QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIP STUDIES ON ANTIDIABETIC AGENTS

A Thesis
Submitted for the Award of Ph.D. degree
in Chemistry
(Faculty of Science)
to the
University of Kota, Kota

By

Raghuraj Parihar

Under the Supervision of

Dr. Brij Kishore Sharma
Department of Chemistry
Government College, Bundi (Raj.)

UNIVERSITY OF KOTA, KOTA

2020

CERTIFICATE

I feel great pleasure in certifying that the thesis entitled "**QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIP STUDIES ON ANTIDIABETIC AGENTS**" is an original piece of work carried out by **Raghuraj Parihar**, under my supervision for the degree of DOCTOR OF PHILOSOPHY. He has completed the following requirements as per Ph. D. regulations of the University.

(i) Course work as per the University rules.

(ii) Residential requirements of the University (200 days).

(iii) Regularly submitted annual progress report.

(iv) Presented his work in the departmental committee.

(v) Published/accepted minimum of one research paper in a referred research journal.

I recommend the submission of the thesis.

Place: Bundi
Date: / /2020

(Brij Kishore Sharma)
Research Supervisor
Associate Professor
Department of Chemistry
Government College, Bundi

ANTI-PLAGIARISM CERTIFICATE

It is certified that the Doctoral thesis entitled "**QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIP STUDIES ON ANTIDIABETIC AGENTS**" by Raghuraj Parihar has been examined by us with the following anti-plagiarism tools. We undertake the follows:

a. Thesis has significant new work/knowledge as compared to already published or are under consideration to be published elsewhere. No sentences, equation, diagram, table, paragraph or section has been copied verbatim from previous work unless it is placed under quotation marks and dully referenced.

b. The work presented is original and own work of the author (i.e., there is no plagiarism). No ideas, processes, result or words of others have been presented as author' own work.

c. There is no fabrication of data or results which have been compiled and analyzed.

d. There is no falsification by manipulating research materials, equipment or processes, or changing or omitting data or results such that the research is not accurately represented in the research record.

e. The thesis has been checked by using ANTI-PLAGIARISM URKUND software and found within limits as per HEC plagiarism policy and instruction issued from time to time.

(Raghuraj Parihar)
Research Scholar
Place: Bundi
Date: / /2020

(Dr. Brij Kishore Sharma)
Research Supervisor
Place: Bundi
Date: / /2020

ABSTRACT

Diabetes mellitus (DM), a very common metabolic disease, is affecting millions of the people around the globe. Due to globalization, mechanization, and changes in human lifestyle and daily routines incidences of diabetes are continuously increasing. The objective of the present study is to investigate the quantitative structure-activity relationships of antidiabetic agents with regard to the future development of such ligands as agonists and antagonists.

The QSAR studies on the imidazolopyrimidine amides, the (2*S*)-cyanopyrrolidine analogues and the derivatives of β-aminoamide bearing subsituted triazolopiperazines as dipeptidyl peptidase IV (DPP-4) inhibitors; PPARγ transactivation profiles of the derivatives of tetrahydroquinolines, benzylpyrazole acylsulfonamides and pyridyloxybenzene-acylsulfonamides; and GPR119 agonistic activity of indole-based derivatives and triazolopyridines provided a rational approach for the development of titled derivatives as inhibitors or agonists. The QSAR rationales for these analogous were obtained in terms of Dragon descriptors using Combinatorial Protocol in Multiple Linear Regression (CP-MLR). The biological actions of these compounds have appeared as the cumulative influence of different structural features which were recognized in terms of individual descriptors.

CANDIDATE'S DECLARATION

I, hereby, certify that the work, which is being presented in the thesis, entitled "**QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIP STUDIES ON ANTIDIABETIC AGENTS**" in partial fulfillment of the requirement for the award of the Degree of Doctor of Philosophy, carried out under the supervision of Dr. Brij Kishore Sharma, Associate Professor, Department of Chemistry, Government College Bundi and submitted to the University of Kota, Kota represents my ideas in my own words and where others ideas or words have been included I have adequately cited and referenced the original sources. The work presented in this thesis has not been submitted elsewhere for the award of any other degree or diploma from any Institutions.

I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will cause for disciplinary action by the University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Date: / /2020 (Raghuraj Parihar)

Place: Bundi Research Scholar

This is to certify that the above statement made by Raghuraj Parihar (Registration No. RS/1755/13) is correct to the best of my knowledge.

Date: / /2020. (Dr. Brij Kishore Sharma)

Place: Bundi Research Supervisor

PREFACE

Diabetes mellitus (DM), a very common metabolic disease, is affecting millions of the people around the globe. Due to globalization, mechanization, and changes in human lifestyle and daily routines incidences of diabetes are continuously increasing. β-Cells secrete insulin in islets of Langerhans as a response of elevated blood glucose. A severe increase in blood glucose induces a rapid release of insulin which is sustained for a short period (known as 1st phase) and then followed by longer period of lower secretion (the 2nd phase)) which accounts for the most part of secretion of insulin. Diabetes and its complications are the outcomes of progressive reduction in β-cell mass or secretory capacity resulting to abnormal glucose metabolism. Thus, DM is an entity of considerable morbidity consisting of a spectrum of multisystem dysfunctions arose from the combination of insulin resistance and inadequate insulin secretion. Type 1 diabetes is caused usually by immune destruction of pancreatic islet cells, on the other hand type 2 is associated with metabolic conditions such as the insulin resistance, hyperglycemia, hypertension, obesity and hyperlipidemia. It is just like a tightrope walk, to manage diabetes as it have need of an ample understanding of numerous factors such as over-all clinical picture, profile related to adverse effects, the multifaceted inter-play of drugs, etc. The ranges of new antidiabetic drugs are continuously increasing with targeted novel facets of diabetes and that calls for ample consciousness by the treating clinicians. New therapeutics must be aimed at to treat diabetic patients at an earlier stage of the disease and able to address the multi-factorial nature of DM.

The objective of the present study is to investigate the quantitative structure-activity relationships of antidiabetic agents with regard to the future development of such ligands as agonists and antagonists. This, in turn, would assist in further proposing the possible mode of action, at molecular level, of these agents at the receptor of concern.

The work embodied in this thesis is focused on the derivation of various QSAR models, their validation, interpretation and forecasting new potential

congeners, if appropriate. Sometimes, the mechanism of drug-receptor interaction has also been addressed. The work carried out is arranged into five chapters in this thesis. A brief introduction to the modeling techniques and the general methodology used to develop QSAR models is presented in the first chapter.

The second chapter gives an overview on the antidiabetic agents, based on the current FDA recommendations. Due to the elevated CVD risk in DM, all new anti-diabetic drugs show exemplary cardiovascular safety profiles. Thus, drugs that target molecular pathways having potential implications in both diabetes and CVD are especially desirable. The targeting of 11β-hydroxysteroid dehydrogenase type 1 (11β-HSD1), GPR119, TGR5, sirtuin 1 (SIRT1), the sodium-glucose co-transporter 2 (SGLT2), and GPR40 are examples of such approaches and rationale of each of these is discussed in this chapter.

In the third chapter, the QSAR studies on the imidazolopyrimidine amides, the (2*S*)-cyanopyrrolidine analogues and the derivatives of β-aminoamide bearing subsituted triazolopiperazines as dipeptidyl peptidase IV (DPP-4) inhibitors have been discussed. The QSAR rationales for these analogous were obtained in terms of Dragon descriptors using Combinatorial Protocol in Multiple Linear Regression (CP-MLR). The DPP-4 inhibitory actions of these compounds have appeared as the cumulative influence of different structural features which were recognized in terms of individual descriptors.

The quantitative analysis carried out on the PPARγ transactivation profiles of the derivatives of tetrahydroquinolines, benzylpyrazole acylsulfonamides and pyridyloxybenzene-acylsulfonamides in terms of topological 0D- to 2D-descriptors have been discussed in the fourth chapter. These analyses have provided a rational approach for the development of titled derivatives as PPARγ agonists.

The fifth chapter represents QSAR rationales for the GPR119 agonistic activity of indole-based derivatives and triazolopyridines. The derived QSAR models have provided a rational approach for the development of these derivatives as GPR119 agonists.

ACKNOWLEDGEMENTS

I express my deep sense of gratitude to my research supervisor Dr. Brij Kishore Sharma, Associate Professor, Department of Chemistry, Government College Bundi. I am greatly indebted to him for introducing to this field and providing me the necessary computational facilities throughout the research work. The enlightening discussions with him widened my imagination and give me a deeper insight of the field.

I am very much grateful to all the faculty members of Chemistry Department and Principal Govt. College Bundi for their help and providing necessary facilities to complete this work. I also sincerely thank to Dr. Pradeep Pilania and Dr. Manju Chaudhary, Department of Chemistry, MLV College Bhilwara and Afsar Jahan for their constant encouragement in this journey. I am grately thankful to Dr. Ghasi Ram Choudhary for his moral support and inspiration.

My special thanks go to Dr. R. M. Quereshi, Former Principal and Dr. J. K. Vijayavargiya, Present Principal, Govt. College Kota for their constant help and encouragement. I pay my sincere thanks to Dr. Sanjay Bhargava, Present Asistant Director, College Education Kota Zone and Associate Professor of Physics, Govt. College Kota and Dr. K M Gavendra, Former Asistant Director, College Education for their wonderful support and help and also to Dr. R. N. Soni, Head of Physics Department, Govt. College Kota for his valuable suggestions, discussions and encouragement. Dr. B. K. Yogi, Head of Physics Department Govt. College Baran deserves for special thanks for his support, ever ready help and inspiration.

I and greatly thankful to Drs. Kalpana and Monika Dakshene, the former and present Head of Chemistry Department respectively, Govt. College Kota for their stimulating guidance and inspiration. My friends and colleagues Dr. S. M. Nafees and Dr. Naveen Mittal deserve for special thanks for their help at various stages of my experimental work and thesis preparation. I also pay my sincere thanks to my Departmental collegues, Dr. Uttra Chandrawat, Dr. Manju Bala Yadav, Dr. Suman Sharma, Dr. Renuka Jain, Dr. Seema Agrawal, Dr. Nitin

Gupta, Dr. Suresh Meena, Dr. Dhanraj and all other faculty members for useful discussions and support. I also extend my special thanks to Dr. Siyaram Meena for his valuable help during my thesis preparation. I am also thankful to Laboratory staff of Department of Chemistry Govt. College Kota for their timely help. I extend my special thanks to Dr. Arun Kumar and Dr. Manish Gaur Department of Mathematics, Dr. Anju Kapoor, Head of Botany Department and Dr. Irshad Ahmed Head of Zoology Department, Govt. College Kota for their help and support.

This part would be incomplete, if I miss to pay my sincere regards to Prof. Ashu Rani, Department of Chemistry, University of Kota for her esteemed guidance and encouragement.

I express my profound regards to my parents for their constant inspiration for academics throughout my carrier. Finally, I extend my deepest affection to my wife Mrs. Shashikala Meena, my son Manmohan and daughter Latika, whose patience beyond tolerance, made it possible to devote maximum time. It would be difficult for me to put their salient contribution.

(Raghuraj Parihar)
Research Scholar
Department of Chemistry
Government College Bundi

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

11β-HSD1	11β-HydroxySteroid Dehydrogenase type 1
3D-MoRSE	3D-Molecule Representation of Structures based on Electron diffraction
AACE	American Association of Clinical Endocrinologists
ACCORD	Action to Control Cardiovascular Risk in Diabetes
ACE	American College of Endocrinology
ACE	Angiotensin Converting Enzyme
AD	Applicability domain
ADA	American Diabetes Association
ADVANCE	Action in Diabetes and Vascular Disease: Preterax and Diamicron MR Controlled Evaluation
AIC	Akaike's Information Criterion
ARB	Angiotensin II Receptor Blocker (ARB)
ASCVD	Atherosclerotic Cardiovascular Disease
Atg	Autophagy-Related Gene
ATS	Autocorrelation of a Topological Structure
BA	Biological Activity
BA	Bile Acid
BMAL	Brain and Muscle Aryl hydrocarbon receptor nuclear translocator-Like
CA	Cholic Acid
cAMP	cyclic AMP
CARMELINA	Cardiovascular safety and Clinical Outcome with Linagliptin
CAROLINA	CARdiovascular Outcome Trial of LINAgliptin
CDCA	ChenoDeoxyCholic Acid
CHO	Chinese Hemester Ovary
CKD	Chronic Kidney Disease
CP-MLR	Combinatorial Protocol in Multiple Linear Regression
CR	Calorie Restriction
CVD	CardioVascular Disease
DCA	DeoxyCholic Acid
DEPICT-1	Dapagliflozin Evaluation in Patients With Inadequately Controlled Type 1 Diabetes
DKA	Diabetic KetoAcidosis
DM	Diabetes Mellitus
DPP	DiPeptidyl Peptidase
EASD	European Association for the Study of Diabetes
EXAMINE	The Examination of Cardiovascular Outcomes With Alogliptin Versus Standard of Care in patients with type 2 diabetes mellitus

	and acute coronary syndrome
FAO	Fatty Acid Oxidation
FBG	Fasting Blood Glucose
FFA	Free Fatty Acid
FIT	Kubinyi Function
FOXO	Forkhead boX O
FXR	Farnesoid X Receptor
G6P	Glucose-6-Phosphate
GATS	Geary Autocorrelation of a Topological Structure
GC	GlucoCorticoid
GDIS	Glucose Dependent Insulin Secretion
GDM	Gestational Diabetes Mellitus
GETAWAY	GEometry, Topology, and Atom-Weights AssemblY
GFR	Glomerular Filtration Rate
GIP	Glucose-Dependent Insulinotropic Peptide
GK	GlucoKinase
GKA	GlucoKinase Activator
GKRP	GlucoKinase Regulatory Protein
GLP-1	Glucagon-Like Peptide-1
GPCR	G Protein-Coupled Receptor
GPR119	G Protein-coupled Receptor 119
GPR40	G Protein-coupled Receptor 40
GR	Glucocorticoid Receptor
HF	Heart Failure
HFD	High Fat Diet
HIF	Hypoxia Inducible Factor
IDDM	Insulin Dependent Diabetes Mellitus
IDF	International Diabetes Federation
IKK	IkB Kinase
IL-6	Interleukin-6
IR	Insulin Receptor
IRS-1	Insulin Receptor Substrate-1
JAK2	Janus-Activated Kinase 2
JNK	c-Jun N-terminal Kinase
KATP	ATP-Dependent Potassium Channel
LC3	Light Chain 3
LCA	LithoCholic Acid
LGO	Leave-Group-Out
LKB	Liver Kinase B
LOF	Friedman's Lack of Fit
LOO	Leave-One-Out

LPC	LysoPhosphatidylCholine
LXR	Liver X Receptor
MATS	Moran Autocorrelation of a Topological Structure
MIM	Molecular Influence Matrix
MODY2	Maturity-Onset Diabetes of the Young type 2
MR	Molar Refractivity
MRA	Multiple Regression Analysis
MWC	Molecular Walk Count
NAD	Nicotinamide Adenine Dinucleotide
NAFLD	Non-Alcoholic Fatty Liver Disease
NASH	Non-Alcoholic Steato-Hepatitis
NF-κB	Nuclear Factor-κB
NHE3	Na^+/H^+-Exchanger 3
NIDDM	Noninsulin Dependent Diabetes Mellitus
NOD	Non-Obese Diabetic
OBA	2-Oxalylamino Benzoic Acid
OEA	OleoylEthanolAmide
OLDA	OLeoyl DopAmine
OXPHOS	OXidative PHOSphorylation
p38 MAPK	p38 Mitogen Activated Protein Kinase
PARP	Poly-ADP-Ribose Polymerase
PER2	Period 2
PGC	Peroxisome proliferator activated receptor-g Coactivator
PI3K	PhosphoInositide 3-Kinase
PKA	Protein Kinase A
PLS	Partial Least Squares
PPAR	Peroxisome Proliferator Activating Receptor
PPRE	Peroxisome Proliferator Response Elements
PRESS	Predictive Residual Sum Of Squares
PTP1B	Protein Tyrosine Phosphatase 1B
PTP	Protein Tyrosine Phosphatases
PYY	Polypeptide YY
QSAR	Quantitative structure-activity relationship
RDF	Radial Distribution Function
ROS	Reactive Oxygen Species
RXR	Retinoid X Receptor
SAR	Structure-Activity Relationship
Savor-TIMI-53	The Saxagliptin Assessment of Vascular Outcomes Recorded in Patients with Diabetes Mellitus–Thrombolysis in Myocardial Infarction 53
SGLT2	Sodium-GLucose coTransporter 2

Sir 2	Silent information regulator 2
SIRT1	Sirtuin 1
SREBP	Sterol Regulatory Element Binding Protein
SRW	Self-Returning Walk
T1D	Type 1 Diabetes
T2D	Type 2 Diabetes
TECOS	The Trial Evaluating Cardiovascular Outcomes With Sitagliptin
TGR5	Takeda G protein Receptor 5
TLCA	Tauro-LithoCholic Acid
TNF-α	Tumor Necrosis Factor-α
Treg	Regulatory T Cell
TRPV1	Transient Receptor Potential Channel
TWC	Total Walk Count
TYK2	TYrosine Kinase 2
UCP	UnCoupling Protein
UDCA	UrsoDeoxyCholic Acid
VIF	Variance Inflation Factor
WHIM	Weighted Holistic Invariant Molecular

Chapter 1
QSAR: Methods and Parameters

CHAPTER 1

QSAR: METHODS AND PARAMETERS

1. INTRODUCTION

From the beginning of civilization or may be long before, human brain is always in search of treatment of ailments. Initially man used herbs as medicines, but then in the mid-nineteenth century serious efforts were made to isolate and purify the active principles of these remedies. Since then a large variety of biological active compounds have been obtained and their structures were determined. Usually this is achieved by molecular modifications using the trial-and-error approach, not by proper information. Though the properties indicating a certain molecule as a drug candidate were known, it was not really feasible to investigate a large number of molecules for such types of properties. Of course, the nature of these properties would be represented by structural features of a molecule and thus examination of certain motifs provided a direction for experimental investigation. The problem with this approach is that it does not always lead to an understanding of why a molecule behaves as a drug against its target or why it does not so. This procedure, with a demand of intensive long journey, is very expensive and yet less efficient. The reasons for this apparently low success are poor pharmacokinetics, toxicity, side effects and lack of efficacy. Thus, it is desirable that only those molecules with a good chance of activity should be prepared and tested.

To this end, the approach based on "rational drug-design" has been developed in which a new active compound is forecasted by modifying a carefully chosen "lead" structure through molecular manipulation. Thus, appropriate theoretical methods may be employed to have new potential compounds prior to their synthesis. First of all, Brown and Fraser [1] laid the foundation of an idea about the relationship between physiological or biological action of a compound and its chemical constitution. Presently, it is well established that the biological activity (*BA*) is a function of physicochemical (physical, structural, and chemical) properties i.e. chemical constitution C, which was later formulated in a mathematical expression of the type shown in Equation (1.1)

$$BA = f(\mathrm{C}) \qquad (\mathbf{1.1})$$

With this concept, structure-activity relationships established the novel strategies of chemical synthesis by inserting new chemical groups into the biomedical compound and test the modifications pertaining to their biological effects. This enables the identification and determination of the chemical groups responsible for inducing a targeted biological effect in the organism. That in turn resulted into new methods of finding, examining and interpreting the SAR, on a quantitative basis in a more systematic manner. This strategy has, therefore, been termed as quantitative structure-activity relationship (QSAR).

At present, a QSAR study has two important objectives one is the diagnostic and the other is predictive. The former, deals with the mechanistic aspect supporting or suggesting theories for the site of action. The later is concerned with the extrapolative and interpolative predictions based on the correlative approach. The interpolative prediction within spanned substituent species (SSS) is thought to be much more reliable as compared to extrapolative prediction outside SSS. In this way, the QSAR study has become an obligatory tool to rationalize the design of new bioactive compounds and to investigate their interaction with the living matter.

2. METHODOLOGY

The prime goals of a correlative approach are (i) to search for the relevant parameters (descriptors) which can account for the variation in the observed biological activities of compounds in a congeneric series, and (ii) to determine the extent of correlation between a set of descriptors and the activity under investigation. Statistical and physical models have been developed for this purpose. However, the choice of a method relies profoundly on various factors such as the quality of biological data, the number of compounds to be tested, the degree of variance in results, and the ratio of time needed in synthesizing and testing the compounds for their biological activity.

To date a large matrix of QSAR methodologies have been developed to assist a medicinal chemist in his esteemed goal to have more potent compounds with lesser dependence on “trial-and-error” synthesis and testing. These are summarized as:

(i) Manual methods: The Craig plots [2], the Topliss schemes [3, 4], the Simplex

search [5] and the Fibonacci search [6]

(ii) Regression analysis methods: The Hansch analysis [7-12], the bilinear model of Kubinyi [13-16] and the Free-Wilson or Fujita-Ban approach [17-18]

(iii) Pattern recognition methods: the learning machines [19], the K-nearest neighbor analysis [20-21] and the discriminate analysis [22] and

(iv) Some other methods such as the probabilistic analysis [23], the factor analysis [24-26] and the cluster analysis [27].

Different correlative approaches based on parametric (2- and 3-dimensional) and non-parametric strategies and methods to compute molecular descriptors which are able to quantify the biological actions of drug molecules are given in Figure 1.1.

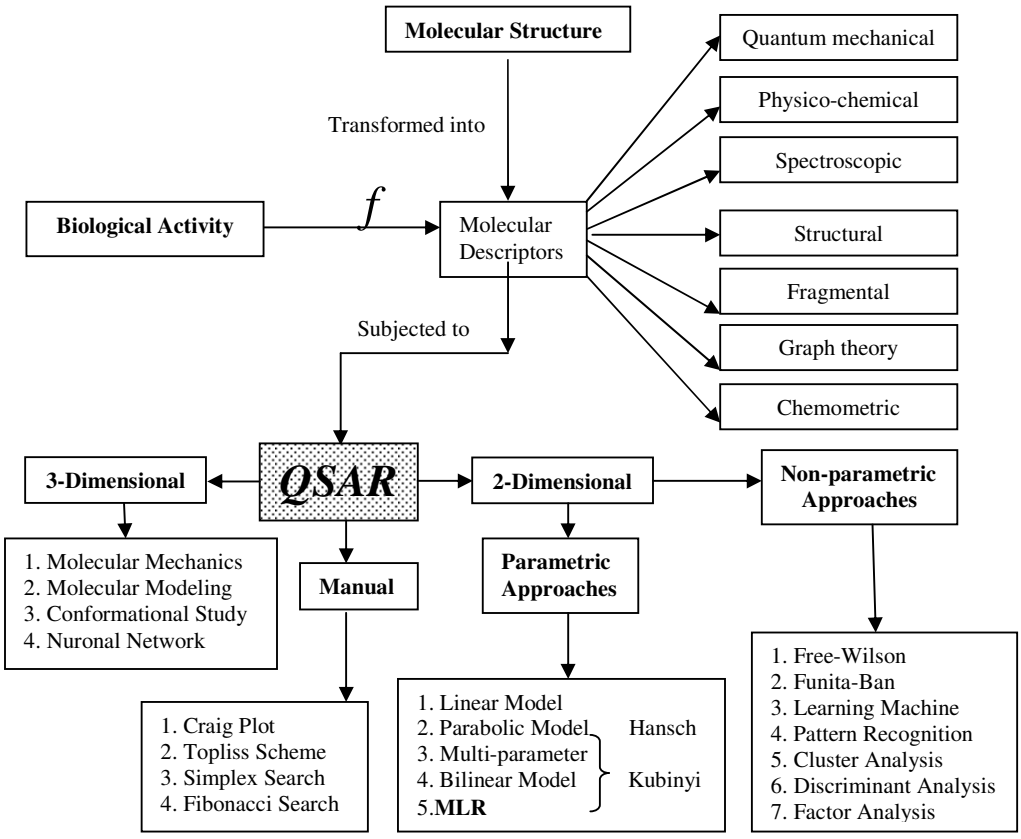


Figure 1.1: Common methodologies used for QSAR analyses and computation of molecular descriptors.

