N.; Yu, H.-Z.; Meng, X.-M. *Chin. Chem. Lett.,* **2018**, *29*, 1033-1042

(112) Aoki, K.; Maeda, M.; Nakae, T.; Okada, Y.; Ohya, K.; Chiba, K. *Tetrahedron* **2014**, *70*, 7774-7779

(113) Viero, C.; Shibuya, I.; Kitamura, N.; Verkhratsky, A.; Fujihara, H.; Katoh, A.; Ueta, Y.; Zingg, H. H.; Chvatal, A.; Sykova, E.; Dayanithi, G. *CNS Neurosci. Ther.,* **2010**, *16*, e138-e156

(114) Antoni, F. A.; Chadio, S. E., *Biochem.,* **1989**, *257*, 611-614

(115) Elabd, S.; Sabry, I. *Front. Endocrinol.,* **2015**, *6*, 121

(116) Gimpl, G.; Fahrenholz, F. *Physiol. Rev.,* **2001***, 81, 629-883*

(117) Benavente, M. A.; Bianchi, C. P.; Imperiale, F.; Aba, M. A., *Front Vet. Sci.,* **2016**, *3*, 119-119

(118) Lin, H.-J.; Chen, Y.-S.; Chen, Y.-J.; Lin, Y.; Mersmann, H.; Ding, S.-T. *J. Biomed. Sci. Eng.,* **2017**, *10*, 37-50

(119) Klement, J.; Ott, V.; Rapp, K.; Brede, S.; Piccinini, F.; Cobelli, C.; Lehnert, H.; Hallschmid, M. *Diabetes* **2017**, *66*, 264-271

(120) Khoo, K. K.; Norton, R. S. *Amino Acids, Pept. Proteins Org. Chem.,* 2011, 395-417

(121) Wright, Z. V. F.; McCarthy, S.; Dickman, R.; Reyes, F. E.; Sanchez-Martinez, S.; Cryar, A.; Kilford, I.; Hall, A.; Takle, A. K.; Topf, M.; Gonen, T.; Thalassinos, K.; Tabor, A. B., *J. Am. Chem. Soc.,* **2017**, *139*, 13063-13075

(122) Zhang, C.; Niu, W.; Wang, Z.; Wang, X.; Xia, G., *PLoS One,* **2012***, 7,* e42406

(123) Copland, J. A.; Zlatnik, M. G.; Ives, K. L.; Soloff, M. S. *Biol. Reprod.,* **2002**, *66*, 1230-1236

(124) Greenfield, N. J., *Nat. Protoc.,* **2006**, *1*, 2876-2890

(125) Greenfield, N. J., *TrAC Trends Anal. Chem.,* **1999**, *18*, 236-244

(126) Purdie, N., *J. Am. Chem. Soc.,* **1996**, *118*, 12871-12871

(127) Miles, A. J.; Wallace, B. A., *Chem. Soc. Rev.,* **2016**, *45*, 4859-4872

(128) Whitmore, L.; Wallace, B. A., *Biopolymers,* **2008**, *89*, 392-400

(129) Reed, J.; Reed, T. A., *Anal. Biochem.,* **1997**, *254*, 36-40

(130) Ranjbar, B.; Gill, P., *Chem. Biol. Drug Des.,* **2009**, *74*, 101-120

(131) He, S.-B.; Chen, R.-T.; Wu, Y.-Y.; Wu, G.-W.; Peng, H.-P.; Liu, A.-L.; Deng, H.-H.; Xia, X.-H.; Chen, W., *Microchim. Acta* **2019**, *186*, 778

(132) Siques, P.; Brito, J.; Flores, K.; Ordenes, S.; Arriaza, K.; Pena, E.; Lean-Velarde, F.; Lapez de Pablo, A. n. L.; Gonzalez, M. C.; Arribas, S. *Front. Physiol.,* **2018**, *9*

(133) Vijayakumar, M. V.; Ajay, A. K.; Bhat, M. K., *J. Biosci.,* **2010**, *35*, 525-531

(134) Tan, S. Y.; Mei Wong, J. L.; Sim, Y. J.; Wong, S. S.; Mohamed Elhassan, S. A.; Tan, S. H.; Ling Lim, G. P.; Rong Tay, N. W.; Annan, N. C.; Bhattamisra, S. K.; Candasamy, M., *Diabetes Metab. Syndr.,* **2019***, 13,* 364-372

(135) Galvanin, F.; Barolo, M.; Macchietto, S.; Bezzo, F., *Ind. Eng. Chem. Res.,* **2009**, *48*, 1989-2002

(136) Lee, P. G.; Halter, J. B., *Diabetes Care* **2017**, *40*, 444-452

(137) Lapolla, A.; Fedele, D.; Reitano, R.; Arica, N. C.; Seraglia, R.; Traldi, P.; Marotta, E.; Tonani, R., *J. Am. Soc. Mass Spectrom.,* **2004***, 15,* 496-509

(138) Lopez-Clavijo, A. F.; Duque-Daza, C. A.; Soulby, A.; Canelon, I. R.; Barrow, M.; O'Connor, P. B., *J. Am. Soc. Mass Spectrom.,* **2014***, 25,* 2125-2133

(139) Nazhad, N., Manchester Metropolitan University 2015

(140) Chilukuri, H.; Kulkarni, M. J.; Fernandes, M., *Med. Chem. Comm.,* **2018**, *9*, 614-624

(141) Rhee, S. Y.; Kim, Y. S., *Diabetes Metab. J.,* **2018**, *42*, 188-195

(142) Puddu, A.; Sanguineti, R.; Maggi, D.; Nicola, M.; Traverso, C. E.; Cordera, R.; Viviani, G. L., *J. Diabetes Res.*, *2019*, 6198495

(143) Yang, P.; Feng, J.; Peng, Q.; Liu, X.; Fan, Z., *Oxid. Med. Cell. Longevity*, **2019**, 9570616

(144) Zhu, Y.; Snooks, H.; Sang, S., *J. Agric. Food Chem., 2018, 66, 1325-1329*

(145) Vlassara, H.; Brownlee, M.; Cerami, A., *Diabetes,* **1988**, *37*, 456-461

(146) Chinchansure, A. A.; Korwar, A. M.; Kulkarni, M. J.; Joshi, S. P. *RSC Adv.,* **2015**, *5*, 31113-31138

(147) Beseni, B. K.; Bagla, V. P.; Njanje, I.; Matsebatlela, T. M.; Mampuru, L.; Mokgotho, M. P., *Evid-Based Compl. Alt.*, *2017*, 6453567

(148) Kolekar, Y. M.; Vannuruswamy, G.; Bansode, S. B.; B, S.; Thulasiram, H. V.; Kulkarni, M. J., *RSC Adv.,* **2015**, *5*, 25051-25058

(149) Al-Amoudi, M.; Ahmed, M.; Al Majthoub, M. d.; Adam, A.; Alshanbari, N.; Refat, M., *Res. Chem. Intermed.,* **2015**, *41*, 3089-3108

(150) Yin, S.; Yang, Y.; Li, Y.; Sun, C., *Anal. Methods.,* **2018**, *10*, 9-21

(151) Chan, Y.-M.; Bailey, R.; O'Connor, D. L., *Adv. Nutr.,* **2013**, *4*, 123-125

(152) Greenberg, J. A.; Bell, S. J.; Guan, Y.; Yu, Y.-H., *Reviews in Obstetrics & Gynecology*, **2011**, *4*, 52-59

(153) Alam, C.; Aufreiter, S.; Georgiou, C. J.; Hoque, M. T.; Finnell, R. H.; O'Connor, D. L.; Goldman, I. D.; Bendayan, R., *Proc. Natl. Acad. Sci. U. S. A.*, **2019**, *116*, 17531-17540

(154) Shuvalov, O.; Petukhov, A.; Daks, A.; Fedorova, O.; Vasileva, E.; Barlev, N. A., *Oncotarget,* **2017***, 8,* 23955-23977

(155) Anderson, A. C.; Wright, D. L., *Expert Opin. Ther Pat.,* **2014,** *24,* 687-97

(156) Walling, J., *Invest. New Drugs* **2006**, *24*, 37-77

(157) Fernandez-Villa, D.; Aguilar, M. R.; Rojo, L., *Int. J. Mol. Sci.,* **2019***, 20,* 4996

(158) Scagliotti, G. V.; Selvaggi, G., *Expert Opin. Ther. Pat.,* **2006**, *16*, 189-200

(159) Angelastro, A.; Dawson, W. M.; Luk, L. Y. P.; Loveridge, E. J.; Allemann, R. K., *J. Am. Chem. Soc.,* **2017**, *139*, 13047-13054

(160) Chen, C.; Ke, J.; Zhou, X. E.; Yi, W.; Brunzelle, J. S.; Li, J.; Yong, E.-L.; Xu, H. E.; Melcher, K., *Nature,* **2013**, *500*, 486-489

(161) Ferreira, L. G.; Dos Santos, R. N.; Oliva, G.; Andricopulo, A. D., *Molecules* **2015**, *20*, 13384-13421

(162) Azam, S. S.; Abbasi, S. W., *Theor. Biol. Med. Modell.,* **2013**, *10*, 63

(163) Renwick, S. B.; Snell, K.; Baumann, U., *Structure,* **1998***, 6,* 1105-1116

ABSTRACT

Name of the Student: Harsha Chilukuri   **Registration No. :** 10CC17J26025
Faculty of Study: Chemical Sciences   **Year of Submission:** 2021
AcSIR academic centre/CSIR Lab: CSIR-NCL   **Name of the Supervisor:** Dr. Moneesha Fernandes

Title of the thesis: Alkyl proline and oxytocin derivatives towards development of anti-diabetics and design of potential antifolates derived from guanine

 ----

Designing new molecules for drug development has gathered interests of many researchers. Many diseases, although identified many decades ago, still require new strategies as therapy. For this reason many small molecules, peptides, oligonucleotides, proteins etc have been extensively developed. Antidiabetics and antimetabolites such as antifolates are classes of drugs that have been used in the clinics from a long time but many drugs still fall short of being ideal therapy. This thesis describes the appropriate design, synthesis and evaluation of potential small molecules and peptides as antidiabetics and antimetabolites. One section of the thesis describes glucose uptake activity of several linear and cyclic oxytocin analogues on Chinese Hamster Ovary cells. The disulfide bond in native oxytocin was replaced with an amide bond to overcome serum degradation. Glucose uptake activity was assessed by glucose oxidase-peroxidase assay. All the analogues of the study showed glucose uptake activity. Another section of the thesis describes the synthesis of a new series of non-natural *N*-(3-aminoalkyl)proline derivatives exhibiting antiglycation activity. The compounds were unambiguously characterized by NMR, mass and IR spectroscopy. The *in vitro* antiglycation activity was studied by circular dichroism and fluorescence spectrometry. The mechanism of action was also studied and found to take place by Schiff base formation between the sugar and the protein further inhibiting Amadori product formation. The third section of the thesis describes the design and *in silico* evaluation of antifolates derived from guanine nucleobase. These new analogues of guanine are structurally similar to many antifolates that are currently in use. All the designed ligands were docked against three isoforms of Serine hydroxymethyltransferase (SHMT) which is a ubiquitous enzyme found in all prokaryotes and is the only enzyme yet to be exploited as a target for cancer chemotherapy. Docking study revealed that all the designed ligands showed good binding energy with all isoforms of SHMT. Additionally, the ligands when docked against folate receptor, FRα (4LRH) showed similar and in some cases higher binding energy compared to folic acid. This result was in line with the hypothesis of this study since the designed ligands were envisaged to have a competitive binding   with   folic   acid

1. List of Publications in SCI Journal(s) (published & accepted) emanating from the thesis work, with complete bibliographic details.

Harsha Chilukuri, Yogesh M. Kolekar, Govind S. Bhosle, Rashmi K. Godbole, Rubina S. Kazi, Mahesh J. Kulkarni and Moneesha Fernandes, *N*-(3-Aminoalkyl)proline derivatives with potent antigycation activity, *RSC Adv*., **2015**, *5*, 77332-77340

Harsha Chilukuri, Mahesh J. Kulkarni and Moneesha Fernandes, Revisiting amino acids and peptides as antiglycation Agents *Med.Chem.Commun*., **2018**, *9*, 614-624

2. List of Papers with abstract presented (oral/poster) at national/international conferences/seminars with complete details:

NA

3. List of other publications in SCI Journal(s) with complete bibliographic details:

p-Nitrophenylcarbonates: A New Class of Compounds for Chemodosimetric Colorimetric Fluoride Anion Sensing Detectable by the Naked Eye, Amit Kumar Yadav, **Harsha Chilukuri**, Linthoinganbi Raj Kumari, Muthukumarasamy Karthikeyan and Moneesha Fernandes, *Chemistry Select* **2019**, *4*, 1830 –1833 *(Chemistry Select)*

4. A copy of all SCI publication(s) emanating from the thesis:

N-(3-Aminoalkyl)proline derivatives with potent antigycation activity†

Harsha Chilukuri,[ab] Yogesh M. Kolekar,[a] Govind S. Bhosle,[b] Rashmi K. Godbole,[a] Rubina S. Kazi,[a] Mahesh J. Kulkarni*[a] and Moneesha Fernandes*[b]

The importance of amino acids in the therapy of conditions such as renal failure, neurological disorders and congenital defects has been documented. Some amino acids such as lysine and glycine have also been reported to have antiglycating activity. Herein we report the synthesis of a new series of N-(3-aminoalkyl)proline derivatives which are non-natural in nature. The compounds were unambiguously characterized by NMR, mass and IR spectroscopy. Their *in vitro* antiglycation activity was studied by circular dichroism and fluorescence spectrometry. The mechanism of action was also studied and found to take place by inhibition of Amadori product formation. The inhibition of AGE formation was further confirmed by western blot and LC-MS/MS analyses and the IC_{50} values of the potent compounds were determined. Compounds containing hydroxyl substituents at C4 were found to have superior antiglycation properties than those containing azide substituents at the same position. The compounds were additionally found to possess good anti-oxidant properties, which could lead to further reduction in AGE formation. Moreover, the title compounds were found to have low cytotoxicity in mammalian cells, another important attribute. Thus, the title compounds represent a novel promising class of antiglycating agents.