For the present work, the Hansch analysis model has been employed in obtaining the relevant QSARs for different congeneric series under study. In this

approach, the linear multiple regression analysis (MRA) [28-29], employing the method of least squares, is used to find the "best fit" of a dependent variable, the biological activity to a linear combination of independent variables (descriptors). This is commonly expressed through a model Equation (1.2)

$$y=a_0+a_1x_1+a_2x_2+---+a_n x_n \qquad \textbf{(1.2)}$$

In this equation x_1,x_2------- etc. are the descriptors related to the contribution of substituents to the activity in a congeneric series; the dependent variable, *y,* is related to the biological response of a compound in the series, a_1, a_2, ----- etc., are the weights of descriptors, and a_0 is the intercept. The magnitude of coefficients provides the information about the contribution of variables to the activity. Both the Hansch and the Free-Wilson/Fujita-Ban models are briefly described in the following sections.

2.1. The Hansch Approach

The Hansch approach [7-9] is based on the Hammett's Linear Free-Energy Related (LFER) model. Amongst the various QSAR approaches it is the most widely and effectively used method. In this approach, the biological response is considered as a function of certain physicochemical and/or structural and/or topological properties of a molecule. Mathematically, the same is expressed through Equation (1.3)

$$BA = a + b\pi(\text{or } \log P) + c\pi^2(\text{or } \log P^2) + d\sigma + eEs + fS \qquad \textbf{(1.3)}$$

where π (or log*P*), σ and *Es* are, respectively, the hydrophobic, electronic and steric parameters, and *S* is a structural parameter defining the shape, size or the topography of a molecule. The numerals a, b, c, d, e and f are the regression coefficients associated to independent parameters. The biological activity (*BA*) is measured as negative logarithm of a standard biological response such as IC_{50}, EC_{50}, LD_{50}, K_i etc., on the molar basis. Equation (1.3) expresses a parabolic (rather than linear) dependence of activity on the hydrophobic character of molecules. Most often, the relationship between activity and lipophilicity, π or log*P*, was found to be strictly linear, and in such cases, the square of lipophilicity term in the above equation is dropped.

For the non-linear QSARs, the semi-empirical bilinear model of Kubinyi is a more flexible version of Equation (1.3) that allows for an optimum log*P* (or π)

but provides linear ascending and descending portion of the curve. The same is expressed through Equations (1.4) and (1.5)

$$BA = a\log P - b\log(\beta P + 1) + c \quad (\mathbf{1.4})$$

$$BA = a\pi - b\log(10^{\pi\beta} + 1) + c \quad (\mathbf{1.5})$$

Where all the disposal parameters, a, b, c and β are evaluated using an iterative least square procedure.

2.2. The Free Wilson/Fujita-Ban Approach

The Free-Wilson method [17] is based upon the assumption that the effect of a substituent on *BA* is additive and independent of the presence or absence of substituents at other positions. The same can quantitatively be expressed by Equation (1.6)

$$BA = \sum A_{ij} S_{ij} + \mu_0 \quad (\mathbf{1.6})$$

Here A_{ij} is the activity contribution of substituent i at position j and S_{ij} takes the value of 1 or 0 depending upon the presence or absence of substituent i at position j. The constant μ_0 represents the overall average activity of the series. Solution of the set of simultaneous Equations (1.6), one for each compound, using MRA in a least square manner, gives the "best fit" parameter values of all A_{ij} and μ_0.

Fujita and Ban [18] suggested two modifications in the original formulation. First, the biological activity should be expressed as $-\log C$ or an equivalent measure proportional to a free-energy change so that derived substituent constants might be compared with other free-energy related parameters, and second, that μ_0, the overall average, become analogous to an intercept, that is, the calculated activity of the unsubstituted (or the parent) compound of the series. This obviates the need for the cumbersome restricted equations of the original method.

2.3. Validation Statistics

A number of statistics, for *n* data points, are also derived in conjunction with such calculations, which allows the statistical significance of the resulting correlation to be evaluated. The most important of these are, the standard error of estimates, *s*, which should be minimized, the multiple correlation coefficient squared, R^2, accounting for the variance in observed activities, the explained variance, *EV,* being an adjusted coefficient of determination, measures the fraction

of variance between observed and calculated activities, and it should also be maximized for goodness of fit. In addition, the *F*-value, a statistic for assessing overall significance of the derived equation (statistics Tables list critical values for the appropriate number of degree of freedom and confidence level), the *t*-values (also compared with statistical Tables), and the confidence intervals (usually 95% or 90%) for the individual regression coefficient in the equation are also required to justify the equation statistically. Also very important, in multi-parameter equations, are the cross correlation coefficients, *r* amongst the descriptor variables of the MRA equation. To assure the true "independence" or "orthogonality" of the variables, these must be low. It is a necessary condition for meaningful results.

The derived regression equation should also be subjected to the validation test by the leave-one-out (LOO) [30] and leave-group-out (LGO) methods. In LOO procedure a compound is removed from the parent data set to generate modified data sets, in the form of reduced data sets, in such a manner that each one is excluded once only. Predictions of the response values for the excluded compounds are made on the basis of the developed model. The predictive residual sum of squares (PRESS) is obtained by the addition of the squared difference between predicted and actual values. The formula (1– PRESS/SSY) is used to calculate the Q^2_{LOO}. The variance of the observed responses of data points around the mean value is denoted by SSY. A value greater than 0.5 of this cross-validated Q^2-index, hints out that the model obtained is a reasonable QSAR model. If the value of Q^2-index is greater than 0.9 the model is an excellent statistical model.

Sometimes, the derived statistical equations relative to various methodologies may also be compared using the same data set. The comparisons may be made based on the results of the regression analysis, the predictive capability of the equations generated. The other statistical criteria are *AIC* [31-32], *FIT* [33-34] and *LOF* [35-36]. The Akaike's information criterion *AIC*, that correspond to the total variables, p', is given by Equation (1.7)

$$AIC = \frac{RSS\,.\,(n + p\prime)}{(n - p\prime)^2} \qquad (\mathbf{1.7})$$

Those equations having the minimum *AIC* value are considered as the most useful in statistical sense. The Kubinyi function *FIT*, bearing a resemblance to the *F*-value is given for *k* independent variables by Equation (1.8)

$$FIT = \frac{r^2.\ (n-k-1)}{(n+k^2).(1-r^2)} \quad \textbf{(1.8)}$$

The quality of the derived regression equations may be assessed by this parameter. The highest value of this parameter is representative of a statistical sound model. The number of independent variables (k) decides the sensitivity of the F-value. It is more sensitive when k is small and less sensitive when k is large. On the other hand, *FIT* criterion is less sensitive with small numbered changes in k values and there is substantial increase in sensitivity as the values of k increases. The Friedman's lack of fit, *LOF* parameter is given by Equation (1.9)

$$LOF = \frac{RSS/n}{\left[1 - \frac{k\,(d+1)}{n}\right]^2} \quad \textbf{(1.9)}$$

In this equation *RSS* denotes the summation of the squared differences between the observed and predicted activities. The smoothing parameter d is accountable for the number of terms employed in the equation. This parameter is unbiased regarding large numbers of parameters.

The derived statistical equations were also subjected to external validation test in which a few compounds (nearly 20-30% of the total population), following certain criterion, were removed. The removed compounds then form a test set while remaining compounds considered together represent the training set. The derived regression equation from the training set was further used to predict the activity values of the compounds in test set. The close resemblance of such predicted activity values to that of observed ones validates the model externally. The derived statistical parameter, r^2_{test} [37], may therefore, ascertain the predictive power or external consistency of a generated model. The mathematical expression for this parameter is given below:

$$r^2_{\text{Test}} = 1 - \frac{\sum(Y_{\text{Pred(Test)}} - Y_{\text{(Test)}})^2}{\sum(Y_{\text{(Test)}} - \bar{Y}_{\text{(Training)}})^2} \quad \textbf{(1.10)}$$

where, $Y_{\text{Pred(Test)}}$ are the predicted activities of the test-set compounds and $Y_{\text{(Test)}}$ are observed activity values of the test-set compounds. The mean activity value of the training-set is represented by $\bar{Y}_{\text{(Training)}}$. For a predictive QSAR model, the value of predicted r^2_{Test} should be more than 0.5.

The parameter r^2_{test} may not truly reflect the predictive capability of the model on a new dataset because it is mainly controlled by sum of squared

differences between observed values of test set compounds and mean observed activity values of training data set, i.e., a good value of squared correlation coefficient (r^2) between observed and predicted values of the test set compounds does not necessarily mean that the predicted values are very near to corresponding observed activity.

Thus, two novel parameters, r_p^2 and r_m^2 have been introduced further to account for the acceptability of a predictive QSAR model [38].The modified r^2, [$r_{m(Test)}^2$] is given by the Equation (1.11)

$$r^2_{m(Test)} = r^2 \times \left(1 - \sqrt{r^2 + r_o^2}\right) \quad \textbf{(1.11)}$$

where r_0 is squared correlation coefficient between the observed and predicted values of the test set compounds with intercept set to zero. The value of this parameter should be greater than 0.5 for an acceptable model. Initially, the concept of r_m^2 was applied only to the test set prediction [39], but it can as well be applied for training set if one considers the correlation between observed and LOO predicted values of the training set compounds [40]. More interestingly, this can be used for the whole set considering LOO-predicted values for the training set and predicted values of the test set compounds. The r_m^2(overall) statistic may be used for selection of the best predictive models from among comparable models. Another modified parameter is R_p^2 [41], which penalizes the model R^2 for the difference between squared mean correlation coefficient (R_r^2) of randomized models and squared correlation coefficient (R^2) of the non-randomized model. This parameter may be calculated using following Equation

$$R^2_P = R^2 \times \left(\sqrt{R^2 + R_r^2}\right) \quad \textbf{(1.12)}$$

This parameter ensures that the models thus developed are not obtained by chance and the value greater than 0.5 accounts for an acceptable model.

Finally, a randomization test [42, 43] was performed to observe any chance correlations coupled to the derived models. The test involves the repeated randomization of the biological actions of compounds for each cross-validated model. The multiple regression analysis was made to reassess the randomized response vector of the datasets, in 100 simulation runs. The counts of the

consequential regression equations having correlation coefficients better than or equal to the unscrambled activity data was used as a measure of the percent chance correlation of the derived model.

3. QSAR PARAMETERS

The most essential feature of a QSAR study requires the representations of a substituent/molecule quantitatively in such a manner that it can explain the activity induced in a biological system. Generally, such a representation can be of physicochemical, theoretical, and structural in nature. The physicochemical model of the biological activity that the activity of a compound is a function of three separable factors: the electronic, the steric and the solvent partitioning or the hydrophobic.

The theoretical and/or structural parameter (s) are being used either alone or in conjunction or in lieu of physicochemical parameters to quantify the molecular features of the series. The overall situation is expressed by Equation (1.13)

$$BA = \text{f(electronic)+f(steric)+f(hydrophobic)+ [f(structural)+f(theoretical)]} \quad (\mathbf{1.13})$$

There are three major sources of QSAR parameters i) experimenta1 measurements ii) theoretical estimations and iii) extrapolation from the data base. The discussion and the relevant equations defining the different parameters used in the Hansch approach are further defined in the sections below.

Some of the well-established parameters, being frequently used in QSAR studies, are summarized in Table 1.1.

Table 1.1: Some important QSAR parameters

Substituent parameter/constant	Symbol	Reference
Hydrophobic parameters		
Partition coefficient	log$P$	44-45
Hydrophobic constant	π	44-49
Hydrophobic constant from liquid-liquid chromatography	R_M	50
Elution time in high-pressure liquid chromatography	k'	51
Solubility	δ	52
Electronic parameters		
Ionization constant	pK_a	53

Hammett constants	σ (σ_m, σ_p)	54-56
Distribution constants	D	57
Resonance and Field constants	R, F	58
Steric parameters		
Taft steric parameter	Es	59
Molar refractivity	MR	48
Molar volume	MV	60
Molecular weight	MW	
Theoretical parameters		
van der Waals volume	V_w	61-64
Kier's molecular connectivity indices	${}^m\chi^v$	65-68
Molecular orbital indices		
Charge on an atom	Q (Q_σ, Q_π)	69-70
Nucleophilic & Electrophilic superdelocalizability	S_r^N, S_r^E	71
Electrophilic & Nucleophilic frontier orbital densities	f_r^N, f_r^E	72
Energy of the lowest unoccupied molecular orbital	E_{LUMO}	69-70
Energy of the highest occupied molecular orbital	E_{HOMO}	69-70

3.1. Electronic Parameters

The first important definition of an electronic parameter, σ, came in 1935 when L P Hammett [54] proposed his famous Equation (1.14)

$$\delta\sigma = \log K_X - \log K_H \qquad (\mathbf{1.14})$$

where K_H *and* K_X are the ionization constants for unsubstituted benzoic acid and its ortho/meta/para derivatives in water at 25°C. δ is the reaction constant, which measures the susceptibility of a reaction to polar effects and depends on the nature of the reaction including conditions such as reaction medium and temperature. σ called the Hammett substitution constant, measures the effect of the substituent on K. Positive values of σ represent electron withdrawal by the subsituent from the aromatic ring while negative values of σ indicates electron release to the ring. For benzoic acids, by definition, σ =1.000 when the measurements are made in water in 25°C. Thus σ_p and σ_m values can be experimentally determined using pK_a values for meta/para-substituted benzoic acid in water at 25°C. In fact, the Hammett equation states that the electronic effects of the substituents on the ionization of benzoic acids can be used as a model for the effects of substituents on other reaction centers attached to the aromatic systems. In addition, this parameter may also represent effect on the active receptor site by hydrogen bonding and charge-charge or charge-dipole interactions. Later on, Swain and Lupton [58] proposed

the separation of electronic substituent effect into a field constant, *F,* and a resonance constant, *R,* as

$$\sigma = fF + rR \quad (\mathbf{1.15})$$

where *f* and *r* are, respectively, the field and the resonance weighting coefficients.

3.2. Hydrophobic Parameter

Of all the parameters, the one that is most important in the quantification of *BA* is the lipophilicity or hydrophobicity. A multiplicity of published QSARs would attest this fact. Analogues to the Hammett constant, Hansch school also made a significant contribution in estimating the hydrophobic substitution constant [44-48], π, from measured partition coefficients, log*P* data, using the relation

$$\pi_X = \log P_X - \log P_H \quad (\mathbf{1.16})$$

In this equation, P_X is the partition coefficient of a derivative and P_H that of the parent compound. A positive value of π implies that relative to H the substituent favors the octanol phase and is more hydrophobic in nature while a negative π value indicates its hydrophilic character relative to H. Compounds bearing hydrophobic substituents are supposed to be more prone to polar interaction with the receptor. It has also been showed that 1log*P* or π is, to a first approximation, an additive property, and also has considerable constitutive character [45].

3.3. Steric Parameter

The first successful numerical definition of steric effect was given by Taft [59]. The classic Taft steric constant, *Es*, is defined by the Equation (1.17)

$$Es = \log (K_X/K_H) \quad (\mathbf{1.17})$$

Where, K_X and K_H are the acid-catalyzed hydrolysis constants of unsubstitued and α-substituted acetates (XCH_2COOR) respectively. *Es*(H) = 0.00 is normally standardized to hydrogen. The more positive value of *Es* indicates the greater steric effect affecting intramolecular and intermolecular hindrance to reaction under observation.

3.4. Molar refractivity

For many substituents, the experimental *Es* values may not be available. Thus, an alternative theoretical parameter, the molar refractivity [48], *MR*, is more commonly used in QSAR study. It is obtained by the Lorentz-Lorentz Equation

which is expressed as

$$MR = (n^2-1)/(n^2+2).\ (MW/d) \qquad (\mathbf{1.18})$$

In above equation, n is the index of refraction at the sodium D line, *d*, the density, and *MW*, the molecular weight of the compound. Since *MR* is an additive constitutive property of molecules, fragment values for the atoms/structural moieties have been calculated and are used to obtain the substitutent's *MR* values for QSAR studies. Further, the *MR* having the units of molar volume reflects a linear relationship with molecular "bulk". However, this parameter also takes care of electronic effect and gives hints towards the dipole-dipole interaction at active site of the receptor [62].

3.5. van der Waals volume

Another important substituent parameter, accounting for the size or bulk of a molecule or a substituent, is the van der Waals volume, V_w [61]. It is an easily computable theoretical parameter which may be calculated for the whole molecule or the varying substituents on it. The atoms in a substituent (or in a molecule) are assumed to have spherical shapes, as suggested by Bondi [62], to calculate their total volume. The necessary corrections for the overlap of atomic orbitals and for the branching in hydrocarbon chain are also incorporated [61] to obtain V_w. In a number of cases, the parameters, π or log*P* and *Es* were shown to be linearly correlated with V_w [62-64] which indicates that many times V_w also expresses the hydrophobic and steric effects of substituents. In addition, V_w can also take care of dispersion interaction or polarizability [61] of the molecules.

3.6. Molecular connectivity

Molecular connectivity [65-68, 74], χ, simply signifies the degree of branching or connectivity in a molecule. It is derived from the numerical extent of branching or connectivity in the molecular skeleton. The connectivity index has several versions. However, the first-order simple connectivity index ($^1\chi$) and the first-order valence connectivity index ($^1\chi^v$) are extensively employed in QSAR studies. Both are calculated from the hydrogen suppressed graph of the molecule and are defined as

$$^1\chi=\Sigma(\delta_i\delta_j)^{-1/2} \qquad (\mathbf{1.19})$$

$$^1\chi^v =\Sigma(\delta^v_i\delta^v_j)^{-1/2} \qquad (\mathbf{1.20})$$

where δ is the number of atoms connected to any atom in the graph, and the sum is carried over all the connections or edges in such a graph while δ^v is defined as

$$\delta^v_i = (Z_i^v - h_i)/(Z_i - Z_i^v - 1) \quad (\mathbf{1.21})$$

Here Z_i^v is the number of valence electrons, Z_i is the atomic number and h_i is the number of hydrogen atoms attached to i^{th} atom.

Additionally, where an incident arises, some indicator variables specifying the presence or absence (e.g., R or H) or only binary variations (e.g., Me or Ph) of a structural representative are also used. Such variables, allow the mixing of physicochemical and structural parameters in Hansch type of calculations. Numerical values, 1 and 0, are generally used for denoting such structural variations. Sometimes, these binary values (1 and 0) are also used for the discrimination between the hydrogen-bond donor, *HD,* or hydrogen-bond acceptor, *HA,* nature of substituents.

4. DRAGON DESCRIPTORS

The DRAGON software [75] has been employed for the parameterization of the compounds. This software is able to compute a large number of descriptors from different perspectives corresponding to empirical, constitutional and topological characteristics of the compounds or their structural fragments under multi-descriptor class environment. The first requirement of this software is the structures of the compounds. The structures of compounds have drawn in ChemDraw [76] and then energy was minimized using the standard procedure. Dragon software was used to compute the parameters for these energy minimized structures. This software offers a large number of descriptors belonging to 0D-, 1D-, 2D- and 3D-descriptor classes and these descriptors have been divided into 20 logical blocks, discussed below.

Constitutional descriptors

Constitutional descriptors (0D) are the simplest and most commonly used descriptors. These descriptors provide the information about the compound's molecular composition without considering its molecular geometry [77].

Topological descriptors

Topological descriptors (1D) are based on the graphical molecular representation. Algebraic operators when applied on matrices, which represent H-

depleted molecular graphs, with values independent of vertex numbering result into molecular topology. Thus these descriptors represent the numerically quantified molecular topology and able to sensitively address molecular structural features not only size, shape, symmetry, branching and cyclicity but also about the atom type and bond multiplicity in a molecule [78-106].

Walk and path counts

These descriptors are obtainable from a H-depleted molecular graph. The sequence of pairwise adjacent edges, in a molecular graph that leads from one vertex to another one is known as a walk. The number of edges traversed by the walk is termed as walk length. An edge has the liberty to traverse many times. These descriptors are mainly of three types:

(i) The molecular walk count order k (MWCk) are the total number of graphical walks of a molecule of k^{th} length;

(ii) The total walk count (TWC) is the total number of graphical walks of any length ranging from 0 to 10 and

(iii) The self-returning walk counts (SRWk) as the name implies self-returning walks are the graphical walks with starting and ending on the same vertex. The k^{th} order SRW is the total number of graphical self-returning walks of length k. [107-115].

Connectivity indices

These are molecular descriptors that encode information about size, branching, cyclization, unsaturation and heteroatom content in a molecule [116].

Information indices

These molecular descriptors provide total information about the molecule. The information is based on the criteria that are used to define the equivalence classes in a molecule. The chemical identity, spatial bonding, molecular topology and symmetry may be the one of criteria of equivalency of atoms in a molecule [117-119].

Edge adjacency indices

These are molecular descriptors are calculated from the edge adjacency matrix of a molecule. This matrix is obtainable from the H-depleted molecular graph having predetermined connectivity between graph edges. It is based on the

number of bonds between non-hydrogen atom pairs, *B*. Therefore, the entries of this square symmetric matrix of dimension $B \times B$, are one or zero as the bonds considered are adjacent or otherwise, respectively [120-123].

2D autocorrelations

2D autocorrelations are the spatial autocorrelations calculated on a H-depleted molecular graph weighted by atom physico-chemical properties (i.e. the atom weightings *w*) and include autocorrelations ATS (i.e. *Autocorrelation of a Topological Structure*) proposed by Moreau and Broto; autocorrelations MATS calculated by the Moran coefficient and autocorrelations GATS calculated by the Geary coefficient. These are molecular descriptors which describe how a considered property is distributed along a topological molecular structure [124-126].

BCUT descriptors

These molecular descriptors are helpful in searching similarity or diversity in large data bases. The Burden approach is useful to address different facets of molecular structure. These descriptors are based on the extension of this approach significantly. These are calculated from a modified adjacency matrix, the Burden matrix. The elements of this matrix are derived from the H-included molecular graph. The diagonal elements of this matrix are atomic properties and the off-diagonal elements are the square roots of conventional bond order. The off-diagonal elements represent pairs of bonded atoms. The other remaining matrix elements are set at 0.001. DRAGON provides the first 8 highest eigenvalues BEH*wk* and the first 8 lowest eigenvalues BEL*wk* (absolute values) for each matrix, *k* referring to the eigenvalue rank and *w* to the atomic properties, namely, atomic masses (m), atomic van der Waals volumes (v), atomic Sanderson electronegativities (e), and atomic polarizabilities (p) [127, 128].

Topological charge indices

Topological charge descriptors are derived from an unsymmetric matrix *CT*, whose single elements are defined as

$$CT_{ij} = \delta_i \quad \text{if } i=j; \tag{1.22}$$

$$CT_{ij} = m_{ij} - m_{ji} \quad \text{if } i \neq j \tag{1.23}$$

The derivation of these equations is based on the H-depleted molecular graph. The number of connected atoms in this graph represents the vertex degree of the i^{th} atom i.e. δ_i . The result of the multiplication of the adjacency matrix by the reciprocal square distance matrix, are the elements of the matrix, m_{ij}. The topological valence of the atoms is considered as the diagonal entries of the *CT* matrix. The measure of the net charge transferred from the atom *j* to the atom *i* has been represented through the off-diagonal entries CT_{ij} [129].

Eigenvalue-based indices

These molecular descriptors are calculated as a selected eigenvalue or function of the eigenvalues of a square matrix representing a H-depleted molecular graph [85, 130, 131].

Randic molecular profiles

Molecular profiles are sequences of molecular descriptors, proposed by Randic, and derived from the interatomic geometric distances of a molecule. DRAGON provides two molecular profiles. One is much more related to the global molecular 3D structure: *DPk* and the other to the molecular shape: *SPk*. Each descriptor *DPk* in the DP profile is calculated as

$$DPk = \frac{1}{k!}\frac{\sum_{i-1}^{nAT}\sum_{j-1}^{nAT} r_{ij}^{k}}{nAT} \qquad (\mathbf{1.24})$$

where r_{ij} is the geometric distance between atoms *i* and *j*, *nAT* the number of molecule atoms and *k* the descriptor order (*k* = 1, ... , 20). The effect of the factorial normalization factor diminishes at higher values of *k* and *DP* values tend to zero. Each descriptor *SPk* of the shape profile is calculated in the same way as the *DP* descriptors, but taking into account only atoms on molecular periphery (i.e., atoms with H-depleted connectivity equal to 1 or 2). Randic molecular profile is characteristic of a molecule and thus these are particularly suitable to molecule similarity/diversity analysis [132-134].

Geometrical descriptors

Geometrical descriptors are derived from the 3 dimensional structure of the molecule. An optimized molecular geometry is necessary to calculate these descriptors which may be obtained by computational chemistry or crystallography. Since a geometrical representation of a molecule involves the knowledge of the relative positions of the atoms in 3D space, i. e., the (x,y,z)

atomic coordinates of the molecule atoms, geometrical descriptors usually provide more information and discrimination power, also for similar molecular structures and molecule conformations, than topological descriptors. Being calculated on the graph representation of molecules many geometrical descriptors are commonly known as *topographic indices*, instead of using the geometric distances between atoms [77].

RDF descriptors

The recently proposed RDF (*Radial Distribution Function*) descriptors have their origin in radial distribution function. These functions may be treated as the probability distribution function to find out an atom in a spherical volume of radius *R*. The general form of the radial distribution function, calculated at a number of discrete points with predefined intervals, is represented below by RDF*Rw*.

$$\text{RDF}Rw = f.\sum_{i-1}^{nAT-1}\sum_{j-i+1}^{nAT} w_i.w_j\,.\,e^{-\beta(R-r_{ij})^2} \qquad (\mathbf{1.25})$$

where f is a scaling factor (assumed equal to one in the calculations), w is characteristic property of the atoms i and j, r_{ij} is the interatomic distance and nAT is the number of atoms in the molecule. The exponential term contains the interatomic distance r_{ij} and the smoothing parameter β ($Å^{-2}$), which defines the probability distribution of the individual interatomic distance; β can be interpreted as a temperature factor that defines the movement of atoms [135-148].

3D-MoRSE descriptors

The descriptors, 3D-MoRSE (*3D-Molecule Representation of Structures based on Electron diffraction*) are based on the idea of obtaining information from the 3D atomic coordinates by the transformation used in electron diffraction studies for preparing theoretical scattering curves [149].

WHIM descriptors

Weighted Holistic Invariant Molecular (WHIM) descriptors are geometrical descriptors based on statistical indices calculated on the projections of the atoms along principal axes. These are divided mainly into two classes: directional and global WHIM descriptors. All WHIM descriptors are built in such a way as to capture relevant molecular 3D information regarding molecular size,

shape, symmetry and atom distribution with respect to invariant reference frames [150-151].

GETAWAY descriptors

GEometry, Topology, and Atom-Weights AssemblY (GETAWAY) descriptors have recently been proposed as chemical structure descriptors, derived from a new representation, the *Molecular Influence Matrix* (MIM).The molecular information matrix is a symmetric matrix and shows rotational invariance with respect to the molecular coordinates, thus independent of molecule alignment [152, 153].

Functional group counts

These are simple molecular descriptors defined as the number of specific functional groups in a molecule. They are calculated by knowing the molecular composition and atom connectivities as the same functional group belongs to an aliphatic or an aromatic molecular fragment [77].

Atom-centred fragments

These are simple molecular descriptors known as atom-centred fragments and defined as the number of specific atom types in a molecule. They are calculated by knowing the molecular composition and atom connectivities. Each atom type is an atom in the molecule described by its neighboring atoms. Hydrogen and halogen atoms are classified by the hybridization and oxidation state of the carbon atom to which they are bonded; for hydrogen, hetero-atoms attached to a carbon atom in α-position are further considered. Carbon atoms are classified by their hybridization state and depending on whether their neighbors are carbon or hetero-atoms [155].

Charge descriptors

These are electronic descriptors defined in terms of atomic charges and used to describe electronic aspects both of the whole molecule and of particular regions, such as atoms, bonds, molecular fragments etc. [156-158].

Molecular properties

These are 1D-descriptors, representing molecular properties of a molecule such as molar refractivity, fragment based polar surface area and octanol-water partition coefficient [77]. These are calculated for entire structure of a molecule.

The important descriptor classes from above twenty logical blocks have further been discussed in the follow up section: QSAR modeling.

5. QSAR MODELING IN PRESENT WORK

QSAR modeling is a stepwise process involving following main steps:

5.1 Structure entry and energy optimization

5.2 Descriptor calculations

5.3 Feature selection

5.4 Model development

5.5 *Y*-Randomization

5.6 Prediction, validation and interpretation

5.7 Applicability domain

5.1. Structure entry and energy optimization

The molecules to be used in the study can be available as 2D- or 3D-structures. The structures of the compounds under investigation are drawn in 2D ChemDraw [76] using standard procedure. The drawn structures are then converted into 3D modules using the default conversion procedure implemented in the CS Chem 3D Ultra. The generated 3D-structures of the compounds were subsequently subjected to energy minimization in the MOPAC module, using the AM1 procedure for the closed shell systems, implemented in the CS Chem 3D Ultra. This was done to bring all molecules at common minimum energy level and to ensure a well defined conformer relationship across the compounds under study.

5.2. Descriptor calculations

As mentioned previously, the fundamental assumption of QSAR modeling is that molecular structure may be visualized in terms of physical or biological properties. Thus the essential requirement is to have some method which may encode various structural features of a molecule. The encoding of structural features may be achieved through the calculation of molecular descriptors which are obtained as their numerical representations. Such features can range from very simple ones such as the number of carbons or number of halogen atoms etc. to more complex and abstract features. The DRAGON software [75] has been employed for the parameterization of the molecules under investigation. This

software is able to evaluate several hundreds of descriptors from different perspectives corresponding to empirical, constitutional, and topological characteristics of the compounds or their structural fragments under multi-descriptor class environment.

The energy minimized structures of respective compounds have been ported to DRAGON software for computing the parameters corresponding to 0D-, 1D-, 2D- and 3D-descriptor classes. For most of the present work, the descriptors corresponding to 0D-, 1D- and 2D-classes have been computed and used for correlation purposes as the physical interpretation of parameters from these classes is simpler compared to the descriptors of 3D-class. However, 3D-descriptors have also been employed to obtain QSAR rationales.

A brief description about the definition and scope of 0D-, 1D- and 2D-classes in modeling the biological actions of compounds under study is given in Table 1.2.

Table 1.2: Descriptor classes used for modeling the biological actions

S. No.	Descriptor class (acronyms)[a]	Definition and scope
1	Constitutional (**CONST**)	Dimensionless or 0D descriptors; independent from molecular connectivity and conformations
2	Topological (**TOPO**)	2D-descriptor from molecular graphs and independent conformations
3	Molecular walk counts (**MWC**)	2D-descriptors representing self-returning walks counts of different lengths
4	Modified Burden eigenvalues (**BCUT**)	2D-descriptors representing positive and negative eigenvalues of the adjacency matrix, weights the diagonal elements and atoms
5	Galvez topological Charge indices (**GVZ**)	2D-descriptors representing the first 10 eigenvalues of corrected adjacency matrix
6	2D-autocorrelations (**2D-AUTO**)	Molecular descriptors calculated from the molecular graphs by summing the products of

		atom weights of the terminal atoms of all the paths of the considered path length (the *lag*)
7	Functional groups (**FUNC**)	Molecular descriptors based on the counting of the chemical functional groups
8	Atom centered fragments (**ACF**)	Molecular descriptors based on the counting of 120 atom centered fragments, as defined by Ghose-Crippen
9	Empirical (**EMP**)	1D-descriptors represent the counts of non-single bonds, hydrophilic groups and ratio of the number of aromatic bonds and total bonds in an H-depleted molecule
10	Properties (**PROP**)	1D-descriptors representing molecular properties of a molecule

[a]Reference [77].

5.3. Feature selection

The web version of DRAGON software is able to evaluate 1497 molecular descriptors distributed into eighteen classes covering twenty logical blocks, discussed previously. It is apparent that in such a large descriptor pool a number of descriptors will be highly correlated with other descriptors or else may have the same value for all the molecules and will thus contain no relevant information. Thus prior to their use for model development, the original descriptor pool must be reduced in size by selecting only those descriptors which are information rich and relevant. Such descriptors, in subsequent effort, may only be considered for the development of statistical significant models.

5.4. Model development

Once the descriptors have been calculated and reduced the original pool to a more manageable size the next step is to proceed for building a set of models and choose the best one. The recently developed software [27, 159-160], namely the Combinatorial Protocol in Multiple Linear Regression (CP-MLR) analysis has been used successfully to achieve the same. The strategy followed in CP-MLR approach is presented below.

Combinatorial Protocol in Multiple Linear Regression

The CP-MLR is a 'filter' based variable selection procedure for model development in QSAR. The thrust of this procedure is in its embedded 'filters' which are briefly as follows:

Filter-1: seeds the variables by way of limiting inter-parameter correlations to predefined cutoff level (default acceptable value ≤ 0.3);

Filter-2: controls the variables entry to a regression equation through *t*-values of coefficients (default acceptable value ≥ 2.0);

Filter-3: provides comparability of equations with different number of variable in terms of square-root of adjusted multiple correlation coefficient of regression equation, *r*-bar (for 'baseline model' the minimum value is 0.71); and

Filter-4: estimates the external consistency of the equation in terms of cross-validated R^2 or Q^2 with leave-one-out (LOO) cross- validation as default option (default acceptable limits are $0.3 \leq Q^2 \leq 1.0$).

The filter-2 evaluates the significance of variables of each seed in terms of the *t*-values of regression coefficients. It involves a comparison of estimated regression coefficients of the variables and their standard errors; the seed is skipped if the ratio is below the threshold value. Successive additions of variables to multiple regression equation will increase successive multiple correlation coefficient values. In light of this, filter-3 (*r*-bar value) compensates the increment in correlation coefficient due to the increasing number of explanatory variables in seeds and allows the comparison of different seeds.

The flow chart in Figure 1.2 has demonstrated the strategy for the identification of information rich descriptors corresponding to the phenomenon, the biological activity, under investigation.

This has three stages in it: The first stage sorts the descriptor classes into different categories depending on their ability to form any model to explain the variance in the activity. The second stage collates the information rich descriptor classes to select the individual descriptors significant to the activity. The last stage reuses the selected individual descriptors to discover higher models and/or to explain the phenomenon in a comprehensive manner.

In this process, the first stage has been developed based on the philosophy of elimination through selection. This has three iterations in it. It has been devised to address the multiple descriptor classes' environment in high dimensional modeling studies. It operates by way of categorized treatment of descriptor classes. In this the contributing descriptor classes will be identified using simple models called 'baseline models'. Here, a baseline model represents any entry-level cross-validated regression equation with minimum variables (for example one-to three- descriptors) and capable of explaining at least 50% of variance in the dependent variable. This has been considered with the view that among multi-descriptor models, the two- or at the most three-descriptor equations are the simplest to understand and explain the chosen phenomenon. Also, at this stage the level of importance of the descriptors to the phenomenon of study can be seen clearly.

Moreover, the 'baseline model' concept helps in efficiently handling a large number of variables in each descriptor class and in identifying the information rich descriptors of all classes corresponding to the phenomenon. For the identification of the baseline models, the CP-MLR – a 'filter' based variable selection procedure for model development – has been used in its simplest form with predefined filter thresholds as discussed above. In the first iteration the data files corresponding to each individual descriptor class will be evaluated separately for their ability to form a baseline model and accordingly they will be classified as the primary contributors (category I) and the residual descriptor classes (leftover group).

The second iteration is meant for identifying the collective information content of the leftover descriptor classes´ vis-à-vis the activity under study. In this, the leftover descriptor classes of the first iteration have been merged and recycled for their 'collective' influence in evolving the baseline models. Accordingly, the residual descriptor classes have been classified as the 'collective' contributors (category II) and the leftover of second iteration. At the end of second iteration, if no descriptor class is selected under the categories II and I, the process have been terminated to redefine the filters' threshold in CP-MLR for new baseline models to facilitate the capture of the descriptor class.

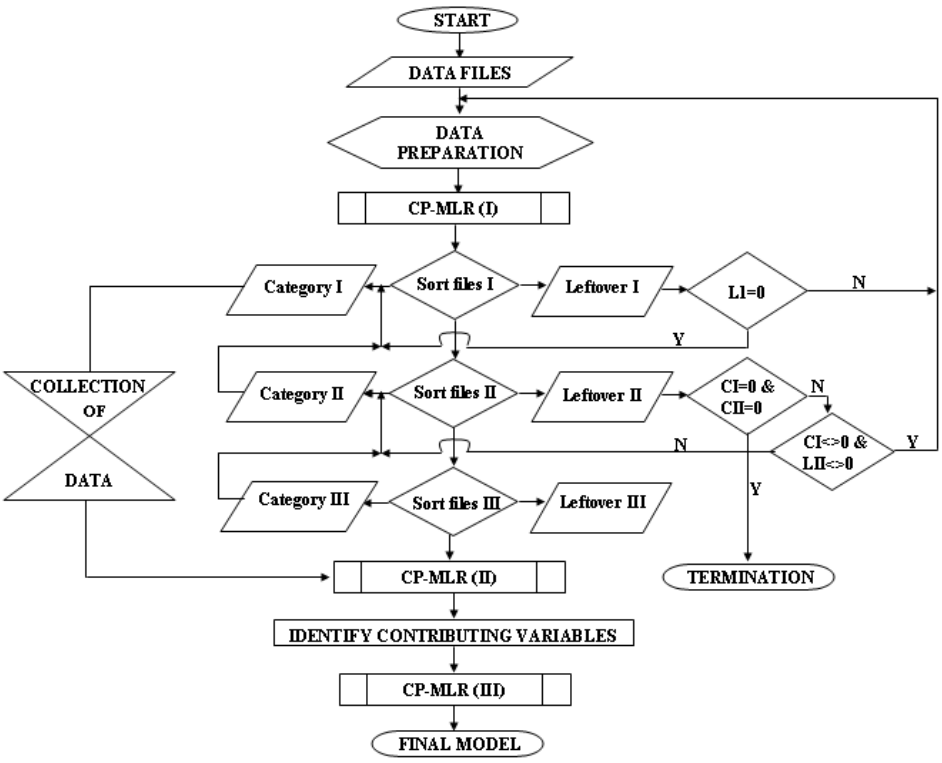


Figure 1.2: Procedure of the model(s) development strategy. It is embedded with Combinatorial Protocol in Multiple Linear Regression (CP-MLR) and shows the progress of selection of descriptors classes into categories I, II and III (CI, CII, CIII) and leftovers I, II and III (LI, LII, LIII) (1^{st} stage), individual descriptors (2^{nd} stage) and final structure-activity models (3^{rd} stage). In this 'Y' stands for 'yes' and 'N' stands for 'no'. In each stage the CP-MLR has been used for distinct function namely categorisation of descriptors classes, sieving contributing descriptors from the identified descriptor categories and finally identifying higher models and descriptors involved therein.

However, if no descriptor class is selected under the category I alone, then the leftovers, if any, of the second iteration were excluded from the study by treating them as noncontributing descriptor classes and the process continued with the second stage of the flowchart. Otherwise, the leftovers of the second iteration have been carried forward to the next generation iteration to examine their possibility of making a 'secondary' contribution in association with the primary descriptor classes (category I). In this way, the third iteration classifies the

corresponding descriptor classes as the 'secondary' contributors (category III) and the non-contributing descriptor classes, which were excluded from the study from this point onward.

On identifying the contributing descriptor classes in the form of categories I, II, and/ or III, the process continues in the second stage with the collated descriptor classes to create all possible baseline models that could possibly emerge from them. These models give out the individual contributing descriptors across the categories. The identified individual descriptors have been recycled in the last stage for higher models and comprehensive diagnosis of the phenomenon.

5.5. *Y*-Randomization

All the models identified in the last stage have been further put to a randomization test [161, 162] by repeated randomization of the activity to discover the chance correlations, if any, associated with them. For this every model has been subjected to 100 simulation runs with scrambled activity. The scrambled activity models with regression statistics better than or equal to that of the original activity model have been counted to express the percent chance correlation of the model under scrutiny.

5.6. Prediction, validation and interpretation

Validation of the derived model is necessary to test its prediction and generalization within the study domain. The data set is randomly divided into training set for model development and test set for external prediction. For each model, besides the statistical parameters *R*, *s* and *F*-ratio, the other indices such as the cross- validated Q^2_{LOO} (leave-one-out) and Q^2_{L5O} (leave-five-out) have also been computed. Additional statistical parameters such as the Akaike´s information criterion, *AIC* [31, 32] the Kubinyi function, *FIT* [33, 34] and the Friedman's lack of fit, *LOF* [35], have also been calculated to further validate the derived models. In case of internal validation, cross validated Q^2_{LOO} and Q^2_{L5O} have been used to ascertain the robustness and predictive ability of the derived model.

5.7. Applicability domain

The usefulness of a model is based on its accurate prediction ability for new congeners. A model is valid only within its training domain and new compounds must be assessed as belonging to the domain before the model is

applied. The applicability domain is the physicochemical, structural or biological information on which training set of the model has been developed and for which it is applicable to make predictions for the new compounds. This domain is evaluated by the leverage values for each compound [163]. A Williams plot [the plot of standardized residuals versus leverage values, h] is constructed which can be used for a simple graphical detection of both the response outliers (*Y* outliers) and structurally influential chemicals (*X* outliers) in the model. In this plot, the applicability domain is established inside a squared area within ± x times the standard deviations and a leverage threshold h^*which is generally fixed at $3(k + 1)/n$ (n is the number of training set compounds and k is the number of model parameters) whereas x = 2 or 3. If the compounds have a high leverage value ($h>h^*$) then the prediction is not trustworthy. On the other hand, when the leverage value of a compound is lower than the threshold value, the probability of accordance between predicted and observed values is as high as that for the training set compounds.

At this point we have in hand a validated model with good predictive ability. The important feature of the model is that it should have incorporated one or more structure activity relationships. The final task of a QSAR modeling methodology is to interpret the model to describe these relationships.

The interpretation of a linear model may also utilizes the PLS technique which dissect the effects of individual descriptors on the dataset and allows a very detailed description of any structure activity relationship captured by the model. A brief description of the PLS technique is provided below.

6. PARTIAL LEAST SQUARE REGRESSION

Partial Least Squares (PLS) regression technique finds a linear regression model by projecting the predicted variables and the observed variables to a new space which is especially useful in quite common case where the number of descriptors (independent variables) is comparable to or greater than the number of compounds (data points). PLS approach leads to stable, correct and highly predictive models even for correlated descriptors instead of the solution of classical least squares problem which does not exist or unstable and unreliable [164-166]. Partial Least Squares regression is based on linear transition from a

large number of original descriptors to a new variable space based on small number of orthogonal factors (latent variables). Latent variables are chosen in such a way as to provide maximum correlation with dependent variable; thus, PLS model contains the smallest necessary number of factors. With increasing number of factors, PLS model converges to ordinary multiple linear regression models. In addition, PLS approach allows one to detect relationship between activity and descriptors even if key descriptors have little contribution to the first few principal components [167-169].

A PLS model will try to find the multidimensional direction in the *X* space that explains the maximum multidimensional variance direction in the *Y* space. Here the *X*- and *Y*-scores are selected so that the relationship between successive pairs of scores is as strong as possible. In principle, this is like a robust form of redundancy analysis, seeking directions in the factor space that are associated with high variation in the responses but biasing them toward directions that are accurately predicted [165].

In principle, the PLS components are extracted from relatively large number of descriptors, the obtained PLS regression models are sensitive to the noise due to the excessive irrelevant descriptors. Thus variable selection procedures have been applied to refine the performance of PLS models. In this procedure the information rich descriptors corresponds to the biological activity are selected by variable selection algorithm, i.e., CP-MLR that can integrate the meaningful variables or to eliminate the redundant variables in final PLS model.

For optimum selection of meaningful variables a new approach Variance Inflation Factor (VIF) can also be used, which is a potent method of detecting the severity of multicollinearity [170-171]. VIF can easily be calculated as:

$$VIF = \frac{1}{1-R^2} \quad (\mathbf{1.26})$$

$$Tolerance = \frac{1}{VIF} \quad (\mathbf{1.27})$$

When VIF value is higher than 5 or tolerance remains under 0.20 then multicollinearity among the descriptors exists. For continuation of process one variable can be selected while others can be left aside from a set of multicollinear variables. The significance of normalized PLS regression coefficients of the descriptors coupled with different statistical measures have also been used to

identify the redundancy in the variables [170]. Thus, the PLS approach is useful in retaining the descriptors with high explanatory/ predictive power in the final models.