Introduction

Advanced Glycation End products (AGEs) are formed due to non-enzymatic reactions between proteins and reducing sugars. They are involved in the pathogenesis of diabetes and its complications. Thus, reducing AGE levels[1] and inhibition of protein glycation is crucial in the prevention of the above complications. Some molecules such as aminoguanidine that reduce AGE levels, but have toxic side-effects have not been approved by the FDA, while other FDA-approved drugs have been shown to display anti-glycation activity[1–5] in addition to their originally intended activities. These offer the possibility of repositioning these drugs for the treatment of diabetes and its complications. Reports from this laboratory have also been made for the use of N-(aminoalkyl)proline- and 4-hydroxy-N-(aminoalkyl)proline-derived compounds towards the synthesis of peptide nucleic acids[6] and cell-penetrating oligomers,[7] where these compounds were used as analogues of amino acids such as lysine and arginine.[7b] Since amino acids such as glycine and lysine have been reported to exert antiglycating activity,[8] we surmised that the use of non-natural analogues such as the title compounds, would have advantages in terms of stability *in vivo*.

In this study, we report the synthesis and anti-glycation activity of selected N-(3-aminoalkyl)proline derivatives and their mode of action by various physicochemical assays such as circular dichroism (CD), fluorescence spectroscopy, MALDI-TOF and LC-MS/MS assays. We also show that these compounds are capable of exerting good anti-oxidant properties and display low cytotoxicity, thus enhancing their value as potential anti-glycation agents.

Results and discussion

Synthesis of title compounds

The title compounds were synthesized using a simple strategy outlined in Scheme 1. Accordingly, 4(S)-hydroxy-2(R)-proline methylester was N-alkylated by treating it with 3-((*tert*-butoxy-carbonyl)amino)propyl methanesulfonate in CH_2Cl_2 in the presence of triethylamine to yield compound **1** which was subjected to acidolytic removal of the Boc protecting group to afford compound **2** in 62% yield. Compound **3** was obtained from **1** upon saponification with lithium hydroxide and subsequent purification by column chromatography on neutral alumina.

[a] Proteomics Facility, Division of Biochemical Sciences, CSIR-National Chemical Laboratory, Pune-411 008, India. E-mail: mj.kulkarni@ncl.res.in; Tel: +91 2025902541

[b] Organic Chemistry Division, CSIR-National Chemical Laboratory, Pune 411008, India. E-mail: m.dcosta@ncl.res.in; Tel: +91 2025902084
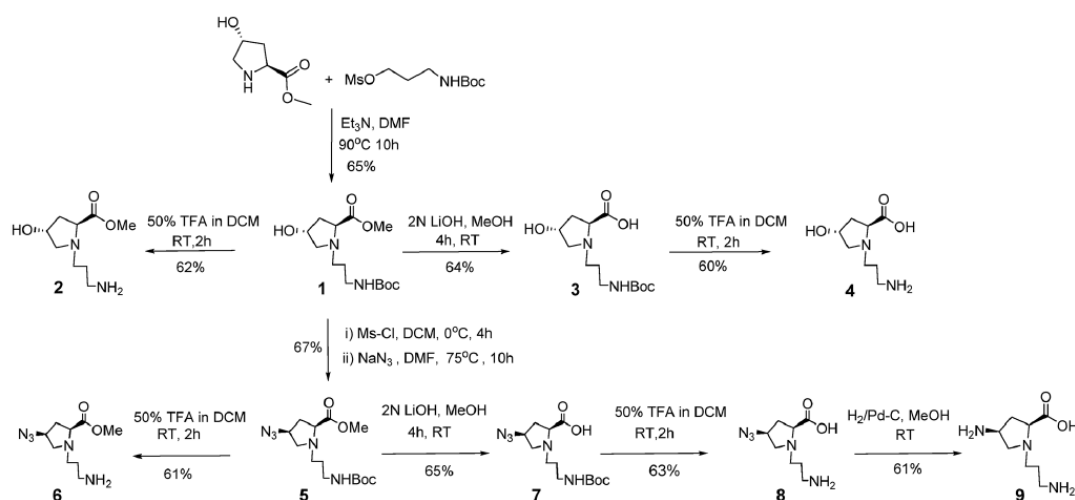
Scheme 1 Synthesis of the title compounds.

The removal of the Boc protecting group in **3** resulted in compound **4**, which was obtained in 60% yield. Mesylation of the 4-OH group in compound **1** yielded the 4-*O*-mesyl derivative, which was further converted to its azide counterpart **5** in 67% yield over two steps after purification by column chromatography. Compound **6** was obtained by acidolytic removal of the Boc protecting group in **5**, and purified by column chromatography. Compound **5**, upon saponification with lithium hydroxide yielded **7** in 65% yield after column purification. Compound **7** was further subjected to removal of the Boc protecting group by treating it with TFA in CH_2Cl_2 to yield compound **8** in 63%. Reduction of the azide function in **8** by hydrogenation in the presence of Pd–C further gave compound **9** in 61% yield. All the compounds were unambiguously characterized by appropriate spectroscopic techniques.

Glycation inhibition studies

Circular dichroism (CD). CD is a powerful tool for investigating the structure and conformational changes of proteins, such as those occurring upon glycation,[9] where an increase in the beta sheet conformation is observed. A decrease in the beta sheet conformation is therefore, indicative of antiglycation ability. Analysis of the CD spectra of BSA upon glycation in the presence of the title compounds (Fig. 1) revealed a decrease in the beta sheet percentage in comparison to glycated BSA, when no compound was present. In particular, the beta sheet conformation in the presence of compounds **2**, **4**, **6**, **8** and **9** was 3%, 1.7%, 2.5%, 2% and 5.2% respectively, while that in glycated BSA was 18.7% and 5.7% in the presence of known antiglycating agent, aminoguanidine (Amg). Thus, the title compounds are able to protect the conformation of BSA and inhibit beta sheet formation significantly, even in comparison to aminoguanidine.

***In vitro* glycation inhibition by fluorescence spectroscopy.** The degree of glycation can be measured by measuring the

fluorescence intensity at 440 nm, using an excitation wavelength of 370 nm,[10] since most AGEs have a characteristic fluorescence with an excitation maximum approximately at 370 nm, and emission around 440 nm. A decrease in the fluorescence emission intensity is thus, indicative of inhibition of AGE formation. BSA was incubated with glucose in the presence and absence of title compounds and the fluorescence emission of the compounds was monitored after excitation at 370 nm (Fig. 2). The formation of AGEs was monitored after 14 days by measuring the fluorescence emission intensity at 440 nm. The fluorescence-based assay for AGEs was also used to determine the IC_{50} values of representative title compounds **2** and **6**. Accordingly, these were found to be 19.2 mM and 29.5 mM respectively. In comparison, the IC_{50} value of aminoguanidine, considering the fluorescence excitation/emission as 370 nm/440 nm is reported to be 10 mM.[11] However, this fluorescence-based method may be less specific[12] because some fluorescent AGEs differ in their excitation–emission wavelengths.

Western blot assay. Western blot analyses of the glycation products of BSA in the presence and absence of the title compounds were carried out using both anti-AGE as well as anti-CML (anti-carboxymethyl lysine) antibodies, since CML is known to be the most abundant non-fluorescent AGE. In both the cases, compounds **2** and **4** were found to exhibit potent antiglycation activity, which was even superior to aminoguanidine (Fig. 3). Ponceau staining of the gels was also carried out (ESI, Fig. S16†) to illustrate the protein (BSA) content in the samples.

From the above studies, compounds **2** and **4** emerged as the most promising title compounds with potent anti-glycation activity, which was found to be superior to known antiglycating agent, aminoguanidine.

Probable mechanism of glycation inhibition. Circular dichroism, AGE fluorescence and western blotting suggested that at least some of the title compounds are capable of inhibiting AGE formation in proteins, using BSA as a
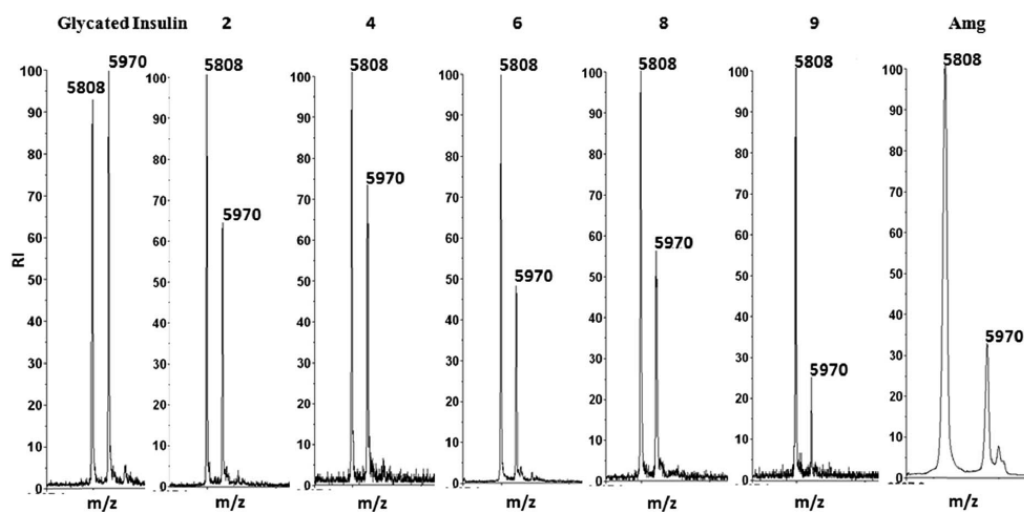
Fig. 4 MALDI-TOF assay for glycation inhibition for glycated insulin, insulin in presence of compounds 2, 4, 6, 8, 9 and aminoguanidine.[5]

incident vascular complications.[14] The fructosamine assay is a simple colorimetric test that measures glycated serum protein concentrations. Colour change is based on the reduction of nitroblue tetrazolium (NBT) to monoformazan (MF) by Amadori rearrangement products.[15] The fructosamine levels in presence of the title compounds were found to be significantly lower than that in glycated BSA (Fig. 6). Specifically, the fructosamine levels were 539, 551, 476, 592 and 526 μmol L^{-1} in the presence of compounds 2, 4, 6, 8 and 9 respectively, which were even lower than that in the presence of aminoguanidine (735 μmol L^{-1}), while the fructosamine level in glycated BSA was 976 μmol L^{-1}. This data is thus, in accordance with that obtained from the MALDI-TOF assay, which suggested that the title compounds inhibited Amadori product formation.

Adduct formation of title compounds with glucose. The title compounds were incubated in the presence of glucose and the

reaction mixture was subjected to LC-MS analysis. All the compounds were found to form adducts with glucose. These are probably Schiff base type of adducts formed by the reaction of the free amino group in the title compounds with the aldehyde group of the sugar (Fig. 7, for a representative title compound, 2. The adduct formation with other title compounds is depicted in the ESI, page S17 and S18†). The formation of adduct with glucose could thus, be one of the ways by which the title compounds inhibit glycation.

Cytotoxicity assay. The cytotoxicity of the title compounds was evaluated in L6 rat muscle cells by measuring the cell viability using the standard MTT assay. It is a colorimetric assay that is based on the reduction of the tetrazolium dye, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide to a purple formazan derivative, that is achieved by enzymes present in viable cells. A decreased colour intensity is therefore indicative of low cell viability as a result of increased cytotoxicity. As
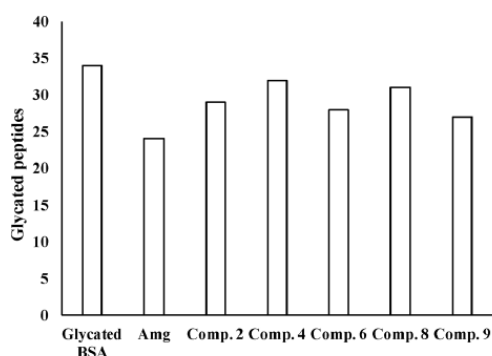
Fig. 5 LC-MS/MS analysis depicting the number of AGE-modified peptides in glycated BSA and glycated BSA after treatment with aminoguanidine or title compounds.
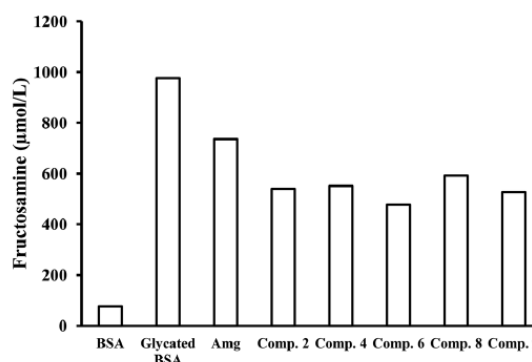
Fig. 6 Fructosamine levels of BSA, glycated BSA and glycated BSA in presence of aminoguanidine or title compounds.
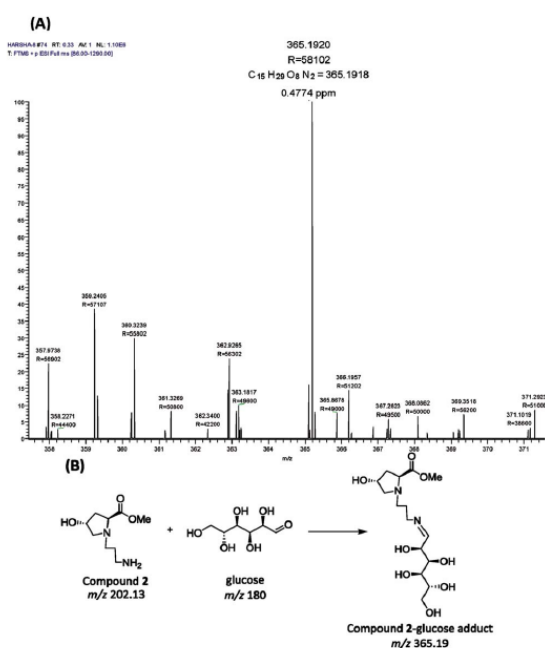
Fig. 7 (A) LC-MS spectrum of compound 2-glucose adduct and (B) probable mechanism of compound 2-glucose adduct formation.

seen in Fig. 8, all the title compounds were non-toxic at almost all concentrations tested, except at 40 mM concentration, where a slight decrease in cell viability was observed, this being more pronounced in compounds **6** and **8**. However, this concentration is double that used for the glycation inhibition studies described herein.