Thus, QSAR has great potential for modeling and designing novel congeners enable to forecast biological activities as a function of structural features and or physicochemical properties. Following the successful utilization of linear free-energy relationships, numerous 2D- and 3D-QSAR methods have been developed, most of them based on descriptors for hydrophobicity, polarizability, ionic interactions and hydrogen bonding. QSAR models used for the prediction of biological activity (or toxicity), as well as the evaluation of absorption, distribution, metabolism, and excretion. It has a particular interest in the preclinical stages of drug discovery to replace tedious and costly experimentation, to filter large chemical databases, and to select a few drug candidates. By quantifying physicochemical properties, it is possible to predict the biological activities of novel analogues prior to their synthesis. The main advantages of QSAR study are:

- it allows the medicinal chemist to target efforts on analogues which should have improved activity and thus cut down the number of analogues which have to be made.
- if an analogue is discovered which does not fit the model equation, it suggests that some other feature is important and offers a lead for further development.

Furthermore, the methodology is not dependent on the original dataset. All that is required is the availability of the original residuals. Another attractive feature is that apart from the threshold residual value, the methodology does not require extra information such as similarity measures or new descriptors, since it restricts itself to using the descriptors that were used in the original quantitative model. Thus, the study has reached to a stage where it can be used as an alternative for both lead identification and optimization. It provides powerful tool for virtual screening and can accompaniment well with the current techniques of combinatorial chemistry and high throughput screening in drug discovery research.

REFERENCES

1. Brown AC and Fraser TR. Trans Roy Soc (Edinburgh) (1869) 25:151-203.
2. Craig PN. J Med Chem (1971) 14:680-684.
3. Topliss JG. J Med Chem (1972) 15:1006-1011.
4. Topliss JG. J Med Chem (1977) 20:463-469.
5. Darvas F. J Med Chem (1974) 17:799-804.
6. Bustard TM. J Med Chem (1974) 17:777-778.
7. Hansch C, Maloney PP, Fujita T and Muir RM. Nature (London) (1962) 194:178-180.
8. Hansch C and Fujita T. J Am Chem Soc (1964) 86:1616-1626.
9. Hansch C. Acc Chem Res (1969) 2:232-239.
10. Hansch C. "Drug Design", (Ariens EJ, Ed.) (1971) Vol- I. Academic Press, New York.
11. Hansch C. in "Structure-Activity Relationships" "International Encyclopedia of Pharmacology and Therapeutics" (Cavolto CJ, Ed.) Peragamon Press, Oxford.
12. Hansch C. J Med Chem (1976) 19:1-6.
13. Kubinyi H. Arzneim-Forsch (1976) 26:1991-1997.
14. Kubinyi H and Kehrhahn O-H. J Med Chem (1976) 19:1040-1049.
15. Kubinyi H. J Med Chem (1977) 20:625-629.
16. Kubinyi H. Arzneim-Forsch (1978) 28:598-601.
17. Free SM Jr and Wilson JW. J Med Chem (1964) 7:395-399.
18. Fujita T and Ban T. J Med Chem (1971) 14:148-152.
19. Nilsson NJ. "Leaning Machines" (1965) McGraw-Hill, New York.
20. Kowalski BR and Bender CF. J Am Chem Soc (1972) 94:5632-5639.
21. Kowalski BR and Bender CF. J Am Chem Soc (1973) 95:686-693.
22. Enslein K, Ralston A and Wife HS. "Mathematical Methods for Digital Computers" (1977) Vol-III. Wiley-Interscience, New York.
23. Hodes L, Hazard GF, Geran RI and Richman S. J Med Chem (1977) 20:469-475.
24. Weiner ML and Weiner PH. J Med Chem (1973) 16:655-661.
25. Cammarata A and Menon GK, J Med Chem (1976) 19:739-748.

26. Menon GK and Cammarata A. J Pharm Sci (1977) 66:304-313.
27. Hansch C, Unger SH and Forsythe AB. J Med Chem (1973) 16:1217-1222.
28. Snedecor GW and Cochran WG. "Statistical Methods" (1967) Iowa State University Press, Ames.
29. Draper NR and Smith H. "Applied Regression Analysis" (1966) Wiley, New York.
30. Wold S. Quant Struct-Act Relat (1991) 10:191-193.
31. Akaike H. “Second International Symposium on Information Theory”, (Petrov B and Csáki F, Eds.) (1973) pp. 267-281, Budapest.
32. Akaike H. IEEE Trans Automat Contr (1974) AC-19:716-723.
33. Kubinyi H. Quant Struct-Act Relat (1994) 13:285-294.
34. Kubinyi H. Quant Struct-Act Relat (1994) 13:393-401.
35. Friedman J. Technical Report No. 102, (Nov. 1988; revised Aug. 1990) Laboratory for Computational Statistics, Department of Statistics, Stanford University.
36. Friedman J. The Annals of Statistics (1991) 19:1-67
37. Wu W, Walczak B, Massart DL, Heuerding S, Erni F, Last I R and Prebble KA. Chemom Intell Lab Syst (1996) 33:35-46.
38. Roy PP, Paul S, Mitra M and Roy K. Molecules (2009) 14:1660-1701.
39. Roy PP and Roy K. QSAR Comb Sci (2008) 27:302-313.
40. Rogers D and Hopfinger AJ. J Chem Inf Comput Sci (1994) 34:854-866.
41. Roy K and Paul S. QSAR Comb Sci (2009) 28:406-425.
42. So SS and Karplus M. J Med Chem (1997) 40: 4347-4359.
43. Afantitis A, Melagraki G, Sarimveis H, Koutentis PA, Markopoulos J and Markopoulou OI. Bioorg Med Chem (2006) 14:6686-6694.
44. Hansch C, Leo A and Elkins D. Chem Rev (1971) 71:525-616.
45. Green JP, Johnson CL and Kang S. Ann Rev Pharmacol (1974) 14:319-342.
46. Fujita T, Iwasa J and Hansch C. J Am Chem Soc (1964) 86:5175-5180.
47. Fujita T, Iwasa J and Hansch C. J Med Chem (1965) 8:150-153.
48. Hansch C, Leo A, Unger SH, Kim KH, Nikaitani D and Lien EJ. J Med

Chem (1973) 16:1207-1216.

49. Hansch C and Leo A. "Substituent Constants for Correlation Analysis in Chemistry and Biology" (1979) Wiley-Interscience, New York.
50. Martin AJP. Biochem Symposia (1949) 3:4-15.
51. Haggerty WJ Jr and Murril EA. Res Develp (1974) 25(8):30-39.
52. Yamazaki M, Kakeya N, Morishita T, Kamada A and Aoki M. Chem Pharm Bull (Japan) (1970) 18:708-714.
53. Fujita T and Hansch C. J Med Chem (1967) 10:991-1000.
54. Hammett LP. "Physical Organic Chemistry" (1940) Ch 7, McGraw-Hill New York.
55. Jaffe HH. Chem Rev (1953) 53:191-261.
56. Ritchie CD and Sager WF. Prog Phys Org Chem (1964) 2:323- 400.
57. Scherrer RA and Howard SM. J Med Chem (1977), 20:53-58.
58. Swain CS and Lupton EC Jr. J Am Chem Soc (1968) 90:4328-4337.
59. Taft RW. "Steric Effects in Organic Chemistry" (1956) (Newman MS, Ed.) Wiley, New York.
60. Exner O. Coll Czech Chem Commun (1967) 32:1-23.
61. Moriguchi I. Chem Pharm Bull (Japan) (1975) 23:247-257.
62. Bondi A. J Phys Chem (1964) 68:441-452.
63. Moriguchi I, Kanada Y and Komatsu K. Chem Pharm Bull Japan (1976) 24:1799-1806.
64. Gupta SP and Prabhakar YS. J Sci Indus Res (1985) 44:189-198.
65. Kier LB, Hall LH, Murray WJ and Randic M. J Pharm Sci (1975) 64:1971-1974.
66. Kier LB, Murray WJ, Randic M and Hall LH. J Pharm Sci (1976) 65:1226-1230.
67. Murray WJ and Randic M. J Pharm Sci (1976) 65:1806-1809.
68. Murray WJ and Randic M. "Molecular Connectivity in Chemistry and Drug Research" (1976) Academic Press, New York.
69. Pullman B and Pullman A. "Quantum Biochemistry" (1963) Wiley-Interscience, New York.
70. Kier LB. "Molecular Orbital Theory in Drug Research" (1971) Academic

Press, New York and London.

71. Fukui K, Yonezawa T and Nagata C. Bull Chem Soc (Japan) (1954) 27:423-427.
72. Fukui K, Yonezawa T and Shingu H. J Chem Phys, (1952) 20:722-725.
73. Dunn WJ. Eur J Med Chem-Chim Ther (1977) 12:109-113.
74. Randic M. J Am Chem Soc (1975) 97:6609-6615.
75. DRAGON software version 3.0. By Todeschini R, Consomni V and Mauri A, (2003) Milano, Italy. http://disat.unimib.it/chm/ Dragon.htm
76. ChemDraw Ultra 6.0 and Chem Bats 3D Ultra, Cambridge Soft Corporation, Cambridge, USA.
77. Todeschini R and Consomni V. "Handbook of molecular descriptors: methods and principals in medicinal chemistry", (2000) Vol-11, (Mannhold R, Kubinyi H and Timmerman H, Eds.) Wiley, VCH.
78. Platt JR. J Chem Phys (1947) 15:419-420.
79. Hosoya H. Bull Chem Soc Jap (1971) 44:2332-2339.
80. Gutman I, Ruscic B, Trinajstic N and Wilcox CF Jr. J Chem Phys (1975) 62:3399-3405.
81. Entiger RC, Jackson DE and Snyder DA. Czech Math J (1976) 26:283-296.
82. Balaban AT. Chim Acta (1979) 53:355-375.
83. Balaban AT. Chem Phys Lett (1982) 89:399-404.
84. Balaban AT. Pure Appl Chem (1983) 55:199-206.
85. Barysz M, Jashari G, Lall RS, Srivastava AK and Trinajstic N. "Chemical Applications of Topology and Graph Theory" (1983) (King RB, Ed.) Elsevier, Amsterdam, The Netherlands.
86. Kier LB and Hall LH. "Structure-Activity Analysis, Research Studies", (1986) Wiley, Chichester, UK.
87. Kier LB. Quant Struct-Act Relat (1987) 6:8-12.
88. Needham DE, Wei IC and Seybold PG. J Am Chem Soc (1988) 110:4186-4194.
89. Skorobogatov VA and Dobrynin AA. Commun Math Comput Chem (MATCH) (1988) 23:105-151.

90. Schultz HP. J Chem Inf Comput Sci (1989) 29:227-228.
91. Buckley F and Harary F. “Distance Matrix in Graphs”, (1990) Addison-Wesley, Redwood City, CA.
92. Schultz HP, Schultz EB and Schultz TP. J Chem Inf Comput Sci (1992) 32:69-72.
93. Plavsic D, Nikolic S, Trinajstic N and Mihalic Z. J Math Chem (1993) 12:235-250.
94. Gutman I, Yeh YN, Lee SL and Luo YL. Indian J Chem (1993) 32A:651-661.
95. Gutman I. J Chem Inf Comput Sci (1994) 34:1087-1089.
96. Voelkel A. Computers Chem (1994) 18:1-4.
97. Trinajstic N, Babic D, Nikolic S, Plavsic D, Amic D and Mihalic Z. J Chem Inf Comput Sci (1994) 34:368-376.
98. Konstantinova EV and Skorobogatov VA. J Chem Inf Comput Sci (1995) 35:472-478.
99. Pogliani L. J Phys Chem (1996) 100:18065-18077.
100. Diudea MV. J Chem Inf Comput Sci (1996) 36:535-540.
101. Randic M. J Chem Inf Comput Sci (1997) 37:1063-1071.
102. Sharma V, Goswami R and Madan AK. J Chem Inf Comput Sci (1997) 37:273-282.
103. Ren B. J Chem Inf Comput Sci (1999) 39:139-143.
104. Gupta S, Singh M and Madan AK. J Chem Inf Comput Sci (1999) 39:272-277.
105. Gramatica P, Corradi M and Consonni V. Chemosphere (2000) 41:763-777.
106. Rücker G and Rücker C. J Chem Inf Comput Sci (1993) 33:683-695.
107. Harary F. “Graph Theory”, (1969) Addison-Wesley, Reading, MA.
108. Randic M, Brissey GM, Spencer RB and Wilkins CL. Computers Chem (1979) 3:5-13.
109. Randic M. Commun Math Comput Chem (MATCH) (1979) 7:3-60.
110. Randic M. J Chem Inf Compu tSci (1984) 24:164-175.
111. Randic M and Jurs PC. Quant Struct-Act Relat (1989) 8:39-48.

112. Balaban AT. "Numerical Modelling of Chemical Structures: Local Graph Invariants and Topological Indices in Graph Theory and Topology in Chemistry" (1987) (King RB and Rouvray DH, Eds.) Elsevier Amsterdam, The Netherlands.
113. Bonchev D, Mekenyan O and Trinajstic N. J Comput Chem (1981) 2:127-148.
114. Bonchev D. "Information Theoretic Indices for Characterization of Chemical Structures" (1983) Research Studies Press, Chichester, UK.
115. Koch R. Toxicol Environ Chem (1983) 6:87-96.
116. Shannon C and Weaver W. "The Mathematical Theory of Communication" (1949) University of Illinois Press, Urbana, ILL.
117. Magnuson VR, Harriss DK and Basak SC. "Topological Indices Based on Neighborhood Symmetry: Chemical and Biological Applications in Studies in Physical and Theoretical Chemistry" (1983), (King RB, Ed.), Elsevier, Amsterdam, The Netherlands.
118. Bertz SH. J Am Chem Soc (1981) 103:3599-3601.
119. Estrada E. J Chem Inf Comput Sci (1996) 36:844-849.
120. Estrada E. J Chem Inf Comput Sci (1998) 38:23-27.
121. Estrada E, Guevara N and Gutman I. J Chem Inf Comput Sci (1998) 38:428-431.
122. Markovic S and Gutman I. J Chem Inf Comput Sci (1999) 39:289-293.
123. Broto P, Moreau G and Vandycke C. Eur J Med Chem (1984) 19:66-70.
124. Moran PAP. Biometrika (1950) 37:17-23.
125. Geary RC. Incorp Statist (1954) 5:115-145.
126. Pearlman RS and Smith KM. "3D QSAR in Drug Design" (1998) Vol-II. (Kubinyi H, Folkers G and Martin YC, Eds.) Kluwer/ESCOM, Dordrecht, The Netherlands.
127. Burden FR. J Chem Inf Comput Sci (1989) 29:225-227.
128. Gálvez J, Garcìa R, Salabert MT and Soler R. J Chem Inf Comput Sci (1994) 34:520-525.
129. Lovasz L and Pelikan J. Period Math Hung (1973) 3:175-182.
130. Balaban AT, Ciubotariu D and Medeleanu M. J Chem Inf Comput Sci

(1991) 31:517-523.

131. Randic M. New J Chem (1995) 19:781-791.
132. Randic M. J Chem Inf Comput Sci (1995) 35:373-382.
133. Randic M and Razinger M. J Chem Inf Comput Sci (1995) 35:594-606.
134. Jug K. Croat Chem Acta (1984) 57:941-953.
135. Bird CW. Tetrahedron (1986) 42:89-92.
136. Mekenyan O, Peitchev D, Bonchev D, Trinajstic N and Bangov IP. Arzneim Forsch (1986) 36:176-183.
137. Bogdanov B, Nikolic S and Trinajstic N. J Math Chem (1989) 3: 299-309.
138. Nikolic S, Trinajstic N, Mihalic Z and Carter S. Chem Phys Lett (1991) 179:21-28.
139. Arteca GA. "Molecular Shape Descriptors in Reviews in Computational Chemistry" (1991) Vol-9, (Lipkowitz KB and Boyd D, Eds.) VCH Publishers, New York.
140. Mihalic Z, Nikolic S and Trinajstic N. J Chem Inf Comput Sci (1992) 32:28-37.
141. Randic M, Kleiner AF and de Alba LM. J Chem Inf Comput Sci (1994) 34:277-286.
142. Krygowski TM, Ciesielski A, Bird CW and Kotschy A. J Chem Inf Comput Sci (1995) 35:203-210.
143. Bath PA, Poirrette AR, Willett P and Allen FH. J Chem Inf Comput Sci (1995) 35:714-716.
144. Katritzky AR, Mu L, Lobanov VS and Karelson M. J Phys Chem (1996) 100:10400-10407.
145. Krygowski TM, Cyranski M, Ciesielski A, Swirska B and Leszczynski P. J Chem Inf Comput Sci (1996) 36:1135-1141.
146. Hemmer MC, Steinhauer V and Gasteiger J. Vibrat Spect (1999) 19:151-164.
147. Randic M and Krilov G. Int J Quant Chem (1999) 75:1017-1026.
148. Schuur JH, Selzer P and Gasteiger J. J Am Chem Soc (1996) 36:334-344.
149. Todeschini R, Lasagni M and Marengo E. J Chemom (1994) 8:263-273.
150. Todeschini R and Gramatica P. "3D QSAR in Drug Design" (1998) Vol. II

(Kubinyi H, Folkers G and Martin YC, Eds.) Kluwer/ESCOM Dordrecht, The Netherlands.

151. Consonni V, Todeschini R and Pavan M. J Chem Inf Comput Sci (2002) 42:682-692.
152. Consonni V, Todeschini R, Pavan M and Gramatica P. J Chem Inf Comput Sci (2002) 42:693-705.
153. Viswanadhan VN, Ghose AK, Revankar GR and Robins RK. J Chem Inf Comput Sci (1989) 29:163-172.
154. Stanton DT and Jurs PC. Anal Chem (1990) 62:2323-2329.
155. Kaliszan R, Osmialowski K, Tomellini SA, Hsu S-H, Fazio SD and Hartwick RA. Chromatographia (1985) 20:705-708.
156. Osmialowski K, Halkiewicz J and Kaliszan R. J Chromat (1986) 361:63-69.
157. Clare BW and Supuran CT. J Pharm Sci (1994) 83:768-773.
158. Karelson M, Lobanov VS and KatritzkyAR. Chem Rev (1996) 96:1027-1043.
159. Prabhakar YS. QSAR Comb Sci (2003) 22:583-595.
160. So SS and Karplus M. J Med Chem (1997) 40:4347-4359.
161. Afantitis A, Melagraki G, Sarimveis H, Koutentis PA, Markopoulos J and Markopoulou OI. Bioorg Med Chem (2006) 14:6686-6694.
162. Wold S. Technometrics (1978) 20:397-405.
163. Gramatica P. QSAR Comb Sci (2007) 26:694-701
164. Martens H and Naes T. "Multivariate Calibration" (1989) John Wiley and Sons Inc., Chichester.
165. Höskuldsson A. J Chemometrics (1988) 2:211-228.
166. Eriksson L, Johansson E, Kettaneh N and Wold S. "Multi- and megavariate data analysis: Principles and applications" (2001) Umeå: Umetrics.
167. Stahle L and Wold S. "Multivariate data analysis and experimental design in biomedical research. In: Progress in Medical Chemistry" (1988) Vol-25, (Eillis GP and West WB, Eds.) Elsevier Science Publishers, BV.
168. Geladi P and Kowalski B.Acta (1986) 185:1-17.

169. Curto JD and Pinto JC. Int Statist Rev (2007) 75:114-121.

170. Johnson RA and Wichern DW. "Applied Multivariate Statistical Analysis" (1992) Prentice-Hall, New York.

171. BUDYPLS (BUlk Data division bY PLS) Algorithm, "A novel network strategy of in situ PLS regression directed recursive data mining for chemo- and bio-informatics" by Prabhakar YS, Central Drug Research Institute, Lucknow 226 001, India.

Chapter 2

Antidiabetic Agents:

An Overview

CHAPTER 2

ANTIDIABETIC AGENTS: AN OVERVIEW

1. INTRODUCTION

β-Cells secrete insulin in islets of Langerhans as a response to elevated blood glucose. A severe increase in blood glucose induces a rapid release of insulin which is sustained for a short period (known as first phase) and then followed by longer period of lower secretion (the second phase)) which accounts for the most part of secretion of insulin. Diabetes and its complications are the outcomes of progressive reduction in β-cell mass or secretory capacity resulting to abnormal glucose metabolism. Diabetes mellitus (DM) is a major health concern all over the world [172]. Due to globalization, mechanization, and changes in human lifestyle and daily routines incidences of diabetes and obesity are continuously increasing [173]. As per the estimate of International Diabetes Federation (IDF), the diabetic population of age group 18-99 years was 451 million and this figure is supposed to be 693 million by 2045 [174]. Diabetes is a metabolic disorder which is characterized by hyperglycemia in a fasted or a fed state. This metabolic disorder is a result of defects in insulin action, insulin secretion, or both, which leads to persistent hyperglycemia [175]. When blood glucose is >130 mg/dl the risk of diabetes increases [176]. Autoimmunity or destruction of insulin-secreting pancreatic β-cells, insulin resistance, obesity, genetic polymorphism, ketoacidosis, sedentary lifestyle and improper diet are the primary causes of diabetes and other causes are enzymatic defects including incretin, dipeptidyl peptidase VI (DPP-VI), peroxisome proliferator activating receptors (PPARs) [177-179].

2. CLASSIFICATION OF DIABETES

Diabetes may be classified, on the basis of insulin deficiency, into the following types:

(i) Insulin Dependent Diabetes Mellitus (IDDM)

It is a result of cellular mediated autoimmune destruction of the pancreatic cells. It is also known as type 1 diabetes or juvenile onset diabetes because usually

occurs in children or young adults and accounts for 5-10% of patients. To control the glucose level in blood, regular supply of insulin injections is desired. The rate of β-cell destruction varies in infants and children, and in adults. The degeneration of β-cells is slower in adults whereas the deterioration of these in infants and children is rapid. As a result of it symptoms like ketoacidosis occur in children and young ones whereas in other individuals modest fasting hyperglycemia is exhibited which in response to stress or infection may change to severe hyperglycemia or ketoacidosis. Such patients are susceptible to higher risk for development of other autoimmune disorders such as Grave's disease, vitiligo, celiac sprue, autoimmune hepatitis, myasthenia gravis, Hashimoto's thyroiditis, Addison's disease and pernicious anemia [180]. This type of diabetes is common in people of African and Asian descent and hereditary pattern is followed [181].

(ii) Idiopathic Diabetes

A small number of type 1 diabetes patients having no etiologies are prone to ketoacidosis and have permanent insulinopenia. The occurrence of ketoacidosis takes places in episodes and the insulin deficiency level fluctuates between episodes. The idiopathic diabetes has genetic predisposition and insulin replacement therapy is absolutely needed depending on the condition of the patient [180].