Anti-oxidant properties of title compounds. AGEs, along with advanced lipoxidation products, have also been implicated in the generation of free radicals and reactive oxygen species
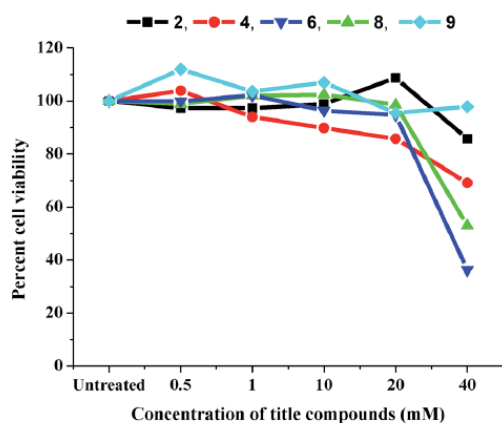
Fig. 8 MTT assay depicting viability of L6 rat muscle cells in presence of title compounds.
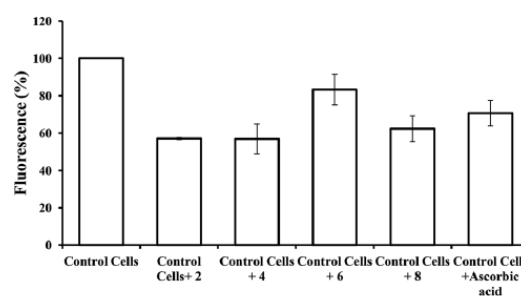
Fig. 9 Anti-oxidant properties of title compounds in comparison to ascorbic acid.

(ROS), that are known to be involved in a variety of cellular processes ranging from apoptosis and necrosis to carcinogenesis.[16] It therefore follows that any candidate, that is capable of inhibiting AGE formation and additionally possess anti-oxidant properties, would have increased therapeutic value. Keeping this in mind, the title compounds were evaluated for their ability to reduce the concentration of intracellular ROS, using ascorbic acid as a control anti-oxidant, in a fluorescence-based assay. The anti-oxidant activity of a compound is evidenced by a decrease in the fluorescence intensity. Fig. 9 shows a comparative study of the title compounds for anti-oxidant activity, wherein, it was found that compounds **2** and **4** exhibited higher anti-oxidant activity than other compounds, including ascorbic acid. More specifically, the percent fluorescence values were 57, 56, 83, 62 and 70 for compounds **2**, **4**, **6**, **8** and ascorbic acid respectively. Compound **9** did not show anti-oxidant activity as observed for other compounds.

Experimental

Synthetic procedures

Methyl(2S,4R)-1-(3-((tert-butoxycarbonyl)amino)propyl)-4-hydroxypyrrolidine-2-carboxylate (1). Methyl(2S,4R)-4-hydroxy-pyrrolidine-2-carboxylate (1.00 g, 6.8 mmol) was dissolved in dry DMF; to it dry Et_3N (3.6 mL, 20.0 mmol) was added and stirred for 5 min at RT. Then to the above solution, 3-((tert-butoxycarbonyl)amino)propyl methanesulfonate (2.09 g, 8.0 mmol) was added drop-wise with continuous stirring. The reaction mixture was heated at 90 °C for 10 h. The reaction was monitored by TLC and after completion of the reaction, DMF was removed under vacuum and the compound was extracted from water with ethyl acetate. The organic layer was concentrated to get the crude compound, which was purified by column chromatography to get yellow coloured gum (1.3 g, 65%). $[\alpha]_D = -48.31$ ($c = 0.135$ in MeOH); ^{1}H NMR (200 MHz, CDCl$_3$) δ: 1.44 (s, 9H), 1.59–1.69 (m, 2H), 2.08–2.17 (m, 2H), 2.38–2.56 (m, 2H), 2.69–2.83 (m, 1H), 3.08 (br s, 1H), 3.19 (br m, 2H), 3.36–3.34 (dd, $J = 5.4$, 10.1 Hz, 1H), 3.50–3.58 (t, $J = 7.6$ Hz, 1H), 3.72 (s, 3H), 4.43–4.46 (m, 1H), 5.38 (br s, 1H) ppm. ^{13}C NMR (50 MHz, CDCl$_3$) δ: 28.0, 28.4, 39.3, 51.8, 60.9, 64.4, 70.0, 78.9, 156.3, 174.4 ppm; ^{13}C DEPT (50 MHz, CDCl$_3$) δ: 28.0 (CH$_2$), 28.4, 38.5 (CH$_2$), 39.3 (CH$_2$), 51.8 (CH$_2$), 51.8, 60.9 (CH$_2$), 64.4, 70.0 ppm. HRMS:

calculated mass for $C_{14}H_{26}N_2O_5$: 302.1842, observed mass (M + H): 303.1915, (M + Na): 325.1713.

Methyl(2S,4R)-1-(3-aminopropyl)-4-hydroxypyrrolidine-2-carboxylate (2). To compound **1** (0.10 g, 0.33 mmol), a 50% solution of TFA in DCM (5 mL) was added with vigorous stirring at RT. After completion of the reaction, solvents were removed under vacuum and crude compound was purified by column chromatography. The resultant TFA salt of the compound was obtained in 62% yield (0.065 g). $[\alpha]_D = -24.47$ (c = 0.11 in MeOH). IR (neat) ν: 3352, 2946, 2833, 1749, 1683 cm^{-1}. ^{1}H NMR (200 MHz, D$_2$O) δ: 2.09–2.21 (m, 2H), 2.33–2.41 (m, 1H), 2.49–2.60 (m, 1H), 3.05–3.13 (m, 2H), 3.28–3.41 (m, 2H), 3.54–3.60 (m, 1H), 3.86 (s, 3H), 3.91–3.99 (m, 1H), 4.66 (br m, 2H) ppm. ^{13}C NMR (50 MHz, D$_2$O) δ: 24.6, 38.0, 38.8, 52.7, 53.2, 60.2, 64.8, 69.2, 175.5 ppm. ^{13}C DEPT (50 MHz, D$_2$O) δ: 34.6 (CH$_2$), 38.0 (CH$_2$), 38.8 (CH$_2$), 52.7, 53.2 (CH$_2$), 60.2 (CH$_2$) 64.8, 69.2 ppm. HRMS: calculated mass for $C_9H_{18}N_2O_3$: 202.1317, observed mass (M + H): 203.1392.

(2S,4R)-1-(3-((tert-Butoxycarbonyl)amino)propyl)-4-hydroxypyrrolidine-2-carboxylic acid (3). Compound **1** (0.2 g, 0.66 mmol) was dissolved in methanol and to it, 2 N LiOH (3 mL) was added with vigorous stirring. The reaction was monitored by TLC and after completion of reaction, the reaction mixture was neutralised using Dowex H$^+$ resin; the resin was subsequently filtered off. The filtrate was concentrated under vacuum and the compound was purified by column chromatography (0.12 g, 64%). $[\alpha]_D = -35.06$, (c = 0.09 in MeOH). ^{1}H NMR (200 MHz, D$_2$O) δ: 1.39 (s, 9H), 1.82–1.86 (m, 2H), 2.11–2.51 (m, 2H), 3.16 (br m, 2H), 3.23 (br s, 1H), 3.31 (br. m, 1H), 3.85–3.93 (dd, J = 4.2, 13.0 Hz, 1H), 4.14–4.23 (m, 1H), 4.59 (br s, 1H) ppm. ^{13}C NMR (50 MHz, D$_2$O) δ: 28.3, 30.1, 40.5, 58.0, 64.2, 71.1, 71.8, 83.6, 160.7, 175.6 ppm. ^{13}C DEPT (50 MHz, D$_2$O) δ: 28.3 (CH$_2$), 30.1, 39.3 (CH$_2$), 40.5 (CH$_2$), 58.0 (CH$_2$), 64.2 (CH$_2$) 71.1, 71.8 ppm. HRMS: calculated mass for $C_{13}H_{24}N_2O_5$: 288.1685, observed mass (M + H): 289.1757.

(2S,4R)-1-(3-Aminopropyl)-4-hydroxypyrrolidine-2-carboxylic acid (4). Compound **3** (0.30 g, 1.04 mmol) was dissolved in a 50% solution of TFA in DCM (5 mL) with vigorous stirring at RT. The reaction was monitored by TLC and after the completion of reaction, solvents were removed under vacuum. Yield 0.19 g (60%). $[\alpha]_D = -12.18$ (c = 0.11 in MeOH). IR (neat) ν: 3375, 2948, 2834, 1684 cm^{-1}. ^{1}H NMR (400 MHz, D$_2$O) δ: 2.12–2.13 (m, 2H), 2.25–2.32 (m, 1H), 2.52–2.56 (m, 1H), 3.06–3.10 (m, 2H), 3.27–3.38 (m, 2H), 3.52–3.54 (m, 1H), 3.94–3.97 (m, 1H), 4.51–4.55 (dd, J = 7.1, 11.3 Hz, 1H), 4.64 (br s, 1H) ppm. ^{13}C NMR (50 MHz, D$_2$O) δ: 23.4, 36.3, 38.0, 48.8, 54.5, 61.8, 69.3, 172.9 ppm. ^{13}C DEPT (50 MHz, D$_2$O) δ: 23.4 (CH$_2$), 36.3 (CH$_2$), 38.0 (CH$_2$), 54.5 (CH$_2$), 61.8 (CH$_2$), 68.6, 69.3 ppm. HRMS: calculated mass for $C_8H_{16}N_2O_3$: 188.1161, observed mass (M + H): 189.1235.

Methyl(2S,4S)-4-azido-1-(3-((tert-butoxycarbonyl)amino) propyl)pyrrolidine-2-carboxylate (5). To a solution of **1** (0.7 g, 2.30 mmol) in dry DCM, dry Et$_3$N (0.9 mL, 6.90 mmol) was added and stirred for 5 min at RT. Ms-Cl (0.3 mL, 0.003 mol) was added dropwise at 0 °C, with continuous stirring. After completion of the reaction, DCM was evaporated under vacuum and the crude mesylated compound was used for further reaction without further purification. The mesylated compound was

dissolved in dry DMF, to it NaN$_3$ was added with continuous stirring (1.5 g, 23.1 mmol). The reaction was heated for 10 h at 75 °C. After completion of the reaction, DMF was removed under vacuum and the compound was purified by column chromatography (0.58 g, 67%). $[\alpha]_D = -0.0686$ (c = 0.1 in CHCl$_3$). ^{1}H NMR (200 MHz, CDCl$_3$) δ: 1.44 (s, 9H), 1.65–1.68 (m, 2H), 2.02–2.15 (br m, 1H), 2.34–2.60 (m, 3H), 2.78–2.88 (m, 1H), 3.15–3.25 (br m, 4H), 3.77 (s, 3H), 3.96 (br s, 1H), 5.47 (br s, 1H) ppm. ^{13}C NMR (50 MHz, CDCl$_3$) δ: 27.8, 28.4, 35.4, 38.4, 51.7, 52.1, 58.0, 58.7, 64.8, 77.2, 78.6, 156.2, 173.3 ppm. ^{13}C DEPT (50 MHz, CDCl$_3$) δ: 27.7 (CH$_2$), 28.4, 35.4 (CH$_2$), 38.4 (CH$_2$), 51.7 (CH$_2$), 52.1, 58.0 (CH$_2$), 58.7, 64.8 ppm. HRMS: calculated mass for $C_{14}H_{25}N_5O_4$: 327.1907, observed mass (M + H): 328.1979, (M + Na): 350.1797.

Methyl(2S,4S)-1-(3-aminopropyl)-4-azidopyrrolidine-2-carboxylate (6). To compound **5** (0.3 g, 0.91 mmol), a 50% solution of TFA in DCM (5 mL) was added with vigorous stirring at RT. The reaction was monitored by TLC and after the completion of reaction, solvents were removed under vacuum and the product was purified by column chromatography (0.19 g, 61%). $[\alpha]_D = -8.83$ (c = 0.105 in MeOH). IR (neat) ν: 3376, 2188, 1745, 1678 cm^{-1}. ^{1}H NMR (200 MHz, D$_2$O) δ: 0.90–1.01 (m, 1H), 1.82–1.94 (m, 2H), 2.23–2.30 (m, 1H), 2.58–2.60 (m, 1H), 2.80–2.87 (m, 2H), 3.17–3.27 (m, 2H), 3.64 (s, 3H), 4.40–4.45 (m, 2H) ppm. ^{13}C NMR (50 MHz, D$_2$O) δ: 23.1, 33.7, 36.2, 52.9, 54.2, 58.6, 59.9, 65.7, 169.2 ppm. ^{13}C DEPT (50 MHz, D$_2$O) δ: 23.0 (CH$_2$), 33.6 (CH$_2$), 36.1 (CH$_2$), 52.8 (CH$_2$), 54.1, 58.5, 59.8 (CH$_2$), 65.7 ppm. HRMS (ESI): calculated mass for $C_9H_{17}N_5O_2$: 227.1382, observed mass (M + H): 228.1457.

(2S,4S)-4-Azido-1-(3-((tert-butoxycarbonyl)amino)propyl) pyrrolidine-2-carboxylic acid (7). Compound **5** (0.2 g, 0.61 mmol) was dissolved in methanol and to it 2 N LiOH (3 mL) was added with vigorous stirring. After completion of the reaction, the reaction mixture was neutralised using Dowex H$^+$ resin; the resin was subsequently filtered off. The filtrate was concentrated under vacuum and the compound was purified by column chromatography (0.125 g, 65%). $[\alpha]_D = -4.54$ (c = 0.11 in MeOH). ^{1}H NMR (200 MHz, D$_2$O) δ: 1.42 (s, 9H), 1.87–1.89 (br m, 2H), 2.34–2.41 (m, 1H), 2.68–2.79 (m, 1H), 3.06–3.39 (br m, 6H), 3.76–3.82 (m, 1H), 4.08–4.12 (m, 1H), 4.57 (br s, 1H) ppm. ^{13}C NMR (50 MHz, D$_2$O) δ: 25.7, 27.6, 34.3, 53.5, 59.0, 59.6, 67.2, 81.1, 158.2, 173.0 ppm. ^{13}C DEPT (50 MHz, D$_2$O) δ: 25.7 (CH$_2$), 27.6, 34.3 (CH$_2$), 36.8 (CH$_2$), 53.5 (CH$_2$), 59.0, 59.6 (CH$_2$), 67.2 ppm. HRMS (ESI): calculated mass for $C_{13}H_{23}N_5O_4$: 313.1750, observed mass (M + H): 314.1817, (M + Na): 336.1635.

(2S,4S)-1-(3-Aminopropyl)-4-azidopyrrolidine-2-carboxylic acid (8). To compound **7** (0.3 g, 0.95 mmol), a 50% solution of TFA in DCM (5 mL) was added with vigorous stirring at RT. The reaction was monitored by TLC and after the completion of reaction, solvents were removed under vacuum. The crude compound was used without further purification (0.132 g, 63%). $[\alpha]_D = -13.90$ (c = 0.105 in MeOH). IR (neat) ν: 3377, 2953, 2122, 1683 cm^{-1}. ^{1}H NMR (200 MHz, D$_2$O) δ: 2.08–2.16 (m, 2H), 2.38–2.46 (m, 1H), 2.70–2.85 (m, 1H), 3.04–3.11 (m, 2H), 3.37–3.41 (m, 2H), 3.80–3.86 (m, 1H), 4.25–4.29 (dd, J = 3.5, 10.8 Hz, 1H), 4.56–4.59 (m, 1H) ppm. ^{13}C NMR (50 MHz, D$_2$O) δ: 23.3, 34.1, 36.4, 52.8, 58.8, 60.0, 66.4, 162.7 ppm. ^{13}C DEPT (50 MHz, D$_2$O) δ: 23.3 (CH$_2$), 34.1

(CH$_2$), 36.3 (CH$_2$), 52.8 (CH$_2$), 58.8, 60.0 (CH$_2$), 66.4 ppm. HRMS (ESI): calculated mass for C$_8$H$_{15}$N$_5$O$_2$: 213.1226, observed mass (M + H): 214.1299.

(2S,4S)-4-Amino-1-(3-aminopropyl)pyrrolidine-2-carboxylic acid (9). Compound 8 (0.2 g, 0.61 mmol) was dissolved in methanol, to it palladium on charcoal (0.02 g) was added and the reaction was allowed to proceed under H$_2$ pressure (35–40 psi). After completion of reaction, the solvent was removed under vacuum and the crude compound (0.112 g, 61%) was purified by HPLC. [α]$_D$ = −13.9 (c = 0.1 in MeOH). ^{1}H NMR (400 MHz, D$_2$O) δ: 1.99–2.23 (m, 3H), 2.96–3.09 (m, 3H), 3.20–3.35 (m, 1H), 3.43–3.58 (m, 1H), 3.70–3.81 (m, 1H), 3.91–4.00 (m, 1H), 4.24–4.34 (m, 2H) ppm. ^{13}C NMR (100 MHz, D$_2$O) δ: 23.2, 32.7, 36.2, 47.2, 52.7, 55.8, 67.6, 163.0, 170.1 ppm. ^{13}C DEPT (100 MHz, D$_2$O) δ: 23.2 (CH$_2$), 32.7 (CH$_2$), 36.2 (CH$_2$), 47.2, 52.7 (CH$_2$), 55.8 (CH$_2$), 67.6 ppm. HRMS (ESI): calculated mass for C$_8$H$_{17}$N$_3$O$_2$: 187.1321, observed mass (M + H): 188.1392.

Materials and methods

All the chemicals were procured from Sigma-Aldrich unless otherwise mentioned. TLC analyses were carried out on pre-coated silica gel 60 F$_{254}$ (Merck). Column chromatographic separations were performed using neutral alumina or silica gel (60–120 mesh or 200–400 mesh, Merck) and using the solvent systems EtOAc/petroleum ether or MeOH/DCM. ^{1}H and ^{13}C NMR spectra were obtained using Bruker AC-200, AC-400 NMR spectrometers. The chemical shifts are reported in delta (δ) values and referred to internal standard TMS for ^{1}H. High resolution mass spectra were recorded on a Thermo Fisher Scientific Q Exactive mass spectrometer. Specific rotations of samples were recorded on a Bellingham Stanley ADP220 Polarimeter, IR was recorded on Bruker Alpha spectrophotometer. Primary anti-AGE antibody, anti-CML antibody and secondary antibody conjugate were purchased from Merck Millipore (India).

Circular dichroism measurements. The CD spectra were measured as described earlier.[5] The protein concentration was constant in all the samples (0.02 mg mL^{-1}). CD spectra were acquired on a JASCO J-815 spectropolarimeter at room temperature. The spectra were collected as accumulations of three scans and were corrected for respective blanks. Results are expressed as molar ellipticity, [θ] (deg cm^2 dmol^{-1}). The CD spectra of the protein samples were analysed to calculate the secondary structure content using CDPro software that has three algorithms: CONTINLL, CDSSTR and SELCON3.[17]

***In vitro* glycation of bovine serum albumin (BSA).** The BSA glycation reaction was performed as described earlier by Kańska and Boratyński[18] with minor modification described recently by Kolekar *et al.*[5] with or without title compounds (20 mM) and kept at 37 °C for 10–15 days. *In vitro* glycation was monitored by fluorescence spectroscopy, excitation at 370 nm and emission at 440 nm by using a spectrofluorometer (Thermo, Varioskan Flash Multimode Reader).

IC$_{50}$ determination of representative title compounds. The percent inhibition of AGE formation was calculated using the formula: % inhibition = $(1 - F_i/F_c) \times 100$, where F_i = fluorescence intensity of glycated BSA treated with inhibitor and F_c = fluorescence intensity of glycated BSA in the absence of any inhibitor. The apparent IC$_{50}$ was determined by plotting the percent glycation inhibition *vs.* inhibitor concentration.

Western blot analysis using anti-AGE and anti-CML antibodies. BSA, glycated BSA or glycated BSA with 20 mM of aminoguanidine or 20 mM of compounds 2, 4, 6, 8, and were incubated at 37 °C for 10–15 days. 5 µg of each protein sample was separated by 12% SDS-PAGE and transferred onto a PVDF membrane. The membranes were blocked with 5% skimmed milk powder dissolved in PBS. The proteins were probed by anti-AGE and anti-CML antibodies at the antibody dilution of 1 : 2000 each. Secondary anti-rabbit antibody conjugated with HRP was used at a dilution of 1 : 5000 for both the blots. Immunoreactive bands were visualized using Western Bright Quantum western blotting detection kit (Advansta) and documented by Syngene Imaging System.

MALDI-TOF-based insulin glycation assay. The reaction mixture (100 µL) comprising phosphate buffer (pH 7.4, 0.1 M) containing title compounds (20 mM), insulin (1.8 mg mL^{-1}, 25 µL) and glucose (250 mM, 25 µL) was incubated at 37 °C. The reaction was monitored till the relative intensity of glycated insulin reached 100% as seen by MALDI-TOF-TOF analysis (AB SCIEX TOF/TOF™ 5800) (5 days). The reaction mixture was mixed with sinapinic acid (ratio 1 : 5) and analysed by MALDI-TOF-TOF in linear mode using Anchor Chip 384 targets as described.[1]

Fructosamine assay. The fructosamine level was measured by the NBT Labkit (Chemelex, S.A.) protocol. 300 µL of 0.75 mM NBT was added to a 96-well microplate containing 30 µL of 0.50 µg BSA, glycated BSA, or glycated BSA with title compounds (20 mM). Glycated BSA with aminoguanidine (20 mM) was taken as a control. The reduction of NBT absorbance by fructosamine was measured at 520 nm immediately after additions (considered as Abs1) and after incubation at 37 °C for 15 min (considered as Abs2). The absorbance was monitored by using a UV spectrophometer (UV 1800, Shimadzu).

Adduct formation with glucose. To elucidate the AGE inhibition mechanism of the title compounds, 2, 4, 6, 8, 9 (20 mM each) were incubated with glucose (0.5 M) in phosphate buffer pH 7.4 at 37 °C for 6 days. The reaction was analysed by LC-MS on Q-Exactive Orbitrap to study and detect the glucose adduct formation with the title compounds.

MTT assay for cytotoxicity. L6 rat muscle cells were seeded at a cell density of 1×10^4 cells per well in a 96 well plate. After the cells adhered and attained their morphology, they were serum starved for 4 h and treated with various concentrations of the title compounds (2, 4, 6, 8, and 9) in triplicate for 16 h, while only serum starved cells served as control. After incubation, cells were given one wash with PBS and 100 µL fresh serum free media was added. 6 µL of 5 mg mL^{-1} MTT (dissolved in PBS) was added to each well and incubated in dark at 37 °C until violet formazan crystals were observed. Media from each well was discarded and crystals were dissolved in 100 µL DMSO. Absorbance was measured at 555 nm using Varioskan flash multimode plate reader. This assay was performed using the Vybrant MTT Proliferation Assay Kit from Invitrogen.

Flow cytometry analysis of intracellular reactive oxygen species in *Saccharomyces cerevisiae* BY4743. Levels of intracellular ROS were measured in the presence of title compounds using ascorbic acid as control with 2′,7′-dichlorodihydrofluorescein diacetate (Molecular Probes). Cells were grown in SD medium with or without compounds for 3 days. Briefly, aliquots were taken after 3 days, cells were pelleted down and washed with PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4). Cells were incubated with the dye (10 μM) for 90 min at 30 °C. Cells were then washed twice in PBS and analyzed by flow cytometry analysis on BD Accuri C6 flow cytometer equipped with a 50 mW argon laser emitting at 488 nm. The green fluorescence was collected through a 488 nm blocking filter. Data acquired from a minimum of 10 000 cells per sample at low flow rate were analyzed[19] using BD accuri C6 software.

Orbitrap analysis, database search and PTM analysis. 100 μg of protein was reduced with 100 mM dithiotreitol (DTT) at 60 °C for 15 min, then alkylated with 200 mM iodoacetamide in dark at room temperature for 30 min. Proteins were digested by adding porcine trypsin in the ratio of 1 : 50 (final enzyme : proteins) at 37 °C overnight. The digestion reaction was stopped by adding concentrated HCL and incubated for 10 min at 37 °C before vortexing and centrifugation and analysed on LC MS/MS (Q-Exactive Orbitrap, Thermo). The chromatographic separation was performed as described in our earlier report.[5] The mass spectrometric data were processed using Proteome discoverer 1.4 (Version 1.4.0.288, Thermo Fisher Scientific, Bremen, Germany). The data were searched against UniProt Bovine Serum Albumin (P02769) sequence database. Carbamidomethylation of cysteine (C) and oxidation at methionine (M) was considered as fixed and variable modification respectively. Additionally glycation modifications at lysine position were searched as dynamic variable modification included Amadori (+162.02 Da); Peptide and fragment mass tolerance were 10 ppm, 0.5 Da respectively with minimum of 2 missed cleavages and 1% false discovery rate (1% FDR). The identification of glycation modifications were selected based on the criteria described earlier.[20]

Conclusions

The studies reported herein represent an important contribution towards the search for new molecules that not only inhibit glycation and AGE formation, but are also effective at controlling the concentration of intracellular reactive oxygen species. Moreover, the title compounds are easily accessible synthetically through simple chemical transformations, that should be amenable to scale-up. The superior antiglycation properties of the title compounds have been demonstrated by circular dichroism, fluorescence spectroscopy, western blot assay and also mass spectrometry. Mass spectrometric analysis and the fructosamine assays suggest that the title compounds form adducts with glucose, indicating that they probably act by inhibiting the formation of Amadori products. The low cytotoxicity of the title compounds is yet another favourable attribute. On the whole, the compounds-methyl(2*S*,4*R*)-1-(3-aminopropyl)-4-

hydroxypyrrolidine-2-carboxylate (2) and (2*S*,4*R*)-1-(3-aminopropyl)-4-hydroxypyrrolidine-2-carboxylic acid (4) exhibited the best activity in the present study. These results suggest the contribution of the C4-hydroxyl group towards the antiglycation activity of the title compounds, in addition to the free amino function. The hydroxyl group offers possibilities of hydrogen-bonding and/or metal ion chelation, which could influence the activity of the respective title compound. Further, C4 being a chiral centre, the role of stereochemistry on the observed activity also cannot be ruled out. Studies towards dissecting and discerning these effects are currently underway in our laboratory and will be reported in due course.

References

1 S. K. Kesavan, S. Bhat, S. B. Golegaonkar, M. G. Jagadeeshaprasad, A. B. Deshmukh, H. S. Patil, S. D. Bhosale, M. L. Shaikh, H. V. Thulasiram, R. Boppana and M. J. Kulkarni, *Sci. Rep.*, 2013, **3**, 2941.

2 S. B. Golegaonkar, H. S. Bhonsle, R. Boppana and M. J. Kulkarni, *Eur. J. Mass Spectrom.*, 2010, **16**, 221.

3 S. B. Bansode, A. K. Jana, K. B. Batkulwar, S. D. Warkad, R. S. Joshi, N. Sengupta and M. J. Kulkarni, *PLoS One*, 2014, **9**, e105196.

4 M. M. Joglekar, S. N. Panaskar, A. D. Chougale, M. J. Kulkarni and A. U. Arvindekar, *Mol. BioSyst.*, 2013, **9**, 2463.

5 Y. M. Kolekar, V. Garikapati, S. B. Bansode, B. Santhakumari, H. V. Thulasiram and M. J. Kulkarni, *RSC Adv.*, 2015, **5**, 25051.

6 M. D'Costa, V. A. Kumar and K. N. Ganesh, *Org. Lett.*, 1999, **1**, 1513.

7 (*a*) K. M. Patil, R. J. Naik, M. Vij, A. K. Yadav, V. A. Kumar, M. Ganguli and M. Fernandes, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 4198; (*b*) K. M. Patil, Ph. D. Dissertation, Pune University, India, 2014.