(iii) Noninsulin Dependent Diabetes Mellitus (NIDDM)

It accounts for nearly 90-95% of all diabetes and also known as adult onset diabetes. Obesity, insulin resistance, and dyslipidaemia are the major metabolic syndromes which led to the epidemic of type 2 diabetes [182]. Oral hypoglycemic drugs are used for the treatment of this type of diabetes which is dietary in nature. Insulin resistance and loss of insulin secretion are contributory to the inception of disease. In developed countries, type 2 diabetes mellitus, the most common form of diabetes, is the fourth leading cause of death with a twofold excess mortality and two- to fourfold increased risk of coronary heart disease and stroke [183].

(iv) Gestational Diabetes Mellitus (GDM)

It was first diagnosed during pregnancy [184] and related to glucose intolerance resulting variable severity of hyperglycaemia [185]. The impaired glucose

intolerance, GDM, affecting nearly 14% women during pregnancy in the United States and is a major risk factor for type 2 diabetes in mothers [186]. The extent of the reported risk varies with the variations in ethnicity, selection criteria and tests for GDM and type 2 diabetes [187]. Respiratory distress syndrome, neonatal hypoglycemia and fetal macrosomia may be developed in gestational diabetes leading to increased rates of birth trauma, shoulder dystocia, and cesarean delivery. Such maternal and fetal complications may be decreased by adequate glycemic control in a strategic manner. The blood sugar in patients with gestational diabetes may successfully be controlled by diet, exercise, oral diabetes medication or insulin.

(v) Catamenial Hyperglycaemia

Inadequate insulin or poor insulin compliance, acute pancreatitis, stroke, drugs, metabolic disturbances within the body, negligence with the treatment and infection may lead to conditions of diabetic ketoacidosis ((DKA) [188]. The occurring of the uncontrolled hyperglycaemia with DKA before the menstrual cycle is called as catamenial diabetic ketoacidosis or catamenial hyperglycaemia. The requirement of insulin increases because of uncontrolled hyperglycemia. The condition is so aggravated that even after the continuous insulin infusion, resulting in vomiting, and leading to significant acidosis, ketonuria and hyperglycaemia. It is the strange fact that several tests including inflammatory markers, blood count renal function, electrocardiogram and chest radiograph, thyroid function and urine and blood cultures were all found to be normal in other words the conditions which lead to catamenial hyperglycaemia remained undiagnosed [189]. Hormonal changes altogether with changes in diet and exercise levels occurred during menstrual cycle may play a significant role [190]. To avoid any diabetic emergencies, the right medication strategy for the treatment of catamenial diabetic ketoacidosis is the increased insulin infusion dosage with effective diet and exercise plans [191].

Type 1 and type 2 diabetes are more common types of diabetes. Risk of developing T2DM is associated with the alteration in glucose metabolism. The risk factors in the development of insulin resistance, loss of pancreatic function, worsening of hyperglycemia and progression to diabetes are excess adiposity,

inflammation and dyslipidemia [192]. Not only type 2 diabetics but prediabetics (presently defined as moderately elevated fasting blood glucose, FBG) also are at increased risk for a wide range of debilitating diseases. Diabetes emerged as the leading cause of kidney failure, blindness and of nontraumatic lower limb amputation. The multitude of cardiovascular disease (CVD) is 2 to 4 times higher in diabetics [193].

The accumulation of fat in hepatocytes (steatosis) which leads to the chronic liver disorder Non-Alcoholic Fatty Liver Disease (NAFLD) and its more advanced form, Non-Alcoholic Steato-Hepatitis (NASH) is the potential fatal complication of T2DM. The addressing of these serious complications of T2DM is important as NAFLD/NASH can progress to hepatitis, cirrhosis, and even liver cancer. The current glucose-lowering treatments are beneficial but the disease related morbidity and mortality remained considerable in patients having T2DM.

Thus there is ardent desire of innovative medications which target the multiple metabolic abnormalities, inflammatory processes and other pathways predisposing to diabetes-associated disorders.

The prevention of long-term complications and the treatment of associated disorders such as NAFLD/NASH and CVD are the challenges in the management of T2DM disease. The association between the degree of hyperglycemia and the risk of micro- and macrovascular complications including fatal CVD events has shown in T2DM prospective studies.

The ACCORD and ADVANCE trials in patients with longstanding T2DM revealed that aggressive glucose control in such patients has no clear benefits, or even may increase CVD events [194] suggesting the existence of other independent risk factors which contribute significantly to CVD risk in T2DM patients. These findings reflect upon the limitations of current anti-diabetic therapies, because the off-target effects countered the potential benefits of glucose lowering. New therapeutics must be aimed at to treat diabetic patients at an earlier stage of the disease and able to address the multi-factorial nature of T2DM.

3. CURRENT AND FUTURE THERAPEUTIC APPROACHES

Metformin (suppressor of hepatic glucose production), sulfonylureas (insulin secretagogues) and thiazolidinedione pioglitazone, (PPAR agonist) are being used at present as the widespread treatments for T2DM. Glucagon-like peptide-1 (GLP-1) mimetics and inhibitors of the enzyme that degrades GLP-1 (dipeptidyl peptidase-4, DPP-4) are being employed in the incretin-based treatment strategies. GLP-1 promotes satiety and weight loss. This intestinally-derived peptide stimulates insulin as a response to food intake and reduces the rate of gastric emptying. Exenatide, the GLP-1 mimetic, was approved by the FDA in 2005 and its indication was extended in 2009 to standalone therapy for T2DM. However, a certain number of gastrointestinal side-effects persist with the exenatide. It showed a reduced incidence of cardiovascular events in a retrospective analysis of almost 40,000 patients [195]. Novel therapeutic approaches in the area of T2DM drug discovery are specifically designed keeping in mind the multi-factorial nature of T2DM by targeting multiple diabetes-related indications and should not be focused simply on the glucose-lowering.

Current FDA recommendations, because of the elevated CVD risk in T2DM, require that all new anti-diabetic drugs show exemplary cardiovascular safety profiles. In this way, drugs that target molecular pathways having potential implications in both diabetes and CVD are especially desirable. The targeting of 11β-hydroxysteroid dehydrogenase type 1 (11β-HSD1), GPR119, TGR5, sirtuin 1 (SIRT1), the sodium-glucose co-transporter 2 (SGLT2) and GPR40 are examples of such approaches. The rationale of each is briefly described below-

3.1. THE 11β-HYDROXYSTEROID DEHYDROGENASE TYPE 1 (11β-HSD1)

Glucocorticoids are steroid hormones which bind to the glucocorticoid receptor (GR) and exert powerful anti-inflammatory and immunosuppressive effects. The treatment of patients with glucocorticoids develops obesity, insulin resistance, glucose intolerance and dyslipidemia [196]. In the USA, more than 2.5 million people are exposed to long-term glucocorticoids [197] and the insulin resistance due to glucocorticoid exposure became a public health problem. Tissue-specific metabolism

of glucocorticoids is catalyzed by two enzymes, one is 11β-hydroxysteroid dehydrogenases type 1 (11β-HSD1) and the other is type 2 (11β-HSD2). The inter-conversion of non-receptor binding cortisone and the receptor binding active form, cortisol is carried out by these enzymes. Inactive cortisone in the liver, adipose tissue, vasculature and brain is converted to the active cortisol by the NADP(*H*)-dependent enzyme 11β-HSD1 [198-200]. On the other hand, the NAD-dependent dehydrogenase, 11β-HSD2, inactivates cortisol to cortisone in the kidney and colon [201]. A well established role of 11β-HSD1 in obesity and metabolic disease in rodents is observed. The adipose tissue-specific aP2 promoter, driven overexpressed 11β-HSD1 showed elevated corticosterone levels in adipose tissue which displayed a phenotype mimicking human metabolic syndrome and that was characterized by visceral obesity, insulin resistance, and hyperlipidemia [202, 203]. A study revealed that administration of glucocorticoids in mice induces metabolic syndrome which was prevented in 11β-HSD1 knockout mice [204]. These findings in addition to tissue-specific expression of 11β-HSD1 implied that the intracellular metabolism of glucocorticoids by 11β-HSD1 is critical to the development of insulin resistance rather than the circulating glucocorticoids. In the treatment of a variety of diseases, 11β-HSD1 emerged as an important therapeutic target for reducing adverse effects of prescribed glucocorticoids.

The mechanism of insulin resistance with the increased levels of 11β-HSD1 is not fully clear. In adipose tissue, the increased levels of leptin, resistin, tumor necrosis factor-α (TNF-α) and interleukin-6 (IL-6) by the overexpression of 11β -HSD1 gene [205, 206] suggested that high local levels of glucocorticoids promote an inflammatory activity through cortisol. The inflammatory function, exerted by glucocorticoids may be regulated through a central player in the insulin signaling in diabetes and insulin resistance that is c-Jun N-terminal kinases (JNK). Findings showed that JNK knockout mice are protected against the development of insulin resistance [207, 208] and in insulin-resistant rodents, the administration of small molecule or peptide inhibitors of JNK significantly improved insulin sensitivity [209, 210]. JNK can be activated by multiple factors including inflammatory cytokines and

free fatty acids. Based on the fact that the increased JNK activity in epithelial cells [211], hippocampal cells [212] and endothelial cells [213] by glucocorticoids, a study using a high fat diet (HFD) mouse model and cultured adipocytes indicated that glucocorticoid-induced insulin resistance was dependent on 11β-HSD1 and resulted in the critical activation of JNK signaling in adipocytes [214].

The mechanism of insulin resistance mediated by 11β-HSD1 is depicted in the Figure 2.1.

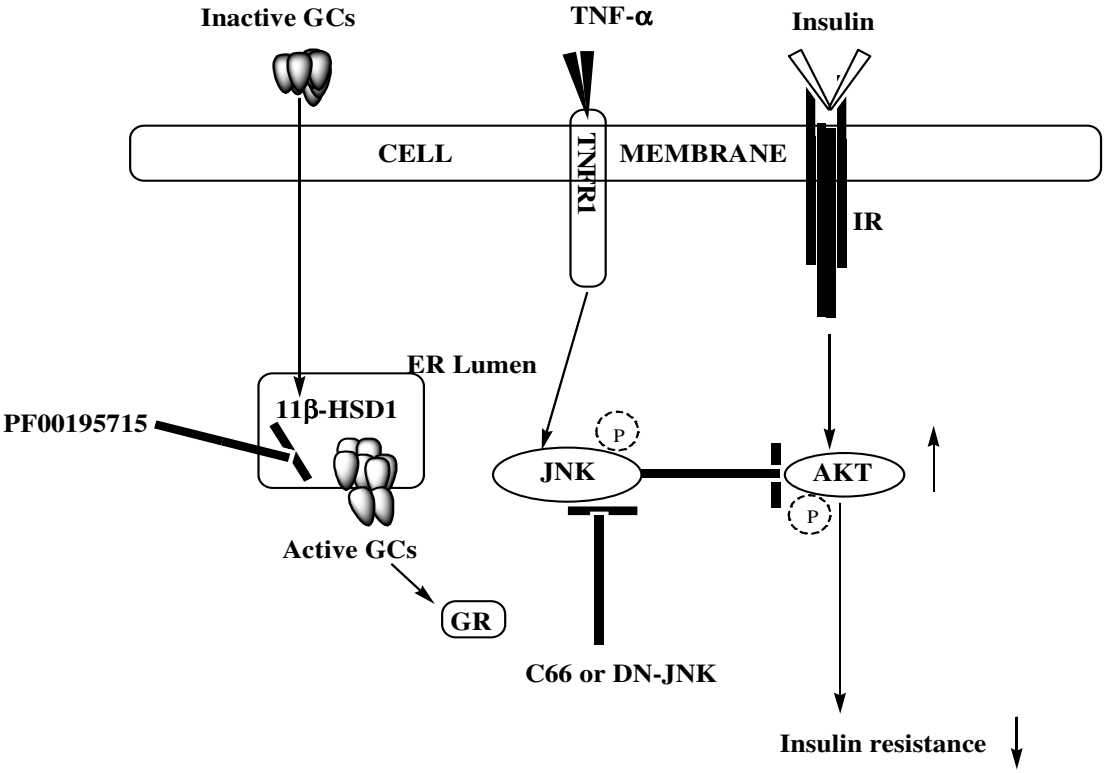


Figure 2.1: Mechanism of insulin resistance mediated by 11β-HSD1. 11β-HSD1 converts inactive glucocorticoid into active glucocorticoid in adipocytes. The JNK pathway is activated by active glucocorticoid-GR signaling complex and JNK inhibits insulin-induced Akt phosphorylation leading to insulin resistance. Inflammatory cytokines such as tumor necrosis factor-α (TNF-α) may also activate JNK. GCs, GR, IR and JNK represent glucocorticoids, glucocorticoid receptor, insulin receptor and c-Jun N-terminal kinase, respectively. PF00915275 and C66 are specific small-molecule inhibitors of 11β-HSD1 and JNK, respectively.

It is expected that 11β-HSD1 inhibitors may address several metabolic syndrome related aspects.The glucocorticoid receptor on excessive activation can generate multiple clinical features such as insulin resistant diabetes, obesity, dyslipidemia and hypertension that are characteristic of the metabolic syndrome. 11β-HSD1 locally generates the active glucocorticoid cortisol from cortisone under normal cellular activity thus there is a notion that specific inhibitors have potential in improving above mentioned conditions [215-221].

Studies reported that inhibition of the enzyme in both liver and adipose tissue is the most beneficial [222, 223]. The inhibition of the enzyme in the liver influences gluconeogenesis [224, 225] whereas in the adipose tissue have positive effects on adipocyte differentiation (reduced weight) and production of adipokines which is entailed in the metabolic syndrome (e.g. adiponectin) [226]. Thus it is expected that prolonged inhibition of 11β-HSD1, in adipose tissue and liver both, would be effectual in treating diabetes with a potential for positive effects on hypertension and dyslipidemia. The nonselective 11β-HSD1/11β-HSD2 inhibitor, carbenoxolone, showed improved insulin sensitivity in rodent and humans [227, 228] but with a limited utility. 11β-HSD2 enzyme converts cortisol to its inactive metabolite cortisone predominantly in the kidney. Selectivity over the 11β-HSD2 enzyme is important to avoid apparent mineralocorticoid excess syndrome which is a result of cortisol action on the mineralocorticoid receptor to induce sodium retention, hypokalemia and hypertension [229]. A number of selective synthetic inhibitors of 11β-HSD1 have been described [230]. These compounds have demonstrated their utility as anti-diabetic agents [231] in preclinical models. In clinical trials, several of the 11β-HSD1 inhibitors showed modest improvements in glycemic control and demonstrated components of the metabolic syndrome [232-236].

The 11β-HSD1, expressed abundantly in metabolically important tissues including adipose, muscle and liver tissue that become resistant to insulin action in type 2 diabetes. The development of the drugs inhibiting 11β-HSD1 is urgently desired since the inhibition of it renders the ability to restore the metabolic action of insulin in these tissues.

3.2. THE G PROTEIN-COUPLED RECEPTOR 119 (GPR119)

Except DPP-4 inhibitors, several deorphanized nonpeptide binding G protein-coupled receptors (GPCRs) are being evaluated for the T2DM as candidate GLP-1 secretagogues [237, 238]. Much attention received from the pharmaceutical industry is the G protein-coupled receptor 119 (GPR119). GPR119 is an attractive drug target for treating T2DM and agonists of it may be represented as new potential insulin secretagogues devoid of the risk of causing hypoglycemia.

GPR119, described as a class A (rhodopsin-type) orphan GPCR, is having no any primary sequence relative in the human genome [239]. The increase in the intracellular accumulation of cAMP on activation of GPR119 results into enhanced glucose-dependent insulin secretion from pancreatic β-cells and increased release of the gut peptides GLP-1 (glucagonlike peptide 1), GIP (glucose-dependent insulinotropic peptide) and PYY (polypeptide YY) [240]. GPR119 agonists have been proposed as a novel therapeutic strategy for diabetes because in preclinical and clinical studies with GPR119 agonists in type 2 diabetes there are indications of lowering blood glucose without hypoglycemia, slowing down of diabetes progression and reducing food intake and body weight.

Based on the data afforded by the Human Genome Project the GPR119 was described in the literature as a Class A receptor with no close relatives. The independent studies described this receptor under various synonyms such as SNORF25 [241, 242], RUP3 [243], GPCR2 [244], 19AJ [245], OSGPR116 [246], MGC119957, HGPCR2 and glucose-dependent insulinotropic receptor (GDIR) [247]. The confusing nomenclature has now been rationalized as “GPR119”. The human GPR119 receptor is encoded by a single exon with introns located on the short arm of X-chromosome (Xp26.1). The homologs of GPR119 have been identified in several vertebrate species such as the rat, mice, hamster, chimpanzee, rhesus monkey, cattle and dog [245]. Fredriksson *et al.* [239] reported that the rat isoform of GPR119 as being 133 amino acids longer than the mouse and human receptors (468 vs. 335 amino acids) [248]. On the other hand another reports by Bonini *et al.* [241, 242] and Ohishi *et al.* [249] documented identical sequences for the rat receptor, which are 335

amino acids in length having 96% amino-acid identity with mouse GPR119.

It has been proposed, using methods to detect receptor GPR119 mRNA, that in human tissues consistently identified major sites of GPR119 mRNA expression are the pancreas and foetal liver and the gastrointestinal tract in several studies. In rodents, mRNA was detected in most of the examined tissues including the pancreas [250] and gastrointestinal tract particularly the colon and small intestine. It is also expressed in certain regions of the rat brain. It is revealed from the *in situ* studies that in pancreatic islets the main site of GPR119 expression are pancreatic β-cells [251] and this observation is supported by the high expression levels in pancreatic β-cell lines NIT-1, MIN6 and RIN5 [252, 253]. With the consistent expression in gut tissues, GPR119 mRNA expressed strongly in several rodent GLP-1 secreting L-cell lines-such as STC-1, FRIC, Hnci-h716 and GLUTag line [253, 254].

Presence of GPR119 mRNA has also evinced in glucosedependent insulinotropic peptide (GIP)-producing murine intestinal K cells [255]. In transfected HEK293 cells high-level expression of GPR119 increases intracellular cAMP levels via activation of adenylate cyclase which indicates efficient coupling of this receptor to $G\alpha_s$. Increase in cAMP levels by the potential endogenous ligands and synthetic small molecule agonists of GPR119 support it. The possible actions of GPR119 have been shown in Figure 2.2.

The first proposed endogenous ligand for GPR119, based on the ability to stimulate glucose-dependent insulin release and increase cAMP in GPR119-transfected cells, was lysophosphatidylcholine (LPC). The potency to promote a concentration-dependent increase in cAMP levels in stably transfected and endogenous GPR119-expressing cell lines of fatty-acid amide oleoylethanolamide (OEA) was more than that of LPC [256]. OEA produced a number of pharmacological effects in rodent studies [257] such as reduced food intake and body weight gain by interaction with the nuclear receptor peroxisome proliferator activated receptor α (PPAR-α) [258], increased fatty acid uptake by adipocytes and enterocytes by increasing fatty acid translocase expression [259], altered feeding behaviour and motor activity through activation of an ion channel (the transient receptor potential

channel, TRPV1) [260].

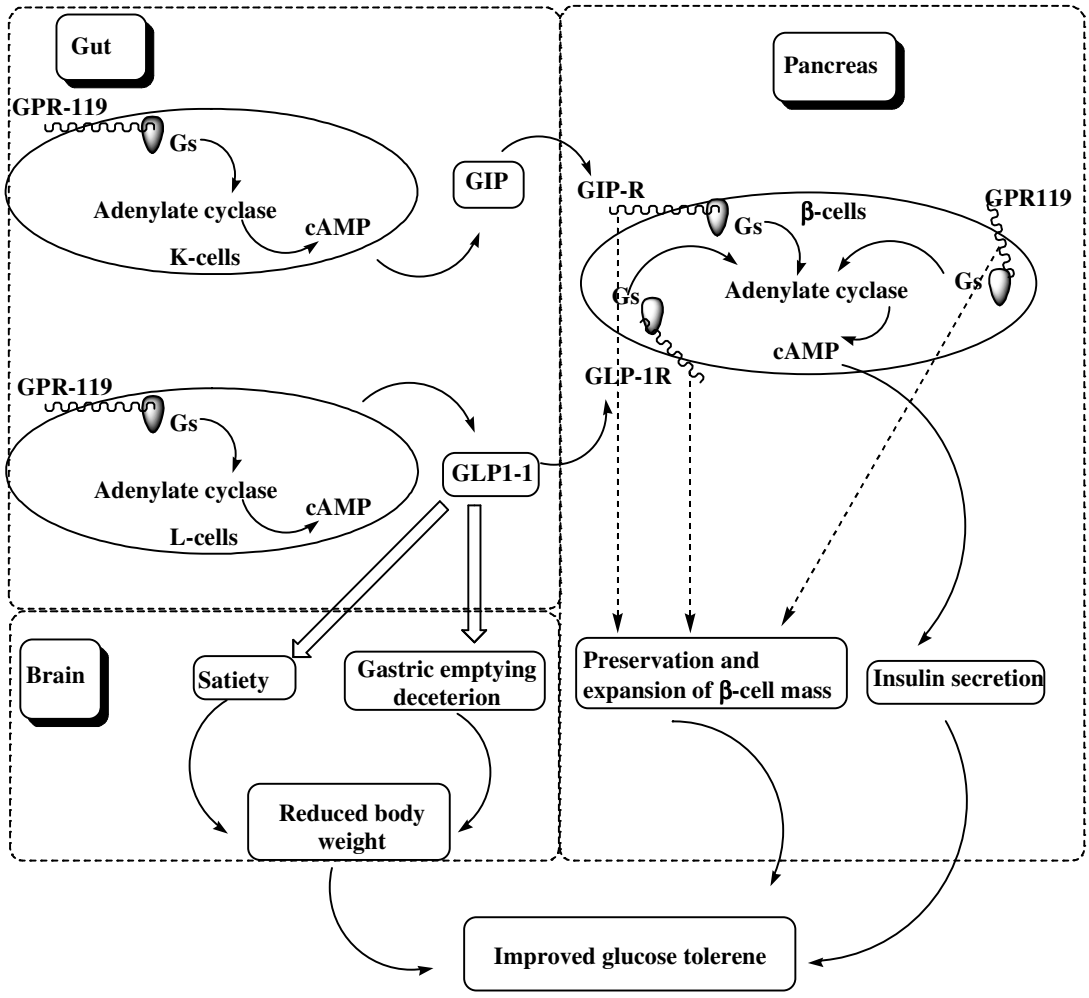


Figure 2.2: Diagram showing the possible actions of GPR119 agonists.