8 S. Ramakrishnan and K. N. Sulochana, *Exp. Eye Res.*, 1993, **57**, 623.

9 F. Monacelli, D. Storace, C. D'Arrigo, R. Sanguineti, R. Borghi, D. Pacini, A. L. Furfaro, M. A. Pronzato, P. Odetti and N. Traverso, *Int. J. Mol. Sci.*, 2013, **14**, 10694.

10 K. Nomoto, M. Yagi, U. Hamada, J. Naito and Y. Yonei, *Anti-Aging Med.*, 2013, **10**, 92.

11 L. Séro, L. Sanguinet, P. Blanchard, B. T. Dang, S. Morel, P. Richomme, D. Séraphin and S. Derbré, *Molecules*, 2013, **18**, 14320.

12 M. P. de la Maza, F. Garrido, N. Escalante, L. Leiva, G. Barrera, S. Schnitzler, M. Zanolli, J. Verdaguer, S. Hirsch, N. Jara and D. Bunout, *J. Diabetes Mellitus*, 2012, **2**, 221.

13 A. Ardestani and R. Yazdanparast, *Int. J. Biol. Macromol.*, 2007, **41**, 572.

14 A. Meeprom, W. Sompong, C. B. Chan and S. Adisakwattana, *Molecules*, 2013, **18**, 6439.

15 J. R. Baker, D. V. Zyzak, S. R. Thorpe and J. W. Baynes, *Clin. Chem.*, 1993, 39, 2460.

16 L. Knels, M. Worm, M. Wendel, C. Roehlecke, E. Kniep and R. H. Funk, *J. Neurochem.*, 2008, **106**, 1876.

17 N. Sreerama and R. W. Woody, *Anal. Biochem.*, 2000, 287, 252.

18 U. Kańska and J. Boratyński, *Arch. Immunol. Ther. Exp.*, 2002, 50, 61.

19 A. Mesquita, M. Weinberger, A. Silva, B. Sampaio-Marques, B. Almeida, C. Leão, V. Costa, F. Rodrigues, W. C. Burhans and P. Ludovico, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 15123.

20 H. S. Bhonsle, A. M. Korwar, S. K. Kesavan, S. D. Bhosale, S. B. Bansode and M. J. Kulkarni, *Eur. J. Mass Spectrom.*, 2012, **18**, 475.

Check for updates

Revisiting amino acids and peptides as anti-glycation agents

H. Chilukuri,[*ab] M. J. Kulkarni [iD][bc] and M. Fernandes [iD][*ab]

The importance of controlling or preventing protein glycation cannot be overstated and is of prime importance in the treatment of diabetes and associated complications including Alzheimer's disease, cataracts, atherosclerosis, kidney aliments among others. In this respect, simple molecules such as amino acids and peptides hold much promise both in terms of ease and scale-up of synthesis as well as in relation to negligible/low associated toxicity. In view of this, a comprehensive account of literature reports is presented, that documents the anti-glycation activity of natural and non-natural amino acids and peptides. This review also discusses the chemical reactions involved in glycation and the formation of advanced glycation end-products (AGEs) and possible/probable intervention sites and mechanism of action of the reported amino acids/peptides. This aspect of amino acids/peptides adds to their growing importance in medicinal and therapeutic applications.

Introduction

Glycation of proteins refers to the reversible reaction of amino groups in proteins, peptides, and lipids with reducing sugars[1] and occurs by a complex series of sequential reactions that involve formation of a Schiff base, Amadori prod-

ucts, and eventually, advanced glycation end-products (AGEs). In conditions of low glucose levels, early glycation products are found to be in equilibrium with plasma glucose, and are capable of dissociating to the native proteins. On the other hand, in conditions of prevalently high glucose levels, molecular rearrangements take place, leading to the irreversible formation of AGEs[2] which contribute to the development of various complications associated with diabetes mellitus (DM), atherosclerosis and neurodegenerative diseases.[3] AGEs have primarily been detected in long-lived proteins, such as collagen and lens crystallins, with post-translational AGE-modifications predominantly occurring at the side-chain amino groups of lysine and arginine residues.[4] This leads to

a Organic Chemistry Division, CSIR-National Chemical Laboratory, Dr. Homi Bhabha Road, Pune 411008, India. E-mail: c.harsha@ncl.res.in, m.dcosta@ncl.res.in
b Academy of Scientific and Innovative Research (AcSIR), CSIR-NCL Campus, Pune - 411008, India
c Proteomics Facility, Division of Biochemical Sciences, CSIR-National Chemical Laboratory, Dr. Homi Bhabha Road, Pune 411008, India

Harsha Chilukuri obtained her Master's degree in Pharmaceutical Chemistry and is currently pursuing her doctoral studies from CSIR-National Chemical Laboratory (CSIR-NCL), Pune, India. She is a recipient of DST-Women Scientists Scheme-A (WOS-A). Her work comprises the design and synthesis of small molecules and peptides with therapeutic potential.

Harsha Chilukuri

Mahesh J. Kulkarni is a Senior Scientist at CSIR-NCL, Pune, India. He obtained his Ph.D. degree from the University of Agricultural Sciences, Bengaluru, India. For the last 10 years, he has been working in the area of mass spectrometry and proteomics. The major focus of his research is to understand the role of advanced glycation end products (AGEs) in the development of diabetic complications. The long-term goal is to identify a diagnostic marker for diabetic complications, identify drug targets and develop intervention strategies.

Mahesh J. Kulkarni

structural alterations of the native protein, thereby impeding both protein function as well as increasing resistance to proteolytic removal.[5] For example, protein glycation has been found to affect the function of cytoskeletal proteins and anti-oxidant enzymes, as well as result in protease-resistant and detergent-insoluble aggregates such as β-amyloid peptide deposits.[4] AGEs, largely *via* the receptor for AGEs (RAGE), activate several signalling mechanisms that cause cell stress, contribute to cellular dysfunction, and damage target organs, leading to complications.[6] RAGE is a transmembrane receptor of the immunoglobulin superfamily, and is prevalent at low concentrations in a variety of healthy human tissues, including the lungs, kidneys, liver, cardiovascular and nervous systems.[7] It is a central signalling molecule in the innate immune system and is involved in the onset and sustainment of the inflammatory response.[8]

Chemistry of glycation

Glycation can be divided into three main stages (Fig. 1). In the initial stage, glucose reacts with an amine from the protein to form a metastable Schiff base which rearranges to the Amadori product. In this phase, glucose shows the slowest glycation rate when compared to other reducing sugars[9] because of the maximal shift of the equilibrium towards its cyclic rather than open-chain aldehyde isoforms, unlike other natural monosaccharides such as ribose. In the intermediate stage, the Amadori product is fragmented to various reactive dicarbonyl compounds such as glyoxal, methylglyoxal and dexoyglucosones.[10] Dicarbonyl compounds react more strongly than their parent sugars with amino groups of proteins to form inter- and intra-molecular cross-links,[11] arginine and lysine being the main amino acid residues involved in cross-linking.[12] The last stage of the reaction involves irreversible AGEs formation through a series of oxidation, dehydration and cyclization reactions. However, not all of the AGEs are derived from protein cross-links. The widely studied N^{ε}-carboxymethyl-lysine (CML), for example, is derived from the modification of a single lysine residue.[13] The heterogeneity of the AGEs formed depends wholly on the glycating

agent. For example, glyoxal-derived AGEs include carboxymethyl-lysine (CML) and glyoxal lysine dimer (GOLD), while methylglyoxal-derived AGEs are carboxyethyl-lysine (CEL), argpyrimidine, and methylglyoxal lysine dimers (MOLD); 3-deoxyglucosone-derived AGEs are pyrraline and deoxyglucosone-derived lysine dimers (DOLD).[10] Accumulation of AGEs is largely seen in long-lived proteins such as lens crystallins and tissue collagens owing to the slow process of their formation.[13]

Inhibitors of AGE formation

The molecular strategies to combat the formation and accumulation of AGEs necessitates consideration of different approaches in order to reduce the AGEs level in the body such as (a) early glycation inhibition (b) advanced glycation inhibition (c) carbonyl quenching (d) masking of lysine residues and (e) inhibition of glycation-induced aggregation,[14] among others. This review is focused on amino acids and peptides against glycation. Besides amino acids and peptides, there exist several small molecules and drugs that possess anti-glycation properties; however, these are beyond the scope of this article.

Amino acids have received considerable attention as therapeutic agents in conditions like respiratory physiology, cardiology, renal failure, neurological disorders, congenital defects, *etc.*[15] Likewise, many peptides have been reported to have properties such as anti-oxidant, anti-hypertensive, anti-atherosclerotic, immunomodulatory[16] and anti-glycation activities[17] in addition to nutrient roles. The role of amino acids and peptides in countering glycation has, however, not been accorded due importance. This review is therefore, aimed towards highlighting the utility benefits and application potential of amino acids and peptides as anti-glycation agents, and hopefully draw the attention of research towards this largely underexploited area.

Amino acid inhibitors of glycation

Some of the regulatory roles of amino acids include gene expression, synthesis and secretion of hormones, nutrient metabolism and oxidative defence, intracellular protein turnover, immune function, reproduction and obesity to name a few.[18] In connection with glycation, free amino acids have been found to mitigate the glycation of lens protein, delay progression of cataract and also bring down blood sugar levels in diabetic rats. Some amino acids inhibit or reduce glycation by hampering the binding of glucose to proteins by competitive inhibition, thereby offering protection, while some amino acids influence pathological pathways resulting in increased tissue sensitivity towards insulin.[19]

Positively charged (cationic) amino acids

Cataract formation, a well-recognized consequence of protein glycation resulting from prolonged hyperglycaemia, was found to be delayed by the anti-glycation effect of free amino

Moneesha Fernandes is a Senior Scientist at CSIR-NCL, Pune, India. Her research interests are in the biomolecular chemistry of peptides, nucleic acids and analogues. Work in her lab focuses on bio-organic chemistry approaches aimed towards the therapeutic and/or diagnostic application of peptides and nucleic acids.
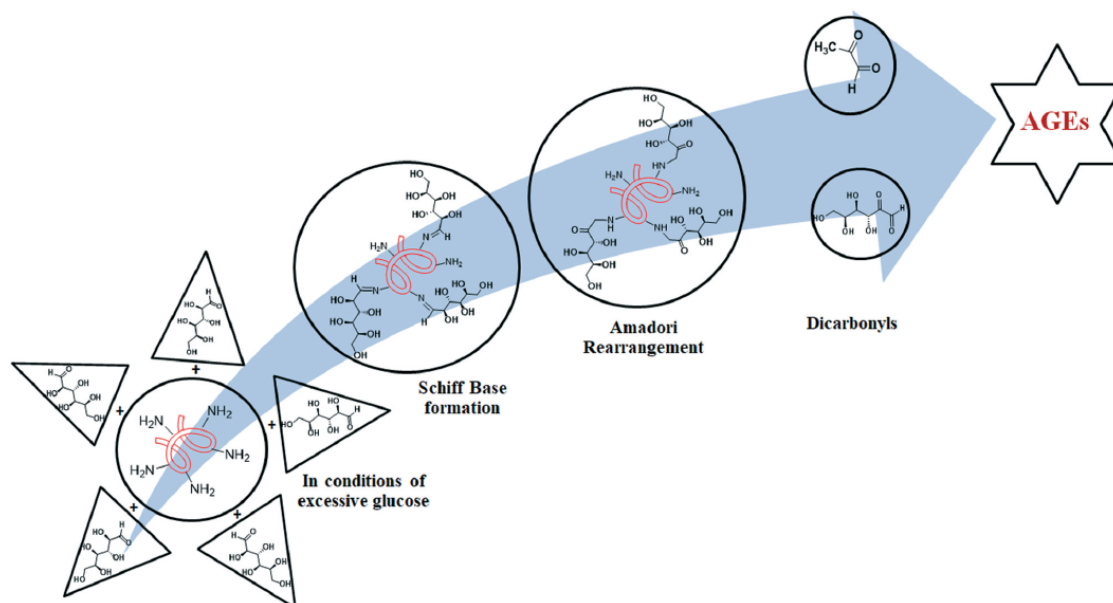
Moneesha Fernandes

Fig. 1 Pictorial representation of glycation.

acid lysine (L-[20,21] and D-lysine[22]). The reduction in the extent of glycation could be attributed to the competitive reaction between free lysine's amino groups and lens protein functionalities with the aldehyde group of glucose forming a Schiff's base, thus scavenging glucose from inside the cells, particularly in hyperglycaemic states.[20] Arginine is known to impede the progression of cataract by acting as an AGE inhibitor and a protein stabilizing agent. The ability of arginine to suppress AGE formation was attributed to its ability to scavenge oxoaldehydes such as glyoxal and methylglyoxal, as well as dehydroascorbic acid (DHA) and its degradation products xylosone, erythrulose, and deoxythreosone.[23] In another study, L-arginine and spermine were found to potently inhibit pyrraline formation in comparison to other polyamines such as putrescine, cadaverine and spermidine *in vitro*.[24] This effect was probably a consequence of the guanidino group and four amino groups in arginine and spermine respectively, suggesting that the anti-glycation effect depended on the number of amino groups present, whereby L-arginine and polyamines are capable of forming a Schiff-base with the carbonyls of open-chain sugars. Glycation inhibition studies under oxidative and non-oxidative conditions led to the conclusion that under oxidative conditions, arginine's anti-glycation property was a result of its guanidinium group, whereas in non-oxidative conditions, it was the α-NH$_2$ group that exerted this effect.[25]

Negatively charged (anionic) amino acids

The anti-glycation property and efficacy of aspartic acid (Asp) to inhibit AGEs formation was established by its protective ef-

fect in shielding the proteins like hemoglobin and albumin from exposure to reducing sugars like glucose, fructose and ribose.[26] Molecular docking studies of BSA–Asp complex revealed hydrophobic – as well as hydrogen bonding interactions between Asp and BSA residues such as Trp, Leu, Val and Arg respectively, to be responsible for the reduced formation of AGEs, which presumably occurs through steric blocking of the reactive Arg guanidine functions from the sugars.

Glycine and hypoglycin A

Glycine exerts anti-diabetic effects in addition to being an insulin secretagogue.[19] It has been reported to counter cataractogenesis by preventing glycation of lens proteins.[20] Alvarado-Vásquez *et al.*[27] reported that glycine and taurine diminished the concentration of glucose, triglycerides, total cholesterol, as well as the non-enzymatic glycation of hemoglobin in streptozotocin-treated rats. Although the specific mechanisms by which glycine or taurine induced diminution of glucose levels remained to be identified, this effect was attributed to the capability of these amino acids of interacting with the insulin receptor. The formation of Schiff base through the reaction of glycine with sugars such as glucose has also been observed,[28] causing a decrease of the free glucose concentration in the medium. Glycine possibly reacts with glucose or dicarbonyl compounds *via* its amino group, thereby resulting in lower production of AGEs.[29]

The unusual amino acid, hypoglycin A (Fig. 2(a)), a plant toxin extracted from unripe fruits and seeds of ackee, *Blighia sapida*, has been reported to cause hypoglycaemia and
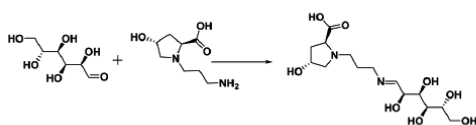
Fig. 5 Probable mechanism for anti-glycation effect of a representative *N*-(3-aminoalkyl)proline derivative.

Fig. 7 Human insulin depicted using one-letter amino acid codes. Disulfide linkages are shown in blue.

parent protein, exert their activity upon release by the action of proteolytic enzymes.[42] The activity largely depends on the inherent amino acid composition and sequence, the length of active sequence varying from two to twenty amino acid residues.[43] Thus peptides may act as alternatives to small molecule drugs offering many advantages such as high bioactivity and biospecificity to targets, wide spectrum of therapeutic action, low levels of toxicity, structural diversity and absence or low levels of accumulation in body tissues.[44]

Dipeptides

The naturally-occurring dipeptide, carnosine (β-alanyl-L-histidine), present in many organisms in muscle and nervous tissues, has been demonstrated to possess anti-glycation activity. Though multiple mechanisms are probably involved, carnosine was found to inhibit AGE formation by quenching reactive carbonyl species.[45] Modifications at the carboxyl terminus or the β-alanine carbon skeleton were found to improve the pharmacokinetic profile without drastically affecting the quenching activity. A C-terminus capping strategy resulted in improved plasma stability as recognition by enzymes was hampered. Conversely, the primary amino group and the imidazole ring were found to play important roles in the observed activity and modifying them greatly reduced the efficiency and/or selectivity towards reactive carbonyl species.[46] Carnosine is also capable of effecting transglycation.[47] Carnosine may play a role in assisting the unfolding of damaged proteins and in solubilizing precipitated protein aggregates, possibly through disruption of complementary binding between AGE-related hydrophobic patches.[48]

γ-Glutamyl-*S*-allyl-cysteine (GSAC, Fig. 6(a)) is a peptide isolated from fresh garlic scales. GSAC attenuated the Maillard reaction during the initial and late stages of glucose-induced protein damage. It significantly reduced structural and functional impairments in proteins caused due to glycation and blocked the exposed hydrophobic surfaces of proteins. Additional iron ion-chelating and radical-
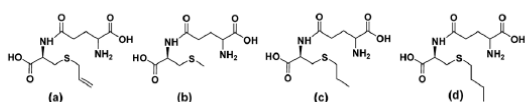
scavenging capacities of the peptide also contribute to its anti-glycation mechanism.[49]

The anti-glycation behavior of γ-glutamylmethylcysteine (γ-GMC) and γ-glutamylpropylcysteine (γ-GPC), extracted from fresh garlic and synthetically prepared γ-glutamylbutylcysteine (γ-GBC) (Fig. 6(b)–(d) respectively) was attributed to their radical scavenging property.[50] This scavenging activity was found to decrease in the order γ-GPC > γ-GMC > γ-GBC.

A computer-aided simulation study of yam dioscorin pepsin hydrolysis resulted in various peptides, of which, the tryptophan-containing dipeptide, Asn-Trp (NW) and its analogue, Gln-Trp (QW), were synthesized and found to possess anti-glycation properties, with NW showing higher activity than QW. Additionally, studies showed that both these dipeptides reduced *in vitro* AGE formation in BSA, although the mechanism is unclear, and could also reduce the production of ROS.[51]

Longer peptides

Animal-derived natural peptides. Insulin[52] (Fig. 7) is highly effective in lowering extremely high glucose levels.[53] Insulin binding to its receptors on insulin-sensitive cells results in activation of a tyrosine kinase which is present intrinsic to the receptor, triggering a signaling cascade and leading to the translocation of glucose transporters from an intracellular location to the cell membrane, which results in internalization and utilization of glucose.[54] While managing hyperglycaemia has greatly improved in recent years, newer strategies focusing on aggressive glucose control have emerged since traditional insulin products do not provide optimal therapy.[55] In spite of their clinical superiority, fundamental safety issues pertaining to these newer analogues are of concern.[56]

Safavi-Hemami and co-workers were the first group to report the presence of insulin (Fig. 8) in the venom of cone snails of the genus *Conus*.[57] The study revealed a class of insulin that acted rapidly and potently to cause severe hypoglycaemia.

Fig. 6 Organosulfur-containing anti-glycation dipeptides – (a) γ-glutamyl-*S*-allyl-cysteine (GSAC), (b) γ-glutamyl-methyl-cysteine (γ-GMC) and (c) γ-glutamyl-propyl-cysteine (γ-GPC) from fresh garlic and synthetically analogous (d) γ-glutamyl-butyl-cysteine (γ-GBC).

Fig. 8 Amino acid sequence of *Conus* insulin with inter- and intra-chain disulfide bridges shown in blue.
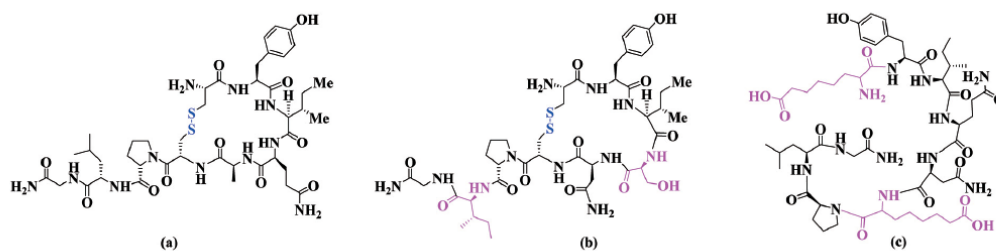
Fig. 9 Oxytocin (a), [Ser4, Ile8]-oxytocin (b) and [Asu1,6]-oxytocin (c).

Oxytocin (Fig. 9(a)), a neuropeptide that plays an important role in labour and lactation, is also known to promote glucose uptake and stimulate insulin secretion.[58–60] Oxytocin analogue [Ser4, Ile8]-oxytocin (Fig. 9(b)) was demonstrated to be superior to native oxytocin in combating glucose intolerance, while another analogue, [Asu1,6]-oxytocin, where Asu is aminosuberic acid (Fig. 9(c)), was found to have similar therapeutic efficacy as oxytocin.[58] The improved activity of [Ser4, Ile8]-oxytocin was suggested to be a consequence of stronger α-helical structure, and consequently, a more stable peptide conformation as a result of Gln → Ser substitution at amino acid position 4.

Plant-derived natural peptides. The purified leaf extract of *Costus igneus* was shown to contain an insulin-like peptide.[61] Interestingly, oral administration of this peptide resulted in better hypoglycaemic activity than intraperitoneal. The hypoglycaemic action was found to occur *via* the insulin signaling pathway.[62]

A novel 68-residue insulin receptor (IR) binding protein was identified from the extract of *Momordica charantia* seeds. The protein was found to bind to the insulin receptor at a site distinct from the insulin-binding site, and could trigger the insulin signal transduction pathway and subsequently lead to reduced hyperglycaemia by enhancing the glucose uptake.[63,64] Residues 50–68 were found to constitute the IR-binding-motif, RVRVWVTERGIVARPPTIG.[65]

Three short peptides from oat hydrolysates-FLQPNLDEH, DLELQNNVFPH and TPNAGVSGAAAGAGAGGKH were reported to reduce hyperglycaemia.[66] These peptides exerted their action by stimulating insulin secretion and insulin sensitivity, and elevating glycogenesis in streptozotocin-induced diabetic mice.

Non-natural peptides. In 2015, a series of analogues of the N-terminal decameric peptide GHPYYSIKKS from *Momordica charantia* L. Var. *abbreviata* Ser. (MCV) were designed, syn-

thesized and evaluated for anti-hyperglycaemic activity, as observed for the parent peptide.[67] This peptide is distinct from the anti-glycating IR-binding peptide from *Momordica charantia* described above, and its mechanism of action is not fully understood. Alanine (Ala) scanning of the parent peptide revealed that the amino acids Pro, Ser, Ile and Ser at positions 3, 6, 7 and 10 respectively, were not necessary for activity. Conformationally constrained analogues of this peptide with a lactam bridge through a Glu-Xaa-Lys scaffold at positions 3, 6 or 10 were found to have increased activity, while introduction of the scaffold at position 7 led to a contrasting effect.[67] Di-proline (-Pro-Pro-) segments, known to be templates for nucleation of folded structures in designed peptides,[68] were further introduced into the peptide at the positions 3, 6, 7 or 10. All the di-proline-containing peptides were found to lower the blood glucose level compared to untreated control mice, with the peptide containing the di-proline segment at position 7 exhibiting significantly higher activity than the parent peptide.[67]

A synthetic selenocysteine analogue of *Conus* insulin, containing a diselenide bond as replacement of the intra-chain disulfide bond (Fig. 10(a)), was found to exhibit activity similar to human insulin.[57] Two analogues of insulin glargine (Fig. 10(b)) containing a 1, 4-disubstituted 1,2,3-triazole group were synthesized to replace the disulfide bridge at positions 7 of the A- and B-chain (Fig. 10(c)) using Cu-assisted azide-alkyne 'Click' (CuAAC) chemistry to efficiently join the peptide chains.[69] However CD studies showed that this modification altered the folding pattern of the native form, resulting in a loss in activity.

Incretins, a group of hormones, stimulate a decrease in blood glucose levels and comprise mainly glucagon-like peptide-1[70] (GLP-1, Fig. 11(a)) and gastric inhibitory peptide or glucose-dependent insulinotropic polypeptide (GIP). The
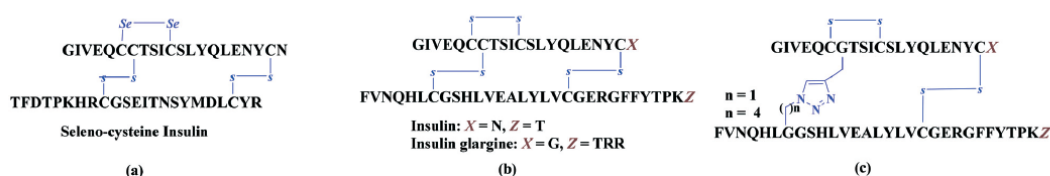

Fig. 10 (a) Seleno-cysteine analogue of *Conus* insulin, (b) insulin and insulin glargine, (c) Triazole analogue of insulin glargine.
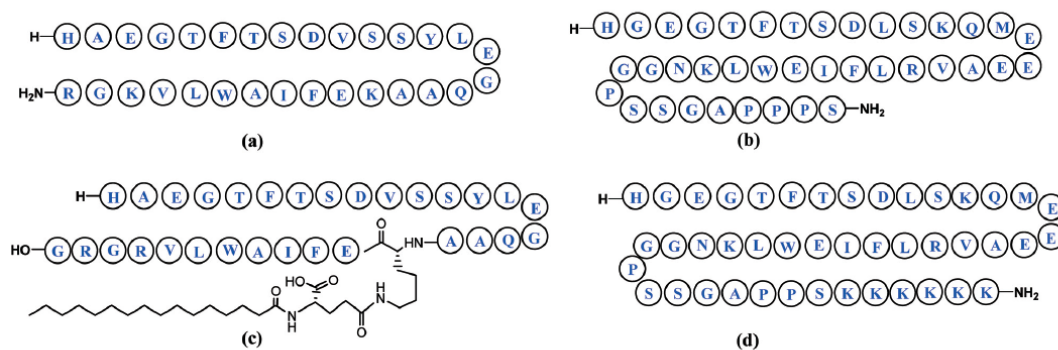
Fig. 11 (a) Glucagon-like peptide-1 (GLP-1)[70] and mimetics (b) exenatide,[77] (c) liraglutide[72] and (d) lixisenatide.[81]

incretin effect may account for 50–70% of total insulin secretion after food intake, playing a crucial role in the maintenance of glycemic control. Incretin mimetics have been intensively developed[71] in anti-diabetic therapies for providing effective glucose control, improved beta-cell function, body weight loss, and lowering of systolic blood pressure.[72] Clinically, GLP-1 has been reported to be very effective in lowering blood glucose levels with very little risk of hypoglycaemia.[73] However, due to rapid inactivation of GLP-1 by a few dominant enzymes like dipeptidyl peptidase IV (DPP IV)[74] and a short lifespan *in vivo* ($t_{1/2} \approx 2$ min), its effectiveness[75] is severely restricted, creating a necessity for the development of longer-acting derivatives.[76]

Exenatide[77] a synthetic analogue of exendin-4,[78] a 39-amino acid peptide produced by the salivary glands of the Gila monster (*Heloderma suspectrum*)[79] (Fig. 11(b)), liraglutide[72] (Fig. 11(c)), lixisenatide[80,81] (Fig. 11(d)), albiglutide, semaglutide, dulaglutide, langlenatide, VRS-859 and CJC-1134-PC are some known GLP-1 agonists.[70] These mimics differ from GLP-1 through the addition of amino acids or other functions, in attempts to increase their stability. Currently GLP-1 receptor agonists available in the market are derived from either the GLP-1 or Exendin-4 backbone.[82] Other exendin-4 analogues reported in literature to possess glucose-lowering potency are conjugates with 2-sulfo-9-fluorenylmethoxycarbonyl group,[83] biotin,[84] hyaluronate (HA),[85] carbohydrate moieties (glycosylated analogues),[86] lithocholic acid,[87] fatty acids (lauric acid or palmitic acid),[88] truncated Evans Blue dye,[89] *etc.*

Further to the discovery of *Xenopus* GLP-1 peptides with potent GLP-1 receptor activation and insulinotropic activities,[90] new *Xenopus* GLP-1 analogues have been reported in literature, such as site-specific mycophenolic acid-modified analogues with differing length of fatty acids,[91] and PEGylated analogues,[92] to name a few.