The endovanilloid compounds, *N*-oleoyl dopamine (OLDA) and olvanil, having similar *in vitro* potencies as of OEA also described as GPR119 agonists. The increased GIP release and improved oral glucose tolerance on oral administration of OLDA (100 mg/kg) in mice has also been observed in in vivo studies which were not present in GPR119 null mice. However, OEA and OLDA are less potent and selective than the natural ligands identified for many other GPCRs, represent the best candidates for endogenous ligands. This work opens the scope for other lipid amides to exert physiological role via GPR119 signaling.

A number of publications focusing on the problem of exogenous influence on the incretin system are available in literature [261-263]. Incretins are gastrointestinal-derived hormones which are released in response to a meal and a key role in the regulation of postprandial secretion of insulin and glucagon by the pancreas is played by them [264]. The incretin system is getting much attention because the incretin effect is severely reduced or absent in patients with T2D [265]. In this context the restoration of adequate incretin biosynthesis and metabolism may be a potential strategy of T2D treatment [266]. This approach is devoted to the development of such drugs which are able to stimulate the incretin secretion by activating the GPCR expressed on the intestinal enteroendocrine cells. Receptors of this group function as the sensors of fatty acids, their derivates and some other digestion products. The stimulation of incretin secretion, by the activation of such receptors, stimulates the synthesis and secretion of insulin leading to a state of postprandial normoglycemia [267, 268]. Glucose-depended activation of insulin secretion [269, 270] is the result of the activation of GPR119, expressed in L- and K-cells of intestine as well as in pancreatic β-cells [271].

Such type of mechanism of agonistic action of GPR119 is supposed to be advantageous because it offered a pronounced antihyperglycemic effect devoid of risk of excess hypoglycemia and rendering such substances as promising candidates for the role of drugs for T2D treatment [272]. The non-clinical studies and investigations performed previously in healthy volunteers has established that GPR119 are capable to increase the level of circulating incretins including GLP-1, GIP and tyrosine-tyrosine peptide (PYY) and reduce the hyperglycemia after oral glucose load [273]. The demonstrated several secondary pharmacodynamic effects such as cerebral, cardiac and endothelial protection in animal studies are contrary to the antidiabetic medications, as that have only the hupoglycemic action. Due to these secondary (or “pleiotropic”) effects the GPR119 agonists might be essential for the prevention of T2D complications [274, 275].

The investigations of several research groups [276, 277] on multiple small-molecule GPR119 agonists led to the development of clinical compounds which

include APD668 [278], GSK1292263 [279] and MBX-2982 [280]. Poor aqueous solubility of GPR119 agonists causes low bioavailability, produces erratic assay results in *in vitro* studies and carries a high risk of not advancing due to potential toxicity which may not be recognized during preclinical studies [281, 282]. The value of GPR119 agonists as a new class of therapeutics for T2D and associated obesity is likely to be determined in due course of time.

3.3. THE TAKEDA G PROTEIN RECEPTOR 5 (TGR5)

The intestinal absorption, emulsification, and transport of lipophilic nutrients and vitamins by bile acids (BAs) are facilitated by the amphipathic steroid molecule possessed by them. BA is the catabolism product of cholesterol in the liver. In recent years, BA showed pleiotropic responses [283] similar as the endogenous molecules such as glucose and energy homeostasis [284]. It is also found that some of the BAs scape the enterohepatic cycling to reach the systemic circulation [285]. Participation of BAs in various functional processes like lipid and glucose homeostasis, energy expenditure, intestinal mobility, inflammation [286], configuration, and the growth of gut microbiome or the skeletal muscle mass [287] is well established. There are also indications of involvement of dysregulated signaling of BAs in various disorders such as diabetes, obesity, dyslipidemia, fatty liver disease, atherosclerosis, cholestasis, gallstones and cancer [288]. BAs furnish these effects in multiple organs basically by binding with the nuclear hormone farnesoid X receptor (FXR) and Takeda G protein receptor 5 (TGR5) [289].

The clinical treatment of T2DM patients with the BA-like agent(s) or bariatric surgery in obese patients, showed a noticeable improvement in glycemic control which are possibly due to changes in TGR5 and signaling. The G protein-coupled receptor, TGR5 is expressed in many tissues such as intestine, gallbladder, adipose tissues, skeletal muscle, brain and pancreas. Thus, the activation of TGR5 by BA induces the formation of the cyclic AMP (cAMP) which in turn may activate protein kinase A (PKA) in cells and tissues [290]. In human TGR5-transfected CHO cells tauro-lithocholic acid (TLCA), lithocholic acid (LCA), deoxycholic acid (DCA), chenodeoxycholic acid (CDCA) and cholic acid (CA) induced cAMP production in

dose-dependent manner. The rank order of potency in terms of EC50 was found to be TLCA (0.33 μM) >LCA (0.53 μM) >DCA (1.01 μM) >CDCA (4.43 μM) >CA (7.72 μM) [291]. CDCA, DCA, LCA and ursodeoxycholic acid (UDCA) may also be involved in activation of FXR [292]. Linolenic acid and oleanolic acid [293], ursolic acid [294] and glycyrrhizic acid [295] also activate TGR5 and triamterene appeared as as the useful blocker of TGR5 [296]. The induction of insulin secretion is more pronounced in oral glucose rather than an isoglycemic intravenous injection.

Thus, entero-endocrine K- and L-cells are identified which are known to secrete the incretins, both glucose-insulinotropic polypeptide (GIP) and glucagon-like peptide (GLP)-1. In L-cells, the action of prohormone convertase 1/3 just after the transcription and translation into proglucagon, leads to secretion of GLP-1, GLP-2, oxyntomodulin and IP2. On the other hand, the action of prohormone convertase 2 in pancreatic α-cells leads to glucagon, glicentin-related polypeptide, IP1 and major proglucagon fragment [297]. The half-life of GLP-1 in blood is about 1.5-5 min because of rapid degradation of it by dipeptidyl peptidase 4 (DPP-4). Therefore in the treatment of T2D, inhibitors of DPP-4 are being used successfully now.

The activation of TGR5 promotes GLP-1 secretion from intestinal L cells as a result of a closure of the ATP-dependent potassium channel (KATP) and a higher mobilization of intracellular calcium to enhance GLP-1 secretion. GLP-1 biosynthesis and secretion is also enhanced by glucose. However, GLP-1 secretion by intestinal L cells is negatively regulated by FXR through inhibition of pro-glucagon gene expression and suppression of GLP-1 secretion through the interfering with pathways activated by glucose [298]. Thus, activation of both TGR5 and FXR by BA in intestinal L cells might induce opposite effects on GLP-1 secretion and production. TGR5 activation in L cells likely occurs rapidly after intaking the food, on the other hand activation of FXR induces a more delayed response which requires transcriptional activation. The pancreatic β-cells having expression of both of TGR5 [299] and FXR [300] promotes glucose-stimulated insulin secretion by increasing intracellular calcium concentration. TGR5 is identified in pancreatic α-cells in pancreatic islet. In pancreatic islet, the activation of TGR5 in pancreatic α-cells,

switches the α-cell secretory phenotype from glucagon to GLP-1 and the result of which is a paracrine effect on β cells to stimulate insulin secretion [301].

T2DM, well known as a heterogeneous group of disorders, is mostly characterized by a decline in insulin-producing pancreatic β-cells, an increase in peripheral insulin resistance, an increase in hepatic glucose production or a combination of all these factors [302]. The T2DM therapies are mostly paying attention on reducing of hepatic glucose production, increasing of insulin secretion and improving insulin sensitivity [303]. TGR5 being a receptor of bile acids effects the regulation of glucose homeostasis. In a murine enteroendocrine cell line, STC-1, activation of TGR5 promoted GLP-1 secretion [290]. The ability to enhance insulin secretion after oral administration of glucose by GLP-1 advocated the potential treatment of T2DM via the management of glucose homeostasis by activatingTGR5. In addition to this, TGR5 might also induce cAMP-dependent thyroid hormone activating enzyme type 2 iodothyronine deiodinase that may cause elevated energy expenditure in brown adipocytes and skeletal muscles [304]. The differential translation of the C/EBPb isoform by AKT-mTOR pathway in macrophages is also induced by TGR5. The insulin action for treatment of T2DM may be improved by the activation of TGR5 through altering adipose tissue macrophage function [305]. The other possible mechanism may be connecting TGR5 signaling and elevated energy expenditure via modifications in the gut microbiome [306]. Thus, TGR5 activation for T2DM is not solely dependent on GLP-1. Furthermore, TGR5 also plays role in inhibiting renal disease in obesity and diabetes through inducing mitochondrial biogenesis and help to prevent renal oxidative stress and lipid accumulation [307]. In obesity, new roles of TGR5 have also documented [308].

The gallbladder volume in mice has been increased due to systemic exposure to TGR5 agonists [309]. The investigation in mice and dogs of an agonist of TGR5, FC-92-EC85, have shown hepatobiliary and cardiovascular effects which limits the utility of systemic TGR5 agonist in diabetes [310]. A novel topical intestinal agonist of TGR5 which was given orally to obese and insulin-resistant mice demonstrated not only a prominent elevation in GLP-1 levels but significant improvement in glucose

tolerance also. In lean mice, intestinal TGR5 agonist did not produced a significant change in gallbladder size [311]. Therefore, it is expected that an ideal TGR5 agonist must be intestinal-specific agonist reaching L cells and must not affecting other systemic tissues. Although, the impact of the intestinal TGR5 agonist on human gallbladder is still remained unclear and the therapeutic potential for T2DM in the clinic needs this issue must be addressed in advance.

3.4. THE SIRTUINS

Aging, which affects all organs, is a universal process. The result of age-related commotions in cellular homeostasis is in the form of decline in the responsiveness to physiological stress such as oxidative stress and inflammation which have implications in the pathogenesis of insulin resistance and T2DM like metabolic diseases. One of the sources of reactive oxygen species (ROS) is mitochondria which play a key role in energy production and responsiveness to nutrient availability [312]. Thus the decline in mitochondrial function is also closely related to the impairment of metabolic homeostasis [313] and oxidative stress [314, 315] that are contributing to the progression of insulin resistance and T2DM associated with aging. The suppression of oxidative stress/inflammation and preservation of mitochondrial function must be considered as therapeutic targets for insulin resistance and T2DM and for anti-aging treatments because oxidative stress is closely linked to inflammation [316, 317].

In yeast, worms, flies and rodents the calorie restriction (CR) retarded aging or extended the life spans [318]. The beneficial effects of CR have also been observed in the suppression of age-related diseases, by improving insulin sensitivity and reducing inflammation and oxidative stress such as glucose intolerance, cardiovascular disease and cancer in rhesus monkeys or humans [319-321]. Sirtuins may play a significant role in modifying lifespan in relation to the benefits of CR, particularly. In a study on aging in yeast [322] silent information regulator 2 (Sir2), a nicotinamide adenine dinucleotide (NAD+)-dependent deacetylase was identified as the possible molecule by which CR promotes lifespan extension. In higher eukaryotic organisms the homologs of Sir2, known as SIRT1, may contribute to CR-induced

longevity [323-325]. In mammals there are seven identified sirtuins (including SIRT1) at present [326, 327)] and these are mentioned, along with their catalytic activity and localization in Table 2.1.

Table 2.1: The seven sirtuins in mammals

S. No.	Sirtuin	Catalytic activity	Localization
1	SIRT1	Deacetylase	Nucleus and cytoplasm
2	SIRT2	Deacetylase	Cytoplasm and nucleus
3	SIRT3	Deacetylase	Mitochondria
4	SIRT4	ADP-ribosyl transferase	Mitochondria
5	SIRT5	Deacetylase	Mitochondria
6	SIRT6	Deacetylase and ADP-ribosyl transferase	Nucleus
7	SIRT7	Deacetylase	Nucleus

In literature, the multiple physiological roles of sirtuins in cellular function like glucose metabolism, mitochondrial function and resistance against cellular stresses such as oxidative stress and inflammation has been documented [326-331]. It provides the basis for the modulation of sirtuin activity as a CR mimetic for insulin resistance and T2DM drug target. It has been reported that chronic inflammation, oxidative stress and impaired mitochondrial function in skeletal muscle, adipose tissue or monocytes/macrophages [332, 333] are intimately related to the pathogenesis of insulin resistance and T2DM. The dysfunction of pancreatic β-cell [334, 335] caused by inflammation and oxidative stress is contributive to the progression of T2DM.

In insulin-resistant and diabetic conditions the activation of monocytes in the circulation, adipocytes and macrophages residing in adipose tissue show the way to release of various inflammatory mediators such as tumor necrosis factor-α (TNF-α), interleukin-6 (IL-6) and chemoattractant protein-1 (MCP-1). Cytokines play a crucial role in the pathogenesis of insulin resistance in adipose tissue and skeletal muscle as inflammatory signaling pathways, such as the inhibitor of IkB kinase (IKK) and c-Jun NH_2-terminal kinase (JNK) pathways, activated by cytokines impair the insulin

signaling pathway by modulating phosphoinositide 3-kinase (PI3K) and Akt [336-338]. The impaired insulin signaling by oxidative stress is also contributory to the insulin resistance in T2DM. In insulin-resistant or diabetic conditions not only the hyperglycemia but other metabolites such as free fatty acids (FFAs) and certain cytokines such as TNF-α also induce the overproduction of ROS from mitochondria. The activation of serine/threonine kinases, such as p38 mitogen activated protein kinase (p38 MAPK), JNK and IKK triggered by ROS induces the serine phosphorylation of insulin receptor substrate-1 (IRS-1) which then degrades IRS-1 and reduces IRS-1 tyrosine phosphorylation that in turn leads to the suppression of insulin signaling [339-342] and inflammation as well. The pancreatic β-cell dysfunction is also related to the inflammatory mediators and oxidative stress as these resulted in the impairment of insulin production or excretion from β-cells.

In the pathogenesis of insulin resistance and progression of T2DM associated with aging the impairment of mitochondrial function in skeletal muscle is also involved. A pivotal role in energy production and responsiveness to nutrient availability is played by mitochondria as it regulates the mitochondrial oxidative phosphorylation (OXPHOS) and fatty acid oxidation. In the patients with insulin resistance and T2DM and in elderly individuals, the rate of mitochondrial OXPHOS is reduced and the intramyocellular lipid accumulation is increased in the skeletal muscle [343-346]. Aging is linked closely to the impairment of metabolic homeostasis like insulin resistance and T2DM and these are closely related to the decline in mitochondria function. The decline in mitochondrial function generates excess ROS from mitochondria linking oxidative stress to inflammation. In this way, oxidative stress, inflammation and mitochondrial dysfunction make a vicious cycle represented in Figure 2.3.

For the treatment of age-related insulin resistance and T2DM the breaking of this cycle may be a therapeutic target.

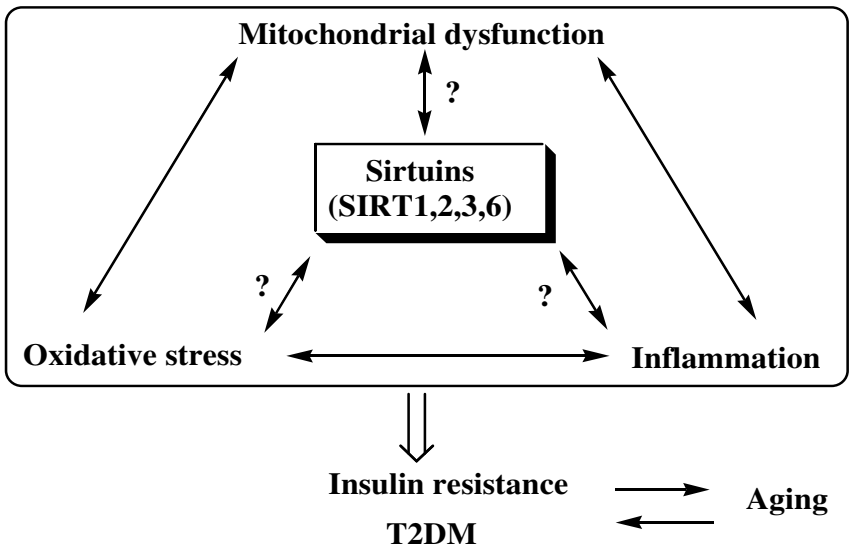


Figure 2.3: A vicious cycle among oxidative stress, inflammation, and mitochondrial dysfunction.

SIRT1

SIRT1, having NAD^+-dependent deacetylase activity, showed existence in the nucleus and cytoplasm [327]. SIRT1 functions as class III histone deacetylases. It binds to NAD^+ and acetyllysine within protein targets and generates lysine, 2'-O-acetyl-ADP-ribose, and nicotinamide as enzymatic products. Nicotinamide acts as a negative-feedback inhibitor of SIRT1. The enzymatic activities of SIRT1 have been depicted in Figure 2.4.

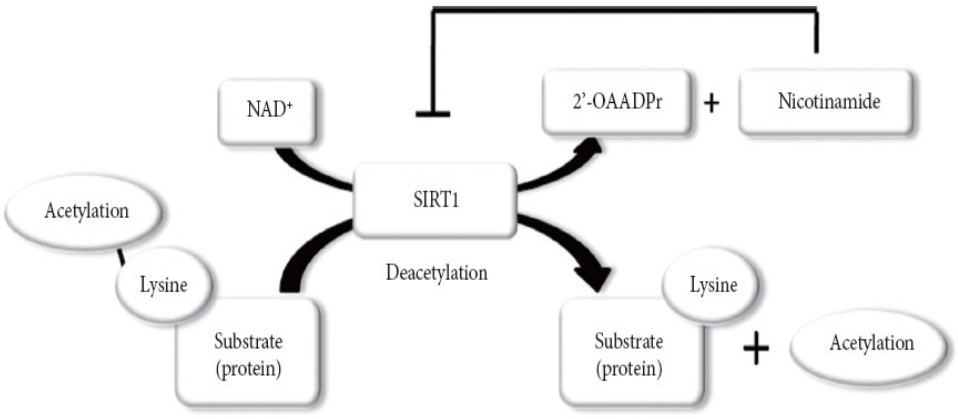


Figure 2.4: Schematic diagram of enzymatic activities of SIRT1.

Various nonhistone proteins associated with a wide variety of cellular processes such as transcription factors, transcriptional coregulatory proteins and histones serve as substrates for SIRT1. The biological functions of SIRT1 are given in Table 2.2.

Table 2.2: Biological functions of sirtuin 1

Activity	Function	Target
Metabolism	Glucose metabolism	PGC-1α, IRS2, PTP1B, UCP2, LKB1
	Lipid metabolism	PGC-1α, PPAR-α, SREBP, LXR, FXR
	Mitochondrial biogenesis	PGC-1α
	Autophagy	Atg5, Atg7, LC3, FOXOs
	Inflammation	NF-κB (p65)
	Circadian rhythms	BMAL1, PER2
Others	Apoptosis	FOXOs, p53, Smad7
	Stress resistance	FOXOs, PARP1, HIF
	Chromatin silencing	H3K9,H3K14, H4K16, H1K26

PGC, peroxisome proliferator activated receptor-γ coactivator; IRS, insulin receptor substrate; PTP1B, protein tyrosine phosphatase 1B; UCP, uncoupling protein; LKB, liver kinase B; PPAR, peroxisome proliferator activated receptor; SREBP, sterol regulatory element binding protein; LXR, liver X receptor; FXR, farnesoid X receptor; Atg, autophagy-related gene; LC3, light chain 3; FOXO, forkhead box O; NF-κB, nuclear factor-κB; BMAL, brain and muscle aryl hydrocarbon receptor nuclear translocator-like; PER2, period 2; PARP, poly-ADP-ribose polymerase; HIF, hypoxia inducible factor.

SIRT1 plays a crucial role in a variety of processes such as regulation of insulin secretion and β-cell protection, repression of the inflammation, and regulation of insulin signaling, mitochondrial biogenesis and subsequent reactive oxygen species (ROS) generation, adipogenesis, adiponectin secretion, hepatic glucose/lipid metabolism, and circadian rhythms. Additionally, SIRT1 may also improve insulin resistance and diabetic status. The role of SIRT1 on glucose/lipid metabolism in

relation to type 2 diabetes mellitus is summarized in Table 2.3.

Table 2.3: The role of SIRT1 on glucose/lipid metabolism in relation to type 2 diabetes mellitus

Pancreas	Insulin secretion ↑ β-Cell protection ↑
Insulin signaling	Insulin sensitivity ↑
Inflammation	Insulin sensitivity ↑
Adipose tissue	Lipid mobilization↑ Adiponectin↑
Skeletal muscle	Mitochondria biogenesis ↑ Glucose uptake ↑
Mitochondria	Biogenesis ↑ ROS ↓ Fatty acid oxidation ↑ Glucose/Lipid metabolism
Liver	Glucose production Fatty acid oxidation ↑
Circadian rhythm	Glucose/Lipid metabolism

Thus, it is noteworthy to mention that SIRT1 might be a pharmacological therapeutic target to treat insulin resistance and T2D [347]. The acquaintance of sirtuins has got extension from the original description of a NAD+-dependent deacetylase that was responsible for longevity in yeast and associated with CR. Sirtuin1 (SIRT1) as described above and other sirtuin family members such as SIRT2, 3, and 6, may also induce beneficial effects in glucose metabolism, partially through improving inflammation, oxidative stress and maintaining mitochondrial function. Therefore, modulation of sirtuins pharmacologically may represent a novel therapeutic approach for improvement of insulin resistance and T2DM. The antidiabetic effects of several SIRT1 activators including resveratrol and synthesized activators, in animal models have been evaluated [348].

Several small trials in humans have revealed that SIRT1 activators exert beneficial effects on glucose metabolism and insulin resistance resembling to the CR effect [349]. There is still lack of sufficient clinical data pertaining to the effect of SIRT1 activators on insulin resistance and T2DM. Additionally, other CR induced sirtuins such as SIRT2, SIRT3 and SIRT6 play critical roles in regulation of cellular processes such as metabolism, inflammation, oxidative stress and mitochondrial function. For the development of new strategies to treat insulin resistance and T2DM

further investigation into the targets and functions of sirtuins SIRT1, SIRT2, SIRT3 and SIRT6 is desired. Other sirtuins family members SIRT4, SIRT5 and SIRT7 in addition to SIRT1, SIRT2, SIRT3 and SIRT6 play pivotal roles in cellular homeostasis and functions, including redox homeostasis, anti-inflammation, cell survival, and mitochondrial quality control [350-355] that may be engaged in the pathogenesis of insulin resistance and T2DM. For the elucidation of the detailed molecular mechanisms further basic studies are necessary.

3.5. THE SODIUM GLUCOSE COTRANSPORTER 2 (SGLT2) INHIBITORS

Sodium glucose cotransporter 2 (SGLT2) inhibitors are a novel class of FDA approved prescription drugs to lower blood glucose levels along with diet and exercise for type 2 diabetic patients. At present these are not approved for use in type 1 diabetic patients. The mechanistic aspect of these drugs is the inhibition of SGLT2 in the early proximal tubule of the kidneys [356, 357]. The reabsorption of glucose, which has been filtered by the glomeruli of the kidneys, is the foremost function of this SGLT2. It accounts for nearly 97% of renal glucose reabsorption and remaining by SGLT1 located in the downstream late proximal tubule, in normoglycemic conditions, in such a manner that urine is nearly free of glucose in healthy individuals [357]. In hyperglycemic patients the inhibition of SGLT2 results in to increase in glucosuria and decline in serum glucose levels.