Thus, it is clear that stable GLP-1 peptide analogues with a long half-life would serve as an efficient therapy option for diabetes.[73] Some noteworthy examples are described in Table 1.

Table 1 Some stable GLP-1 peptide analogues/GLP-1 receptor (GLP-1R) agonists

Sr. no	Study by	Study details	Ref.
1	Deacon *et al.*	N-terminal substitutions with amino acid residues such as threonine, glycine, serine, α-amino isobutyric acid	93
2.	Knudsen *et al.*	Derivatisation with linear fatty acids up to the length of 16 carbon atoms or longer, almost anywhere in the C-terminal region	94
3.	Kaiser *et al.*	Both the N- and C-termini engineered; L-Ala 8 replaced by D-alanine and Arg 36 deleted	95
4.	Knudsen *et al.*	Fatty acid derivatisation and study of effect of polarity and bulkiness	96
5.	Youn *et al.*	Site-specific PEGylation	74
6.	Miranda *et al.*	Evaluation of the effect of helix-favouring amino acid residue substitutions (α-amino-isobutyric acid) and incorporation of a lactam bridge	97
7.	Ueda *et al.*	Glycosylated analogues	98
8.	Mapelli *et al.*	Optimisation of 9-mer peptide closely related to the N-terminal; substituted C-terminal biphenyl dipeptide	99
9.	Haque *et al.*	Study involving 11-amino acid-containing non-natural peptides, including analogues with homohomophenylalanine at the C-terminal	100
10.	Murage *et al.*	Multiple lactam bridges that stabilized both the N- and C-terminal α-helices simultaneously	101
11.	Han *et al.*	Dicoumarol conjugates	75
12.	Johnson *et al.*	Strategic replacement of native α-residues with conformationally constrained β-amino acid residues	102
13.	Yang *et al.*	Structural modification at C-terminal with poly lysine, poly serine and poly valine	103
14.	Hoang *et al.*	Cyclic constraints containing lactam bridges in short hydrophobic peptides	104

Amino acids/peptides in clinical trials for anti-glycation activity

Since the discovery of insulin, its introduction in clinical practice has revolutionized the management of blood glucose levels.[77] Newer insulin analogues like long-acting analogues (glargine, detemir) and rapid-acting analogues (lispro, aspart, glulisine), with subtle changes in their structure, have successfully made it beyond clinical trials and are commercially available in the market.[56] To date, a few stable GLP-1 agonists have received FDA approval as anti-diabetic drugs while some are in clinical trials.[69] Exenatide (approved by the US FDA as the first in the new class of incretin mimic drugs),[78] liraglutide,[72] albiglutide,[70] semaglutide,[105] dulaglutide,[106] *etc.*, that act through the incretin effect, having overcome the clinical limitations of GLP-1, have received FDA approval and are currently available in the market. GLP-1 receptor agonists in clinical trials include, langlenatide, VRS-859, CJC-1134-PC, among others.[70] As for amino acids, a double blind pilot clinical trial exhibited that oral supplementation with free amino acids for patients with type 2 diabetes mellitus appeared to decrease post-prandial plasma glucose without any change in plasma insulin levels.[21] However, no further studies were reported, and to the best of our knowledge, no amino acid or derivative has been approved by the FDA as yet anti glycation.

Discussion

It is evident that the emergence of AGEs and associated complications is a consequence of persistent hyperglycaemia. This may be countered by multiple approaches – [1] increasing the glucose utilization/uptake, which in turn, could be achieved by increasing insulin sensitivity or stimulating its secretion (insulin secretagogues) or increasing its activity and production (insulinotropic agents). This may be achieved by insulin and analogues, which includes those that bind allosterically to the insulin receptor. [2] Decreasing gluconeogenesis. [3] Suppressing glucagon secretion. [4] Decreasing the extent of AGE formation in proteins, through competitive binding to glucose or reactive carbonyls. [5] Reduction in AGE formation through transglycation or reversal of glycation. [6] Shielding the exposed regions of proteins from exposure to sugars. [7] Preventing protein aggregation by increasing solubilization. [8] Ion-chelating. [9] Through the associated anti-oxidant properties of the agents. The heterogeneity of the AGE products, ranging from well-defined structures such as protein-linked – or free pentosidine and CML to ill-characterized AGE peptides, makes the study of their fate *in vivo* a difficult task.[107] As a consequence, little attention has been paid to AGE disposal from the body.

The amino acids and derivatives prevent the formation of AGEs predominantly through competitive Schiff base formation by their amino groups, rather than the amino groups of proteins, with the carbonyls of sugars, trapping of reactive carbonyls, radical scavenging and through their anti-oxidant properties. These effects are more of a general nature. More specific mechanisms by which amino acids and derivatives act to prevent glycation are by binding to the insulin receptor (taurine, glycine) and acting as insulin secretagogues (glycine, cysteine). 4-Hydroxy-isoleucine is reported to be insulinotropic, while hypoglycin A indirectly effects anti-glycation through causing lower glucose levels by decreasing gluconeogenesis as a result of impaired long chain fatty acid metabolism. Some amino acids such as arginine also help to counter the effects of glycation through prevention of protein aggregation and increasing solubilization. Aspartic acid could shield the exposed hydrophobic regions of proteins from exposure to sugars and thus, from potential aggregation. Many of the amino acids and derivatives also possess considerable anti-oxidant activity, contributing to their potential utility value against glycation. The dipeptide carnosine is known to act by multiple mechanisms including an ion-chelating effect. Chelation of transition metal ions such as copper and iron by synthetic or natural compounds is one of the promising mechanisms for the inhibition of AGEs formation because transition metals, in presence of oxygen, catalyze autoxidation of glucose or lipid peroxidation and also assist the formation of oxygen free radicals such as hydroxyl radicals during Fenton reaction in states of hyperglycaemia.[108] Carnosine is also involved in neutralization of reactive carbonyls, preventing protein aggregation and in the reversal of glycation through transglycation. The dipeptide GSAC was found to effectively shield exposed hydrophobic regions of proteins, preventing their aggregation, in addition to inhibiting early and late-stage Maillard reactions between protein and sugars, and having a radical-scavenging ability, this last property being shared with other dipeptides from garlic – γ-GMC and γ-GPC.

The anti-glycation effect of peptides described herein is more specific in nature, and is mostly insulin-related. Thus peptides and synthetic analogues could act as insulin secretagogues (oxytocin and oat hydrolysate peptides), bind to insulin receptor (insulin and analogues and insulin receptor-binding protein from *Momordica charantia*) and/or cause increased glucose uptake and utilization. Alternatively, peptides could act through the incretin effect, as incretin mimetics (*e.g.*, GLP-1 mimetics), to bring about a decrease in the levels of blood glucose.

Scope for further development

The present review summarizes findings on the efficacy of some natural and non-natural amino acids and peptides to reduce, delay or even prevent complications that arise due to hyperglycaemia. They can be used to reduce glycaemia directly by inhibiting various stages of the Maillard reaction, while some of them act as insulin secretagogues or incretin mimetics, thereby reducing glucose levels. Considering the varied modes of action of the various amino acids and peptides described herein, it appears that application of a combination of mechanistic approaches would be desirable. Thus, the potency of amino acids in swiftly reducing the reactive

This journal is © The Royal Society of Chemistry 2018

Med. Chem. Commun., 2018, **9**, 614–624 | **621**

sugar concentrations by Schiff base formation, for example, would benefit from being coupled to more specific receptor-based approaches, such as those involved in glucose uptake and utilization. If these can also be expanded to include anti-oxidant properties and prevention of protein aggregation, one can expect significantly improved anti-glycation, both in terms of potency as well as specificity. We believe that these would constitute directions for future research in this area, thus contributing to the already growing therapeutic appeal of peptides.

Conflicts of interest

The authors have no conflicts of interest to declare.

Acknowledgements

References

1 A. Frolov, R. Schmidt, S. Spiller, U. Greifenhagen, R. Hoffmann and J. Agric, *Food Chem.*, 2014, 62, 3626–3635.

2 A. M. Schmidt, O. Hori, J. Brett, S. D. Yan, J. L. Wautier and D. Stern, *Arterioscler., Thromb., Vasc. Biol.*, 1994, 14, 1521–1528.

3 Q. Zhang, J. M. Ames, R. D. Smith, J. W. Baynes and T. O. Metz, *J. Proteome Res.*, 2009, 8, 754–769.

4 H. Kaur, M. Kamalov and M. A. Brimble, *Acc. Chem. Res.*, 2016, 49, 2199–2208.

5 R. Singh, A. Barden, T. Mori and L. Beilin, *Diabetologia*, 2001, 44, 129–146.

6 R. Ramasamy, S. F. Yan and A. M. Schmidt, *Ann. N. Y. Acad. Sci.*, 2011, 1243, 88–102.

7 S. Bongarzone, V. Savickas, F. Luzi and A. D. Gee, *J. Med. Chem.*, 2017, 60, 7213–7232.

8 G. Fritz, *Trends Biochem. Sci.*, 2011, 36, 625–632.

9 X. Peng, J. Ma, F. Chen and M. Wang, *Food Funct.*, 2011, 2, 289–301.

10 A. A. Chinchansure, A. M. Korwar, M. J. Kulkarni and S. P. Joshi, *RSC Adv.*, 2015, 5, 31113–31138.

11 S. Alam, A. Ahsan and S. Alam, *J. Biochem. Technol.*, 2013, 5, 666–672.

12 S. Seo and S. Karboune, *J. Agric. Food Chem.*, 2014, 62, 12235–12243.

13 V. P. Reddy and A. Beyaz, *Drug Discovery Today*, 2006, 11, 646–654.

14 S. Awasthi and N. T. Saraswathi, *RSC Adv.*, 2016, 6, 24557–24564.

15 C. S. Shantharam, D. M. S. Vardhan, R. Suhas, M. B. Sridhara and D. C. Gowda, *Eur. J. Med. Chem.*, 2013, 60, 325–332.

16 C. H. Han, Y. S. Lin, S. Y. Lin and W. C. Hou, *Food Chem.*, 2014, 147, 195–202.

17 F. Shi, B. Bai, S. Ma, S. Ji and L. Liu, *Food Chem.*, 2016, 194, 538–544.

18 G. Wu, *Amino Acids*, 2009, 37, 1–17.

19 C. V. Anuradha, *Curr. Protein Pept. Sci.*, 2009, 10, 8–17.

20 S. Ramakrishna and K. N. Sulochana, *Exp. Eye Res.*, 1993, 57, 623–628.

21 K. N. Sulochana, S. Lakshmi, R. Punitham, T. Arokiasamy, B. Sukumar and S. Ramakrishnan, *Med. Sci. Monit.*, 2002, 8, CR131–CR137.

22 M. Sensi, F. Pricci, M. G. De Rossi, S. Morano and U. Di Mario, *Clin. Chem.*, 1987, 35, 384–387.

23 X. Fan, L. Xiaoqin, B. Potts, C. M. Strauch, I. Nemet and V. M. Monnier, *Mol. Vision*, 2011, 17, 2221–2227.

24 J. D. Méndez and L. I. Leal, *Biomed. Pharmacother.*, 2004, 58, 598–604.

25 D. A. Servetnick, D. Bryant, K. J. Wells-Knecht and P. L. Wiesenfeld, *Amino Acids*, 1996, 11, 69–81.

26 G. Prasanna and N. T. Saraswathi, *J. Biomol. Struct. Dyn.*, 2016, 34, 943–951.

27 N. Alvarado-Vásquez, P. Zamudio, E. Ceron, B. Vanda, E. Zenteno and G. C. Sandoval, *Comp. Biochem. Physiol., Part C: Toxicol. Pharmacol.*, 2003, 134, 521–527.

28 F. Bahmani, S. Z. Bathaie, S. J. Aldavood and A. Ghahghaei, *Mol. Vision*, 2012, 18, 439–448.

29 S. A. Noe, G. L. Mario, G. D. Reyes, V. J. E. Iván, A. A. F. Javier and G. O. J. Luis, *J. Exp. Clin. Med.*, 2013, 5, 109–114.

30 K. Tanaka, *J. Biol. Chem.*, 1972, 247, 7465–7478.

31 P. Manna, J. Das and P. C. Sil, *Curr. Diabetes Rev.*, 2013, 9, 237–248.

32 A. T. A. Nandhini, V. Thirunavukkarasu and C. V. Anuradha, *Indian J. Med. Res.*, 2005, 122, 171–177.

33 J. Maturo and E. C. Kulakowski, *Biochem. Pharmacol.*, 1988, 37, 3755–3760.

34 P. S. Devamanoharan, A. H. Ali and S. D. Varma, *Mol. Cell. Biochem.*, 1997, 177, 245–250.

35 S. Mahdavifard, S. Z. Bathaie, M. Nakhjavani and H. Heidarzadeh, *Food Res. Int.*, 2014, 62, 909–916.

36 Y. Sauvaire, P. Petit, C. Broca, M. Manteghetti, Y. Baissac, J. F. Alvarez, R. Gross, M. Roye, A. Leconte, R. Gomis and G. Ribes, *Diabetes*, 1998, 47, 206–210.

37 C. Broca, M. Manteghetti, R. Gross, Y. Baissac, M. Jacob, P. Petit, Y. Sauvaire and G. Ribes, *Eur. J. Pharmacol.*, 2000, 390, 339–345.

38 P. Rajasekar and C. V. Anuradha, *Acta Diabetol*, 2007, 44, 83–90.

39 C. S. Shantharam, D. M. S. Vardhan, R. Suhas, M. B. Sridhara and D. C. Gowda, *Eur. J. Med. Chem.*, 2013, 60, 325–332.