The effect is more pronounced in the setting of hyperglycemia as the latter increases the filtered load of glucose to the proximal tubule and enhances glucose reabsorption via SGLT2 and as a consequence leads the glucosuric effect of SGLT2 inhibition. In diabetic patients, this glucosuric effect may further increase because of a diabetes-associated increase in renal SGLT2 expression. However, it is a debatable matter due to availability of positive and negative data [357-360]. In diabetes, the renal glucose reabsorption via SGLT2 increases that is contributing to maintain hyperglycemia whereas, the inhibition of SGLT2 opposes these effects.

Phlorizin, a flavonoid contained in the bark and fruit of fruit trees, discovered over 100 years ago was the first SGLT2 inhibitor. It is a nonspecific SGLT inhibitor which inhibits both SGLT2 and SGLT1 but SGLT2 with a tenfold higher affinity. It

is worth mentioning and in contrast to SGLT2 that SGLT1 have expressions in various body tissues, in addition to the renal tubules, such as the small intestines where SGLT1 is expressed in the luminal membrane and responsible for glucose reabsorption [361]. Consequently, inhibition of SGLT1 with phlorizin may produce extrarenal side effects like diarrhea. Phlorizin derivates have been developed as more specific inhibitors of SGLT2 to avoid SGLT1-dependent side effects. Dapagliflozin (Forxiga™/Farxiga™), canagliflozin (Invokana™), ertugliflozin (Steglatro) and empagliflozin (Jardiance™) are the phlorizin derivates which have been approved by FDA for the treatment type 2 diabetic patients in the USA. These derivatives have shown an expected HbA_{1c}-lowering effect of 0.7–0.8% from a baseline of around 8.0% [356, 357, 362]. Under normal circumstances, in the late proximal tubule SGLT1 is mostly inactive because of the upstream reabsorption of filtered glucose via SGLT2 that permits very little glucose to pass by, and only a small fraction of the glucose transport capacity of SGLT1 is active.

The reabsorptive capacity of SGLT1 has been unmasked by the increased glucose load to the late proximal tubule by SGLT2 inhibition. Consequently, when SGLT2 is inhibited in euglycemic conditions the renal glucose reabsorption remains at around 40-50% of filtered glucose and with preserved glomerular filtration rate (GFR) is due only to SGLT1 [357]. SGLT2 become ineffective once the filtered load falls below the transport capacity of SGLT1 (nearly 80 g/day) as the glucosuric effect of SGLT2 inhibitors is coupled to the filtered load of glucose. Hence, SGLT2 inhibitors are not FDA approved for use in type 2 diabetic patients having severely reduced GFR. Dapagliflozin, empagliflozin and canagliflozin are FDA approved for use in T2D patients. The hepatic gluconeogenesis is enhanced due to the counter-regulatory mechanisms triggered by SGLT2 inhibitors that prevent a further reduction in blood glucose levels and increases in glucagon levels. SGLT2 inhibitors do not cause hypoglycaemia as they do not stimulate insulin secretion or action and their effect ebbs as blood glucose levels fall [363-365]. The potential mechanism contributing to the protective effects of these compounds on the renal and cardiovascular system is lowering of blood glucose without increasing the risk of

hypoglycemia [366].

SGLT2 Inhibitors and T2DM

The beneficial cardiovascular and renal effects of the SGLT2 inhibitors in type 2 diabetic patients were shown in the EMPA-REG OUTCOME trial [367]. In this trail, the effects of empagliflozin on cardiovascular and renal outcomes were 7020 high cardiovascular risk type 2 diabetic patients with an estimated GFR (eGFR) of $\geq$ 30 mL/min^{-1} (1.73 m^2)$^{-1}$. Empagliflozin, using angiotensin converting enzyme (ACE) inhibitor or angiotensin II receptor blocker (ARB) as standard of care therapy reduced the rate of nephropathy. The rate of nephropathy was defined as progression to macroalbuminuria, doubling of serum creatinine, initiation of renal replacement therapy or death from renal disease with a relative risk reduction of 39%. It also reduced the rate of doubling of creatinine by 44% and progression to end stage renal disease by [367]. The SGLT2 inhibitor also reduced the rate of death from cardiovascular disease by 38, hospitalization for heart failure by 35 and death from any cause by 32%, in addition to renal benefits [368].

The second trial was the CANVAS program which involved 10142 type 2 diabetic patients and that showed significant cardiovascular and renal benefits of an SGLT2 inhibitor. In this case canagliflozin was compared to placebo. The composite outcome of sustained reduction in eGFR, the need for renal replacement therapy or death from renal causes as renal benefits were less in canagliflozin group than the placebo group with a hazard ratio of 0.6, similar to the EMPA-REG-OUTCOME trial. In both the trials the rate of heart failure was significantly lowered. In terms of cardiovascular death, there were no significant differences in the canagliflozin and placebo groups [369].

The primary outcome in both trials was a combination of death from cardiovascular causes, nonfatal myocardial infarction or nonfatal stroke. The relative reduction of risk, in both trials, of this outcome between the treatment and placebo arms was significant at 14%. The difference between these two trials in terms of cardiovascular death may curtail from the fact that the EMPA-REG OUTCOME trial included a higher prevalence (99%) of cardiovascular disease at baseline compared to

the CANVAS program (65%). As far as the adverse effects concerned the main difference between the two trials was the increased risk of amputations seen with canagliflozin. Canagliflozin taking patients had nearly 2 times higher risk of demanding an amputation, mainly toe and metatarsal, compared to placebo [369]. Further studies are needed to better understand this issue as it was not reported in previous trials involving canagliflozin.

SGLT2 Inhibitors and T1DM

The progressive loss of pancreatic islet β-cells via an autoimmune mechanism results in insulin deficiency and ultimately hyperglycemia in type 1 diabetic patients. Thus, due to insulin deficiency, intake of insulin is a required part of the treatment regimen for type 1 diabetic patients. The recommended glycemic control goal (HbA1c<7.0%) by the American Diabetes Association could not be achieved by nearly 75% of type 1 diabetes adults [370]. The hypoglycemia caused by insulin can lead to death and in the long term enhance cardiovascular risk in these patients [371].

Pramlintide is the only FDA approved non-insulin drug for use in type 1 diabetic patients and it lowers the glucose levels by inhibiting glucagon secretion [372]. Efficient therapies for type 1 diabetic patients are strongly desired due to a high cardiovascular risk [373]. SGLT2 inhibitors may provide an attractive addition to the typical insulin-only regimens prescribed for poorly controlled T1D patients who are compliant with their insulin therapy and suffer from frequent episodes of hypoglycemia. To assess the efficacy SGLT2 inhibitors in T1D patients, three prospective, well powered, doubleblind, placebo-controlled trials have been completed and published.

Henry and colleagues performed the first trial consisting of 351 patients with type 1 diabetes, randomized into three groups receiving daily doses of 100 or 300 mg of canagliflozin or placebo. The primary endpoint was the proportion of patients who achieved HbA1c reduction from baseline of more than 0.4% and no weight gain. In this 18-week trial, significantly more patients in the 100 and 300 mg canagliflozin groups achieved goals as compared to placebo (36.9 and 41.4 vs 14.5%, respectively; $p<0.001$). Furthermore, both the doses of canagliflozin reduced HbA1c, body weight

and the total required insulin dose significantly as compared to placebo [374].

In the second study, the DEPICT1 trial, 833 patients were randomized into three different treatment arms: doses of 5 and 10 mg/day of dapagliflozin and were compared with placebo. After 24 weeks of treatment, both doses of dapagliflozin significantly and to a similar extent reduced HbA1c levels as compared to placebo. This study also revealed that total daily insulin dose and body weight were significantly reduced in the treatment arms as compared to placebo [375].

The biggest phase 3 trial comprising 1402 patients assessing the efficacy and safety of an SGLT2 inhibitor in type 1 diabetic patients to date was the inTandem3 trial and was published in September 2017 [376]. In this trail, patients were assigned a combination of SGLT2 and SGLT1 inhibitor to insulin therapy plus placebo versus insulin therapy plus sotagliflozin (400 mg per day) for 24 weeks. 28.6% of patients in sotagliflozin group and 15.2% of patients in the placebo group met the primary end point. Additionally, sotagliflozin group patients achieved a significant lowering in HbA1c, systolic blood pressure and body weight and less daily insulin compared to the placebo group [376].

The selectivity for SGLT2 over SGLT1 for sotagliflozin, dapagliflozin, canagliflozin and empagliflozin is approximately 20:1, 1167:1, 263:1 and 2667:1, respectively [377, 378]. The oral application of sotagliflozin decreases glucose absorption by SGLT1 inhibition in the small intestine that causes postprandially glucose lowering which is the added benefit of sotagliflozin compared to the more selective SGLT2 inhibitors. The SGLT1 inhibition, in the small intestine, might also induce a sustained postprandial increase in glucagon like peptide 1 (GLP1) which may elevate glucose-dependent insulin secretion in T2D [361]. In the absence of endogenous insulin secretion this effect becomes inappropriate. The inTandem trial revealed that patients reporting diarrhea were double in the sotagliflozin group compared to the placebo group [376] and this risk was lower than expected. Sotagliflozin inhibits SGLT2 after reabsorption into the systemic circulation in kidney. It is still unclear weether the tubular sotagliflozin concentrations following its oral application are high enough to inhibit SGLT1 in the kidneys. It would be

advantageous to know for better understanding the role of the kidney in the potential differences in side effects between highly and less selective SGLT2 inhibitors.

SGLT2 Inhibitors in T1DM: Safety Aspects

The T1D patients are more prone to DKA than T2D patients, so the use of SGLT2 inhibitors in the T1D population is of particular concern. It is still not fully understood that in what manner SGLT2 inhibitors increase ketone levels in the serum and induce DKA. According to one proposed hypothesis that due to glucosuric and blood-glucose lowering effects of SGLT2, endogenous insulin release decreases and glucagon levels increases [379]. In type 1 diabetic patients, insulin dose is lowered as drugs cannot reduce endogenous insulin levels. The enhanced lipolysis caused by the changes in insulin and glucagon releases more free-fatty acids from adipose tissue and these are then used for ketogenesis by the liver. At low blood glucose levels, the ketone bodies are released into the systemic circulation to provide an alternative energy substrate. When the level of plasma ketone bodies is high, the facilitated renal retention of ketone bodies by lowering GFR by SGLT2 inhibitors reduce the filtered amount of ketone bodies below the renal tubular reabsorption capacity [380]. Therefore, in the absence of hyperglycemia SGLT2 inhibition causes DKA and increased ketonemia [379, 380]. The low basal endogenous insulin levels increase the risk of DKA in type 1 diabetes. A number of potential ketoacidosis triggers major illness, reduced food and fluid intake, concomitant mild infection, increased physical activity and/or reduced food intake and acute insulin dose reduction or omission have been identified in type 1 diabetes patients. On the other hand, in some cases there were no identified contributing factors [380, 381]. Based on the adverse events of ketoacidosis reported to the FDA related to the use of SGLT2 inhibitors, the FDA revised the label on SGLT2 inhibitors in May 2015 that these inhibitors can potentially cause DKA and patients should stop taking the drug if DKA is diagnosed [382-384].

In the early phase of type 1 and 2 diabetes, glomerular hyperfiltration is the proposed risk factor for the later development of albuminuria and diabetic nephropathy [385, 386]. The proportions of hyperfiltration cases are higher in type 1

diabetic patients (nearly 75%) as compared to type 2 [385]. In type 2 diabetic patients, the proportions of hyperfiltration is more variable because of the difference in glycemic control, duration of diabetes, age and GFR measurement method used for obese patients [385]. At moderate levels, hyperglycemia induces a 'primary' increase in proximal tubular reabsorption by providing substrate for SGLTs or by causing the tubule to undergo hypertrophy [387].

The reduction in diabetic hyperfiltration, using selective and non-selective SGLT2 inhibitors has been shown in preclinical studies. Diabetic hyperfiltration was first shown by Vallon *et al.* [388] in 1999 in micropuncture studies in rats using local application of phlorizin into Bowman's space and latterly by acute or chronic systemic application of selective SGLT2 inhibitors [389]. The inhibition of SGLT2, pharmacologically or genetically, suppressed the hyperfiltration on the whole-kidney level in mouse models of diabetes [390]. It is also found that in each case diabetic hyperfiltration suppression was not dependent on blood glucose, however, associated with an increase in NaCl concentration at the macula densa [391] and in hydrostatic pressure in Bowman's space. In addition to this, SGLT2 inhibitors can reduce renal growth, albuminuria and inflammation mainly through their glucose-lowering effect similar to as observed in a genetic rodent model of type 1 diabetes [392]. For the rodent models of type 2 diabetes, there were similar results [357].

The clinical investigations carried out by Cherney and colleagues [393] are in consistency with the assumption that SGLT2 inhibitors lower hyperfiltration by attenuating an increased tone of SGLT2-mediated tubular hyperreabsorption. It is important to know other mechanism that may be added to the cardiovascular benefits of SGLT2 inhibitors in type 1 and type 2 diabetic patients besides lowering blood glucose levels with minimum risk of hypoglycemia and preserving kidney function. Taking in to consideration the prominent benefits related to heart failure, one explanation might be the reduction in blood pressure seen with these SGLT2 inhibitors in addition to body fat and weight loss. Sotagliflozin has lowered the blood pressure in type 1 diabetic patients [376] and all the three trials in type 1 diabetic patients have shown significant weight loss compared to placebo [374-376]. The

glucose based osmotic diuresis (100-470 mL/day), natriuresis and weight loss are the underlying mechanism of this blood pressure-lowering effect [394, 395]. The sustained weight loss may be assumed primarily due to increased lipolysis resulting to a decrease in fat body content.

In SGLT2 inhibitors prescribed patients, because of the reduction in glucose, there is a primary shift from carbohydrate utilization to lipids which in turn leads eventual lipolysis and weight loss. The studies in rodents and clinical studies on type 2 diabetic patients have revealed it [396, 397]. An early proximal tubule transporter Na^+/H^+-exchanger 3 (NHE3) has co-expressions with SGLT2 therefore, SGLT2 inhibition may also inhibit NHE3 as proposed recently [357, 398]. The blood pressure lowering effect of SGLT2 may be attributed to the interaction between SGLT2 and NHE3. On the other hand, this interaction may moderately impair renal acid excretion. According to another potential mechanism which correlates to the fact that inhibition of SGLT2 shifts more glucose transport to SGLT1 in the late proximal tubule which in turn may reduce the oxygen tension in the outer medulla. It leads to enhanced erythropoietin release and red blood cell production that together with the diuretic effect, increases hematocrit and might facilitate oxygen delivery to the kidney and the heart [399]. There was also a small but statistically significant increase in hematocrit in type 1 diabetic patients treated with the SGLT2 inhibitor empagliflozin [393]. Thus, increase in ketogenesis caused by SGLT2 inhibitors may result detrimental effects in the form of DKA. On the other hand, additional energy substrates in the form of ketone bodies for the heart and kidney potentially provided by mild ketosis may be organ protective [400]. In this regard, more studies are desirable for better understanding of these issues and potential mechanisms and that may be applicable to both type 1 and type 2 diabetic patients. In summary, the pleiotropic effects of SGLT2 inhibitors are depicted in Figure 2.5.

The mechanisms that are currently assumed to contribute to the protective effect of SGLT2 inhibitors are also expected to take place in type 1 diabetic patients (Figure 2.5).

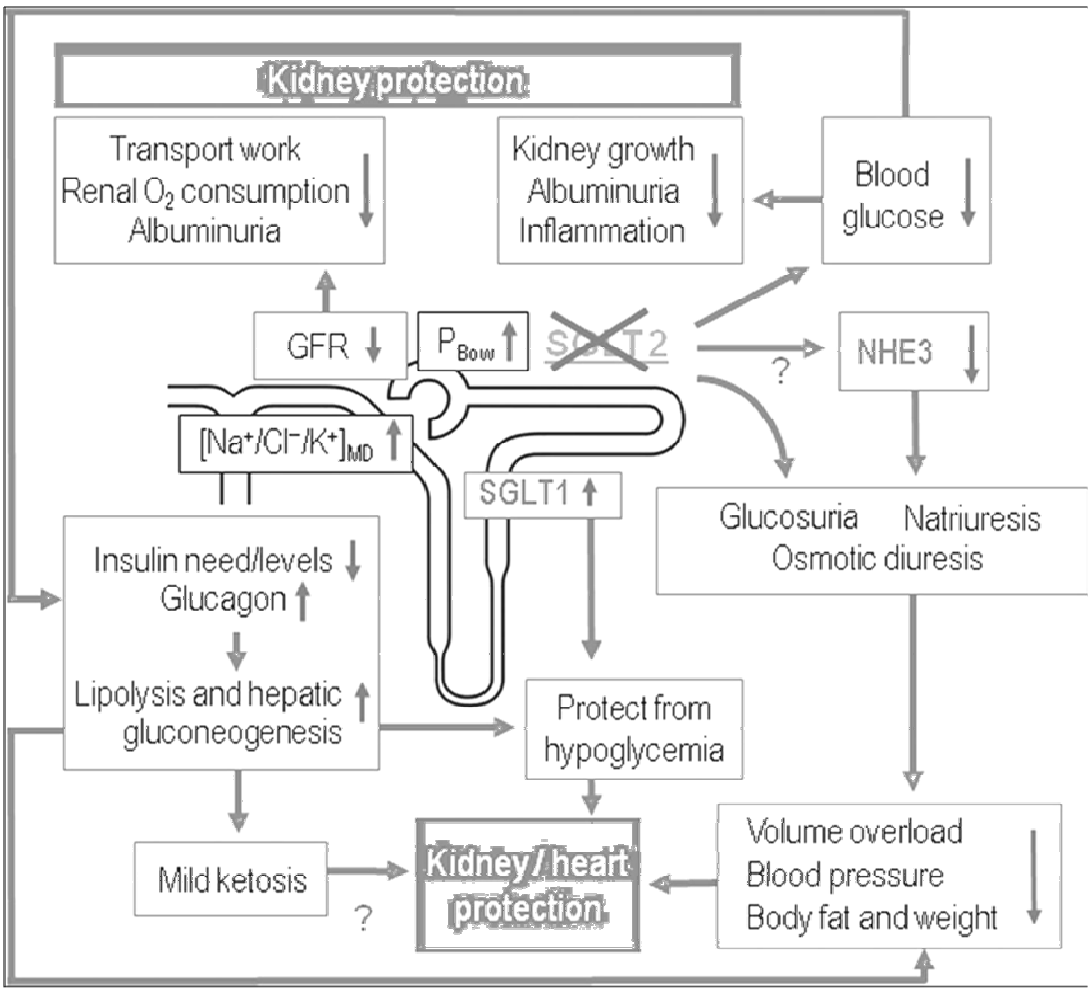


Figure 2.5: Proposed mechanisms of kidney and heart protection induced by SGLT2 inhibition in both T1 and T2D patients.

The assessment of the risk-benefit relationship will be the key aspect for the use and potential approval of SGLT2 inhibitors in type 1 diabetic patients in near future. Do the enhancements in HbA1c levels and the prospective favorable effects on the kidney and heart prevail over the risks of diabetic ketoacidosis that looks as the most serious and adverse happening identified in type 1 diabetic patients.

It was shown in previous studies that SGLT2 inhibitors improve death from cardiovascular causes, hospitalization for heart failure and death from any cause in type 2 diabetic patients. By far the effects of SGLT2 inhibitors on cardiovascular

results have not yet been evaluated in type 1 diabetic patients and it requires a long-term dedicated trials.

It may be summarized that SGLT2 inhibitors are effective glucose-lowering drugs in addition to insulin in type 1 diabetic patients and it was shown by high-powered prospective, double-blind and placebo-controlled clinical trials. Additionally, SGLT2 inhibitors also have the potential to provide renal and cardioprotective benefits to type 1 diabetic patients by reducing blood glucose levels with low hypoglycemia risk, reduction in glomerular hyperfiltration, decrease in blood pressure and volume overload as well as weight loss. Although the inhibition of SGLT2 increases ketogenesis which may lead to DKA in susceptible type 1 diabetic patient particularly and in the presence of precipitating factors such as volume depletion. Furthermore, long-term trials and studies are desirable to better understand how to prevent DKA episodes in these patients, if the dual inhibition of SGLT2 and SGLT1 has any additional value to reveal whether the renal and cardiovascular benefits of SGLT2 inhibitors exposed in type 2 diabetic patients also happen in type 1 diabetic patients and to determine whether these effects compensate the risk and danger of DKA.

3.6. THE G PROTEIN-COUPLED RECEPTOR 40 (GPR40) AGONISTS

The antidiabetic drugs, in short term, have proved to be very effective in improved management of patients' blood glucose levels. Due to the progressive nature of the disease and the unavoidable worsening of pancreatic beta-cell function, the glucose-lowering effects of these agents are not sustained for the long time. Thus, intense research efforts have been made on the discovery of novel therapeutic drugs which can preserve beta-cell function, restore metabolic homeostasis and ameliorate T2DM in a sustainable manner.

Free fatty acids (FFAs), being the structural components of biological membranes, are of great physiological importance to human body and these are important source of energy as well. As a biologically active molecule FFAs exert a wide variety of functions. The participation of FFAs in the regulation of metabolic homeostasis contributes in the development of many metabolic diseases such as

T2DM, obesity, and atherosclerosis. It was revealed in the deorphanization of several G-protein coupled receptors (GPCRs) that GPR40, GPR41, GPR43, GPR84 and GPR120 act as receptors for extracellular FFAs. These were having various carbon chain lengths and mediated a number of their physiological actions [401-405]. The activation of these receptors was dependent on the carbon chain length. GPR40 is activated by medium-chain and long-chain [406] and highly expressed in pancreatic β-cells and participates in the induction of glucose dependent insulin secretion (GDIS) by FFAs. As a result, it has received considerable attention as a potential therapeutic target for the management of T2DM [407-414].