40 D. M. S. Vardhan, C. S. Shantharam, R. Suhas and D. C. Gowda, *J. Saudi Chem. Soc.*, 2017, 21, S248–S257.

41 H. Chilukuri, Y. M. Kolekar, G. S. Bhosle, R. K. Godbole, R. S. Kazi, M. J. Kulkarni and M. Fernandes, *RSC Adv.*, 2015, 5, 77332–77340.

42 H. Korhonen and A. Pihlanto, *Int. Dairy J.*, 2006, 16, 945–960.

622 | *Med. Chem. Commun.*, 2018, 9, 614–624

This journal is © The Royal Society of Chemistry 2018

43   H. Meisel and R. J. Fitz Gerald, *Curr. Pharm. Des.*, 2003, 9, 1289–1295.

44   D. Agyei and M. K. Danquah, *Biotechnol. Adv.*, 2011, 29, 272–277.

45   E. D. Pepper, M. J. Farrell, G. Nord and S. E. Finkel, *Appl. Environ. Microbiol.*, 2010, 76, 7925–7930.

46   G. Vistoli, D. De Maddis, V. Straniero, A. Pedretti, M. Pallavicini, E. Valoti, M. Carini, B. Testa and G. Aldini, *Eur. J. Med. Chem.*, 2013, 66, 153–160.

47   B. S. Szwergold, *Biochem. Biophys. Res. Commun.*, 2005, 336, 36–41.

48   N. W. Seidler, G. S. Yeargans and T. G. Morgan, *Arch. Biochem. Biophys.*, 2004, 427, 110–115.

49   D. Tan, Y. Zhang, L. Chen, L. Liu, X. Zhang, Z. Wu, B. Bai and S. Ji, *Nat. Prod. Res.*, 2015, 29, 2219–2222.

50   F. Shi, B. Bai, S. Ma, S. Ji and L. Liu, *Food Chem.*, 2016, 194, 538–544.

51   C. H. Han, Y. S. Lin, S. Y. Lin and W. C. Hou, *Food Chem.*, 2014, 147, 195–202.

52   F. G. Banting, C. H. Best, J. B. Collip, W. R. Campbell and A. A. Fletcher, *Can. Med. Assoc. J.*, 1991, 145, 1281–1286.

53   I. B. Hirsch, R. M. Bergenstal, C. G. Parkin, E. Wright Jr. and J. B. Buse, *Clin. Diabetes*, 2005, 23, 78–86.

54   C. R. Kahn, *Annu. Rev. Med.*, 1985, 36, 429–451.

55   I. Hartman, *Clin. Med. Res.*, 2008, 6, 54–67.

56   V. Valla, *Exp. Diabetes Res.*, 2010, 2010, 178372, DOI: 10.1155/2010/178372, 14 pages.

57   H. Safavi-Hemami, J. Gajewiak, S. Karanth, S. D. Robinson, B. Ueberheide, A. D. Douglass, A. Schlegel, J. S. Imperial, M. Watkins, P. K. Bandyopadhyay, M. Yandell, Q. Li, A. W. Purcell, R. S. Norton, L. Ellgaard and B. M. Olivera, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, 112, 1743–1748.

58   H. Zhang, C. Wu, Q. Chen, X. Chen, Z. Xu, J. Wu and D. Cai, *PLoS One*, 2013, 8, e61477.

59   E. S. Lee, K.-O. Uhm, Y. M. Lee, J. Kwon, S.-H. Park and K. H. Soo, *Regul. Pept.*, 2008, 151, 71–74.

60   S. Elabd and I. Sabry, *Front. Endocrinol.*, 2015, 6, 121, DOI: 10.3389/fendo.2015.00121.

61   B. N. Joshi, H. Munot, M. Hardikar and A. A. Kulkarni, *Biochem. Biophys. Res. Commun.*, 2013, 436, 278–282.

62   M. R. Hardikar, M. E. Varma, A. A. Kulkarni, P. P. Kulkarni and B. N. Joshi, *Phytochemistry*, 2016, 124, 99–107.

63   H.-Y. Lo, T.-Y. Ho, C. Lin, C.-C. Li and C.-Y. Hsiang, *J. Agric. Food Chem.*, 2013, 61, 2461–2468.

64   H.-Y. Lo, T.-Y. Ho, C.-C. Li, J.-C. Chen, J.-J. Liu and C.-Y. Hsiang, *J. Agric. Food Chem.*, 2014, 62, 8952–8961.

65   H.-Y. Lo, C.-C. Li, T.-Y. Ho and C.-Y. Hsiang, *Food Chem.*, 2016, 204, 298–305.

66   H. Zhang, J. Wang, Y. Liu and B. Sun, *J. Biomed. Sci.*, 2015, 4, 1–7.

67   B. Yang, X. Li, C. Zhang, S. Yan, W. Wei, X. Wang, X. Deng, H. Qian, H. Lin and W. Huang, *Org. Biomol. Chem.*, 2015, 13, 4551–4561.

68   B. Chatterjee, I. Saha, S. Raghotama, S. Aravinda, R. Rai, N. Shamala and P. Balaram, *Chem. – Eur. J.*, 2008, 14, 6192–6204.

69   G. M. Williams, K. Lee, X. Li, G. J. S. Cooper and M. A. Brimble, *Org. Biomol. Chem.*, 2015, 13, 4059–4063.

70   B. Manandhar and J.-M. Ahn, *J. Med. Chem.*, 2015, 58, 1020–1037.

71   S. Son, S. Y. Chae, C. W. Kim, Y. G. Choi, S. Y. Jung, S. Lee and K. C. Lee, *J. Med. Chem.*, 2009, 52, 6889–6896.

72   J. Lau, P. Bloch, L. Schaffer, I. Pettersson, J. Spetzler, J. Kofoed, K. Madsen, L. B. Knudsen, J. McGuire, D. B. Steensgaard, H. M. Strauss, D. X. Gram, S. M. Knudsen, F. S. Nielsen, P. Thygesen, S. Reedtz-Runge and T. Kruse, *J. Med. Chem.*, 2015, 58, 7370–7380.

73   L. B. Knudsen, *J. Med. Chem.*, 2004, 47, 4128–4134.

74   Y. S. Youn, S. Y. Chae, S. Lee, J. E. Jeon, H. G. Shin and K. C. Lee, *Biochem. Pharmacol.*, 2007, 73, 84–93.

75   J. Han, L. Sun, Y. Chu, Z. Li, D. Huang, X. Zhu, H. Qian and W. Huang, *J. Med. Chem.*, 2013, 56, 9955–9968.

76   A. J. Garber, *Diabetes Care*, 2011, 34, S279–S284.

77   S. Pechenov, H. Bhattacharjee, D. Yin, S. Mittal and J. A. Subramony, *Adv. Drug Delivery Rev.*, 2017, 112, 106–122.

78   M. B. Davidson, G. Bate and P. Kirkpatrick, *Nat. Rev. Drug Discovery*, 2005, 4, 713–714.

79   A. Evers, T. Haack, M. Lorenz, M. Bossart, R. Elvert, B. Henkel, S. Stengelin, M. Kurz, M. Glien, A. Dudda, K. Lorenz, D. Kadereit and M. Wagner, *J. Med. Chem.*, 2017, 60, 4293–4303.

80   M. Christensen, F. K. Knop, J. J. Holst and T. Vilsboll, *IDrugs*, 2009, 12, 503–513.

81   J. Han, Y. Fei, F. Zhou, X. Chen, W. Zheng and J. Fu, *Mol. Pharmaceutics*, 2017, 14, 3954–3967.

82   E. J. Park, S. M. Lim, K. C. Lee and D. H. Na, *Expert Opin. Ther. Pat.*, 2016, 26, 833–842.

83   Y. Shechter, H. Tsubery and M. Fridkin, *Biochem. Biophys. Res. Commun.*, 2003, 305, 386–391.

84   C.-H. Jin, S. Y. Chae, S. Son, T. H. Kim, K. A. Um, Y. S. Youn, S. Lee and K. C. Lee, *J. Controlled Release*, 2009, 133, 172–177.

85   J.-H. Kong, E. J. Oh, S. Y. Chae, K. C. Lee and S. K. Hahn, *Biomaterials*, 2010, 31, 4121–4128.

86   T. Ueda, T. Ito, K. Tomita, H. Togame, M. Fumoto, K. Asakura, T. Oshima, S.-I. Nishimura and K. Hanasaki, *Bioorg. Med. Chem. Lett.*, 2010, 20, 4631–4634.

87   S. Y. Chae, C.-H. Jin, J. H. Shin, S. Son, T. H. Kim, S. Lee, Y. S. Youn, Y. Byun, M.-S. Lee and K. C. Lee, *J. Controlled Release*, 2010, 142, 206–213.

88   S. Y. Chae, Y. G. Choi, S. Son, S. Y. Jung, D. S. Lee and K. C. Lee, *J. Controlled Release*, 2010, 144, 10–16.

89   Y. Liu, G. Wang, H. Zhang, Y. Ma, L. Lang, O. Jacobson, D. O. Kiesewetter, L. Zhu, S. Gao, Q. Ma and X. Chen, *Bioconjugate Chem.*, 2016, 27, 54–58.

90   D. M. Irwin, M. Satkunarajah, Y. Wen, P. L. Brubaker, R. A. Pederson and M. B. Wheeler, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, 94, 7915–7920.

91   J. Han, J. Fu, L. Sun, Y. Han, Q. Mao, F. Liao, X. Zhenga and K. Zhu, *Med. Chem. Commun.*, 2018, 9, 67–80.

This journal is © The Royal Society of Chemistry 2018

Med. Chem. Commun., 2018, 9, 614–624 | 623

92 J. Han, Y. Wang, Q. Meng, G. Li, F. Huang, S. Wu, Y. Fei, F. Zhou and J. Fu, *Eur. J. Med. Chem.*, 2017, **132**, 81–89.

93 C. F. Deacon, L. B. Knudsen, K. Madsen, F. C. Wiberg, O. Jacobsen and J. J. Holst, *Diabetologia*, 1998, **41**, 271–278.

94 L. B. Knudsen, P. F. Nielsen, P. O. Huusfeldt, N. L. Johansen, K. Madsen, F. Z. Pedersen, H. Thogersen, M. Wilken and H. Agerso, *J. Med. Chem.*, 2000, **43**, 1664–1669.

95 G. Uckaya, P. Delagrange, A. Chavanieu, G. Grassy, M.-F. Berthault, A. Ktorza, E. Cerasi, G. Leibowitz and N. Kaiser, *J. Endocrinol.*, 2005, **184**, 505–513.

96 K. Madsen, L. B. Knudsen, H. Agersoe, P. F. Nielsen, H. Thogersen, M. Wilken and N. L. Johansen, *J. Med. Chem.*, 2007, **50**, 6126–6132.

97 L. P. Miranda, K. A. Winters, C. V. Gegg, A. Patel, J. Aral, J. Long, J. Zhang, S. Diamond, M. Guido, S. Stanislaus, M. Ma, H. Li, M. J. Rose, L. Poppe and M. M. Veniant, *J. Med. Chem.*, 2008, **51**, 2758–2765.

98 T. Ueda, K. Tomita, Y. Notsu, T. Ito, M. Fumoto, T. Takakura, H. Nagatome, A. Takimoto, S.-I. Mihara, H. Togame, K. Kawamoto, T. Iwasaki, K. Asakura, T. Oshima, K. Hanasaki, S.-I. Nishimura and H. Kondo, *J. Am. Chem. Soc.*, 2009, **131**, 6237–6245.

99 C. Mapelli, S. I. Natarajan, J.-P. Meyer, M. M. Bastos, M. S. Bernatowicz, V. G. Lee, J. Pluscec, D. J. Riexinger, E. S. Sieber-McMaster, K. L. Constantine, C. A. Smith-Monroy, R. Golla, Z. Ma, D. A. Longhi, D. Shi, L. Xin, J. R. Taylor, B. Koplowitz, C. L. Chi, A. Khanna, G. W. Robinson, R. Seethala, I. A. Antal-Zimanyi, R. H. Stoffel, S. Han, J. M. Whaley, C. S. Huang, J. Krupinski and W. R. Ewing, *J. Med. Chem.*, 2009, **52**, 7788–7799.

100 T. S. Haque, V. G. Lee, D. Riexinger, M. Lei, S. Malmstrom, L. Xin, S. Han, C. Mapelli, C. B. Coope, G. Zhang, W. R. Ewing and J. Krupinski, *Peptides*, 2010, **31**, 950–955.

101 E. N. Murage, G. Gao, A. Bisello and J.-M. Ahn, *J. Med. Chem.*, 2010, **53**, 6412–6420.

102 L. M. Johnson, S. Barrick, M. V. Hager, A. McFedries, E. A. Homan, M. E. Rabaglia, M. P. Keller, A. D. Attie, A. Saghatelian, A. Bisello and S. H. Gellman, *J. Am. Chem. Soc.*, 2014, **136**, 12848–12851.

103 X. Yang, Y. Li, Y. Wang, X. Zheng, W. Kong, F. Meng, Z. Zhou, C. Liu, Y. Li and M. Gong, *Mol. Pharmaceutics*, 2014, **11**, 4092–4099.

104 H. N. Hoang, K. Song, T. A. Hill, D. R. Derksen, D. J. Edmonds, W. M. Kok, C. Limberakis, S. Liras, P. M. Loria, V. Mascitti, A. M. Mathiowetz, J. M. Mitchell, D. W. Piotrowski, D. A. Price, R. V. Stanton, J. Y. Suen, J. M. Withka, D. A. Griffith and D. P. Fairlie, *J. Med. Chem.*, 2015, **58**, 4080–4085.

105 Novo Nordisk: Company Announcement of the FDA approval of Semaglutide (Ozempic®) in the USA, Dec 2017, https://www.novonordisk.com/bin/getPDF.2154210.pdf.

106 M. Sanford, *Drugs*, 2014, **74**, 2097–2103.

107 T. Miyata, Y. Ueda, K. Horie, M. Nangaku, S. Tanaka, C. V. Y. De Strihou and K. Kurokawa, *Kidney Int.*, 1998, **53**, 416–422.

108 S. Khangholi, F. A. A. Majid, N. J. A. Berwary, F. Ahmad and R. B. A. Aziz, *Planta Med.*, 2016, **82**, 32–45.