Role in metabolic homeostasis

GPR40, also known as FFA receptor 1, is having expressions in both human and rodent tissues [409, 412] and belonging to the A class of GPCRs that are characterized by a seven-transmembrane domain structure spanning α-helices with three hydrophilic intracellular and three hydrophilic extracellular loops. Besides the high levels of expression in pancreatic beta cells, GPR40 is also expressed, although to a lesser extent, in other tissues such as the intestinal tract, brain and in monocytes [410, 412]. It can be activated by medium-chain and long-chain fatty acids, either saturated or unsaturated, in a dose-dependent manner [414, 417]. The promotion of the induction of GDIS by FFAs in pancreatic beta cells is the most important and well-documented function of GPR40 [416]. Studies suggested that among the FFA receptors, GPR40 is the primary mediator of this effect [415-417]. Furthermore, the role of GPR40 in insulin secretion has been established by using receptor antagonists such as GW1100 and that was shown to inhibit GPR40-mediated augmentation of insulin secretion from MIN6 cells [418]. A natural variant of GPR40 (Gly180Ser) blocks the sensing ability of β-cells to lipids and impairs fatty acid induced insulin secretion from pancreatic β-cells [408]. In this way, there is a demonstrated importance of GPR40 in FFA induced augmentation of insulin secretion from β-cells [419]. In the gastrointestinal tract, GPR40 is expressed in enteroendocrine cells including those which secrete the incretin hormones glucagon like peptide 1 (GLP-1)

and gastric inhibitor peptide (GIP). The incretin hormones GLP-1 and GIP are secreted from the gut upon ingestion of nutrients and stimulation of GDIS from pancreatic β-cells is their primary function. The secretion of these hormones is regulated by the activation of GPR40 [410, 420, 421]. It has also been reported that activation of GPR40 may enhance the secretion of GLP-1 in primary human colonic cultures [422]. Thus, GPR40 in addition to directly increasing GDIS from pancreatic beta cells, has an indirect effect on GDIS through the potentiation of GLP-1 from enteroendocrine cells of the gastrointestinal tract. The insulin release from pancreatic β-cells mediated through GPR40 is represented in Figure 2.6.

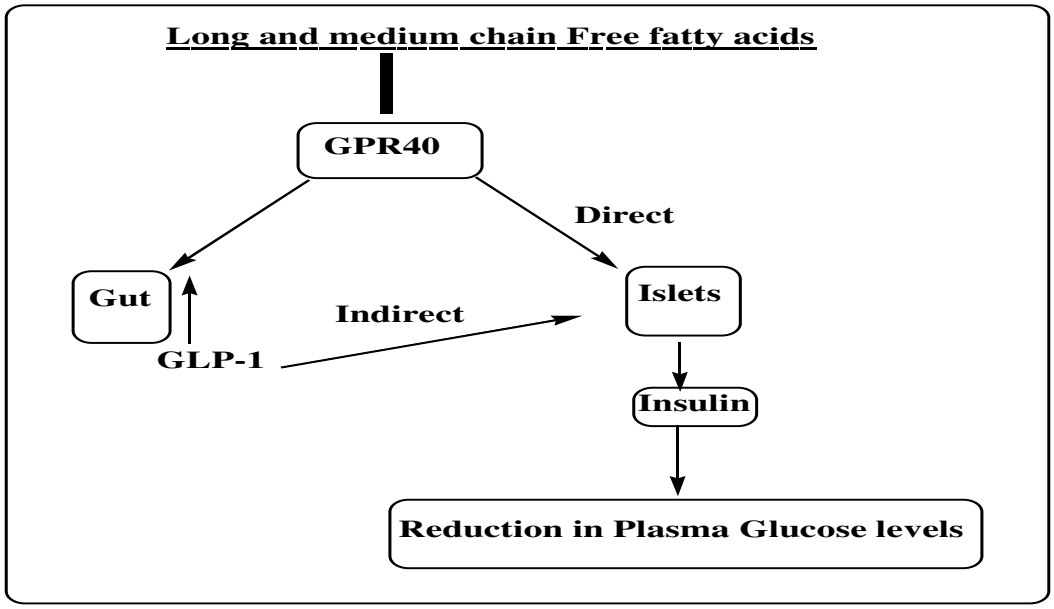


Figure 2.6: GPR40 mediated insulin release from pancreatic β-cells.

GPR40 is ubiquitously expressed in various regions of human brain where it may facilitate a number of important physiological functions [404, 423-427] and the physiological role of GPR40 in brain has been reported in the literature [428-431]. It has implications in mediating trans-arachidonic acid induced neuro-microvascular degeneration in rat pups [432]. GPR40 may also be expressed in the tissues which are sensitive to insulin including liver, muscle and white adipose [433-441] but there is still no sufficient knowledge about its role in such tissues. The involvement in

potentiating insulin signaling in human and chicken hepatocytes where it has expressions has been reported [442-444]. To establish the function of GPR40 in insulin-sensitive tissues that are inevitable to preserve the metabolic health exhaustive studies are desirable.

Emergence of GPR40 agonists

The fact that FFA-induced augmentation of insulin secretion from beta cells is mediated by GPR40 proved helpful in developing agonists of potential therapeutic value in T2DM. These agonists mimic the FFAs structurally, having the acidic head group and the hydrophobic tail. The low bioavailability and vulnerability to beta-oxidation was the main concern in the compounds of first generation. Extensive efforts have been made to design novel GPR40 agonists by optimization of structures so that potential for beta-oxidation may be reduced, improve bioavailability keeping safety aspects in mind [407, 445-447]. As a result, a large number of GPR40 agonists have been synthesized and were tested (Table 2.4). Many of these appeared as to reiterate the actions of FFAs on pancreatic insulin secretion insulin secretion [418, 448-450]. In various rodent models of T2DM, the administration of these molecules improve glucose tolerance and restore metabolic homeostasis, by augmenting insulin secretion from pancreatic beta cells [445, 446, 451-456]. The delayed onset of fasting hyperglycemia in Zucker diabetic fatty (ZDF) rats through increased insulin secretion and preservation of beta cell integrity has also shown by GPR40 agonists [455, 457]. TAK-875 and AMG 837 are the representative GPR40 agonists which reached clinical trials and it was shown that TAK-875 improves glycemic control in Type 2 diabetic patients.

Table 2.4: GPR 40 agonists and their physiological actions

Agonist	Physiological Actions	References
AM-1638	Improved Glycemic Control In T2DM mice Stimulation of incretin (GIP and GLP-1) secretion from enteroendocrine cells	[454, 460]
AMG 837	Improved Glycemic Control In T2DM mice Stimulation of Insulin secretion from MIN6 cells Reduced plasma glucose and HbA1C levels in *ob/ob*	[445-448, 461]

	mice	
AS2034178	Normalized Glycemic Control and improved β-cell function in ZDF rats	[455]
AS2575959	Augmented GDIS in vitro and improved OGTT in diabetic mice	[451]
GW9508	Augmented GDIS in vitro, Iimproved glycemic control in HFD mice Inhibited lipopolysaccharide-induced interleukin-6 secretion	[452, 462, 463]
TUG-424	Reduction of hepatic lipid accumulation in HFD mice Improved Glucose tolerance in mice	[464]
TUG-469	Augmented GDIS in INS-1 cells Improves glucose tolerance in pre-diabetic NZO mice Antagonized palmitate-induced β-cell death.	[465-467]
TUG-770	Improved Glucose tolerance in HFD mice	
TAK-875	Augmented GDIS Improved Glycemic control in rodent and human T2DM	[407, 457, 468-473]

However, phase III clinical trials with TAK-875 were terminated recently due to concerns of liver toxicity [458]. The nature of the liver-related toxicity using TAK-875 is not clear and also that is it specific to TAK-875 or occurs with other agonists. GPR40 regulates metabolic homeostasis by potentiating GDIS from pancreatic beta cells. The role in the metabolic regulation of GPR40 and as a potential therapeutic target for T2DM is evident from the cell culture studies and rodent models of T2DM. No any study provided the evidence of hypoglycemic risk by GPR40 activation, rendering it as an attractive therapeutic approach. Further investigational studies must be aimed to have detailed insight on extra-pancreatic functions of GPR40, especially in those tissues that take part in maintaining metabolic homeostasis. Srivastava *et al.* has reported a high resolution structure of human GPR40 bound to TAK-875 [459]. The structural information, derived from the drug-receptor complex, may provide insight into lipid entry of the ligand and binding mode to receptor to amplify the agonist signal. It would be the potential useful in developing newer and more effective GPR40 agonists which might serve as efficient anti-diabetic agents devoid of any toxic side effects.

3.7. THE PEROXISOME PROLIFERATOR-ACTIVATED RECEPTORS (PPARs)

An autoimmune disease T1D is caused by the dysfunction of pancreatic beta or by killing autoreactive T cells and resulting in to reduced insulin production and hyperglycemia [474, 475]. The increasing incidence of T1D varies geographically. Studies revealed that environmental factors including diet and microorganisms play a pivotal role in the pathology of T1D [476, 477]. It was supposed that there was almost complete loss of beta cells at the inception of the disease but recent studies have shown retention of insulin-positive islets, up to 40%, in new-onset patients of T1D [478-480]. Additionally, the isolated islets may regain their ability to secrete insulin when cultured *in vitro* in a nondiabetogenic environment [481]. Therefore, dysfunction of beta cell may play an important part in the pathology of T1D. The therapeutic approaches being used currently in T1D are having limited clinical efficacy and are mainly focused on to suppress the ongoing immune attack or to stimulate beta cell regeneration [482, 483]. Therefore, there is highly need of such strategies that might dampen the immune response and promote beta cell function. PPARs having anti-inflammatory properties both regulate beta cell biology and modulate the pancreatic lipidome hence may serve as an ideal target for such a strategy.

PPARs, mediators of peroxisome proliferation were identified in the 1990s [484] belonging to the nuclear receptor class II superfamily of transcription factors and regulate a wide variety of biological processes through modulating gene expression. Three isoforms, namely PPARα (NR1C1), PPARβ/δ (NR1C2), and PPARγ (NR1C3) have been identified in mammals that control predominately genes involved in lipid metabolism such as transport, storage, lipogenesis and fatty acid oxidation (FAO) [484]. PPARs are vital targets for metabolic disorders and multiple drugs targeting PPARα, which include fibrates (e.g. fenofibrate, bezafibrate and clofibrate) and PPARγ, which include thiazolidinediones (e.g. troglitazone, rosiglitazone, pioglitazone and ciglitazone) are being used for the treatment of hyperlipidemia and type 2 diabetes. Although, PPARs are present in the nucleus

constitutively but are dynamic as they shuttle between the nucleus and cytoplasm [485, 486] and this shuttling is regulated by binding of PPAR ligands to the C-terminal domain. The induced conformational change, upon binding of ligands, leads to heterodimerization with members of the retinoid X receptor (RXR) family [487, 488]. This complex binds to specific DNA sequences, termed as peroxisome proliferator response elements (PPRE) via the highly conserved zinc finger DNA-binding domain in the N terminus [489]. Upon binding of ligands, target gene transcription increases due to dissociation of corepressors and recruitment of coactivator proteins [490]. In the absence of ligands, PPARs as an alternative recruit corepressors which repress transcription of target genes [491].

PPARs are involved in a ligand-dependent but PPRE-independent mechanism of gene repressions, termed as "transrepression" via interactions with other proteins such as NFκB, AP1, and STAT [492–494]. This mechanism generates and stabilizes corepressing complexes, which typically bind to and repress proinflammatory genes [488]. The pattern of expression of the PPAR isoforms differ although, they have a high structural and functional overlap. The isoform PPARα is highly expressed in tissues such as liver, kidney and adipose that are active metabolically. It is activated during fasting and is involved in controlling many physiological processes such as ketogenesis, lipoproteins, gluconeogenesis, amino acid catabolism, FAO and inflammatory responses [495].

PPARβ/δ expressed ubiquitously and involved in FAO and its activation exerts an anti-inflammatory effect with reduced secretion of proinflammatory cytokines [496]. PPARγ is, having expressions in various tissues such as adipose, intestine, liver and kidney [497, 498] and involved in the regulation of fat cell differentiation, lipid storage and differentiation of monocytes into macrophages [499, 500]. Because of immune regulatory functions of PPARs these are associated to various autoimmune diseases including multiple sclerosis [501], lupus erythematosus [502], autoimmune thyroiditis [503], Graves ophthalmopathy [504], rheumatoid arthritis [505], psoriasis [506] and Guillain-Barré [507]. PPARs may also serve as targets to treat chronic inflammatory diseases [487, 508].

The role of PPAR isoforms as potent regulators of inflammation and in beta cell biology are depicted in Figure 2.7.

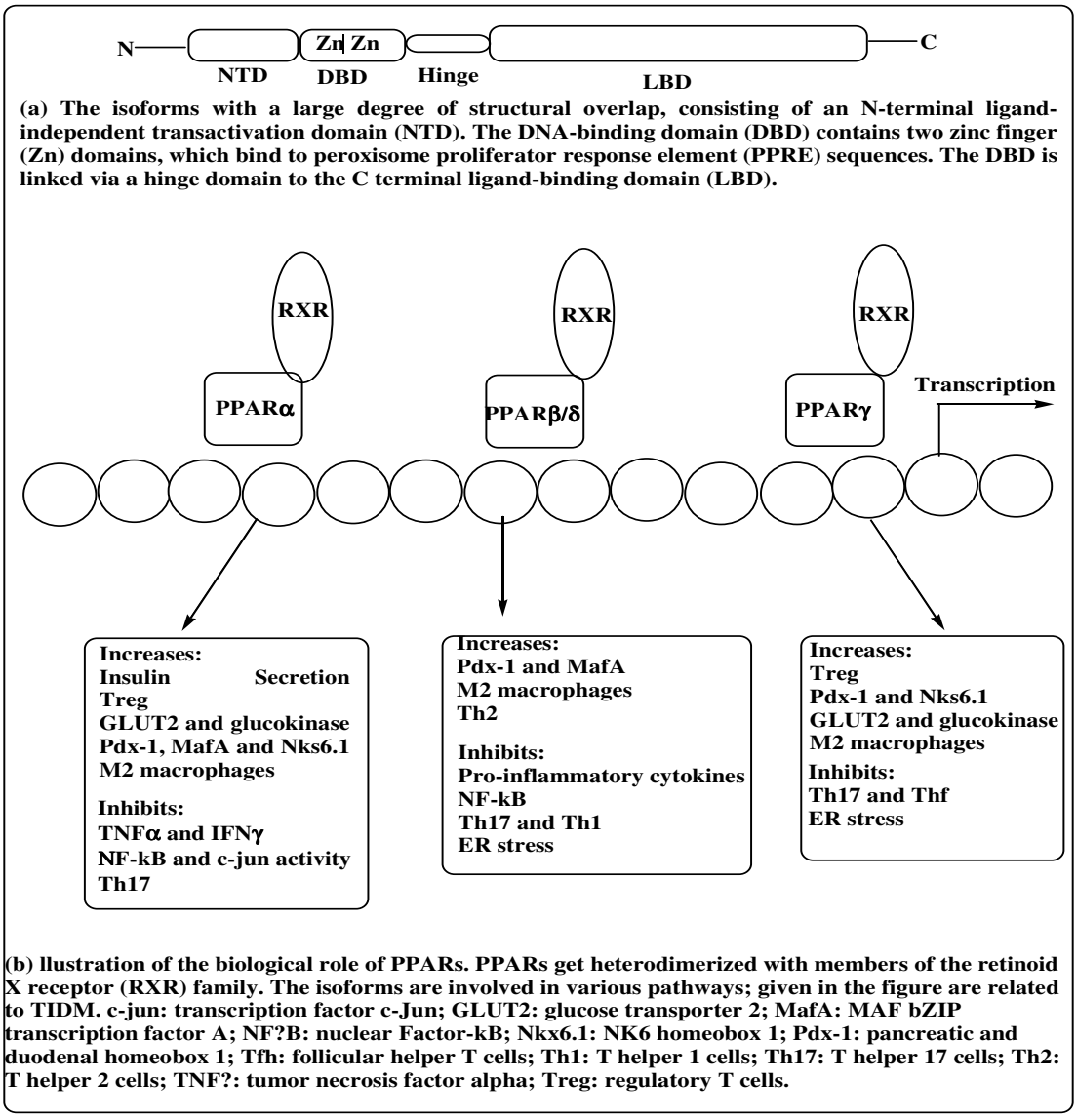


Figure 2.7: Structure and functions of PPARs.

The susceptibility of women than men to develop autoimmune diseases [509] may be linked to PPAR expression. It is found in mouse studies that male mice have higher expression of PPARα in T cells as compared to female mice and it was androgen sensitive [510]. The genetic predisposition to T1D may be attributed to the polymorphisms in PPARβ/δ and PPARγ promoter regions which affect the severity of

islet autoimmunity [511]. Moreover, PPARγ are also linked to the development of insulin resistance and type 2 diabetes [512].

PPARs and the Immune System

The pathogenesis of T1D relies on the interactions between beta cells and components of both of the innate and adaptive immune system [513]. Various immune cells have implications such as B cells and macrophages [514, 515], but T cells are primarily focused as defects in regulatory T cell (Treg) function develops T1D [475, 513]. The islet infiltrate populated predominantly with $CD8^+$ T cells followed by macrophages, $CD4^+$ T cells, B cells, and plasma cells, revealed by the studies of postmortem pancreas samples from T1D patients [516]. The metabolic pathway for the production of ATP plays an important role in the regulation of functions of immune cell. The modulation of FAO through PPARs may induce immunological changes as these are expressed in various types of immune cells, such as macrophages, dendritic cells, B cells, and T cells, and all three isoforms have anti-inflammatory activities [517].

The major contributor to T1D is beta cell dysfunction rather than beta cell death and thus, to find possibilities of restoring the beta cell function became an appealing research area. Based on the the expressions of PPAR isoforms in pancreatic islets [518-520], PPARs are possible drug targets and emerged to have key roles in regulation of beta cell biology. PPARα is expressed in pancreatic islets and its expression in beta cell lines is glucose dependent [521]. In isolated rat islets and INS-1E cells, high level of glucose represses PPARα [522]. The glucose-dependent upregulation of insulin expression might is due only to PPARα because glucose did not increase insulin expression in islets from PPARα knockout mice [523]. Although, in beta cells, PPARβ/δ present abundantly, but its role in beta cell biology is not fully recognized yet [519, 524]. It seems to have significance in pancreas development because pancreatic PPARβ/δ knockout mice showed increased number of pancreatic islets and a 2-fold increase in beta cell mass [525]. It was attributed to increased plasma insulin levels, hypoglycemia, and improved glucose tolerance, as isolated islets found to have an increased second-phase insulin secretion advocating that

PPARβ/δ in the mature pancreas, it is a negative regulator of insulin secretion.

The role of PPARγ in insulin secretion and pancreas development is not fully clear yet. It was demonstrated that there was suppression in insulin secretion and proinsulin biosynthesis by activation or by overexpression of PPARγ [526-530]. PPARγ pancreatic knockout mice were found hyperglycemic in spite of having routine pancreas morphology [531]. *In vivo* studies in rats and mice revealed that long-term treatment of rosiglitazone or troglitazone (both PPARγ agonist) maintained beta cell proliferation and prevents the age-related loss of pancreatic mass [532-534]. Troglitazone also found to prevent age-related pancreatic abnormalities and increases in fasting insulin levels [535, 536]. There are also reports which showed that PPARγ agonists improve beta cell function and prevent mitochondrial alterations and diabetes in obese mice and rats [537]. Moreover, PPARγ activation protects against cytokine-induced apoptosis [538], lipotoxicity [539] and human islet amyloid polypeptide toxicity [540, 541]. A feasible explanation of these findings at molecular level, that activation of PPARγ is associated with a reduced amount of reactive oxygen species by inhibiting iNOS through NFκB [538]. PPARγ activation also reduces islet ER stress in db/db mice and a diabetic ER stress mouse model [542, 543].

Extensive studies revealed that PPAR agonists prevent diabetes in the non-obese diabetic (NOD) mouse model of type 1 diabetes. A vast knowledge has been obtained regarding the identification of genetic and environmental risk factors as NOD mice have many autoantigens and biomarkers similar to human [544]. Experiments were carried out primarily on female NOD mice owing to nearly 80% diabetes incidence than in males [545]. Higher incidences in females may be correlated to the gender-specific changes in the PPARα and PPARγ expressions.

Female NOD mice had increased PPARα and decreased PPARγ expressions in macrophages and $CD4^+$ lymphocytes as compared to male NOD mice [546]. Not only this but NOD mice have altered PPARα and PPARγ expression in $CD4^+$ or $CD8^+$ lymphocytes and macrophages than to non-obese diabetic-resistant (NOR) mice [547]. It was demonstrated that incidence of autoimmune diabetes have been reduced by activation of PPARα by fenofibrate or PPARγ by troglitazone and

rosiglitazone [548]. The treatment with fenofibrate, initiated after the onset of disease, could even reverse the diabetes in 46% of NOD female mice [511]. Additionally, troglitazone, following streptozotocin injections, prevents hyperglycemia and reduces insulitis in mice [549].

PPARs have shown regulation by various naturally occurring agonists also and several of them had been examined for their effect on autoimmune diabetes in NOD mice including epigallocatechin [550, 551], curcumin [552, 553], cannabidiol [554, 555], omega 3 fatty acids [556] and capsaicin [557, 558] and these induced PPAR activity and protect against autoimmune diabetes in NOD mice. The stimulation of PPARα by taurine in the diet during gestation and lactation reduces incidence of diabetes development in offspring of NOD mice [559, 560]. Similarly, a gluten-free diet leading to increase PPARα and PPARγ expression [561] was found to reduce incidence of diabetes in NOD mice [562] even after the exposure of the diet exclusively *in utero* [563, 564].

The role of PPARs as regulators of lipid metabolism and inflammation, and in beta cell biology has been examined in numerous studies. However, the PPAR activation effects of PPAR activation on T cell survival, activation, and differentiation are beneficial in a T1D setting as evinced from the various studies, and studies of pancreas biology mostly conducted with relation to type 2 diabetes but to determine the precise role of PPARs in pathology of diabetes there is still scope for further extensive studies.

The promising beneficial effect on NOD mice of PPAR agonists advocated that modulation of PPARs might represent a novel treatment strategy targeting both the immune system and the pancreas.

3.8. THE PROTEIN-TYROSINE PHOSPHATASE 1B (PTP1B) INHIBITORS

In the treatment of T2DM, insulin sensitizers, such as thiazolidinediones (TZDs or glitazones) are being used as effective drugs [565]. The enzyme, responsible for the dephosphorylation of insulin receptors, has been identified and that is known as called protein-tyrosine phosphatase 1B (PTP1B). Therefore, PTP1B inhibitors as insulin sensitizer agents might be promising anti-diabetic drugs [566].

Increase in insulin sensitivity by PTP1B gene disruption in mouse models confirmed this hypothesis. Similar results were found when PTP1B antisense nucleotides suppressed PTP1B gene expression [567]. Protein tyrosine phosphatases (PTPs) represent a vast and structurally variable family of highly regulated enzymes and most of them have been proposed as advanced drug discovery targets.

The one of the well-established enzymes among PTPs is PTP1B [568-570] which was the first isolated member of the PTP superfamily, having links with insulin resistance, obesity and T2DM. It has been shown in various studies that PTP1B can negatively regulate insulin and leptin signaling pathways. In the insulin signaling pathway PTP1B dephosphorylates both the insulin receptor (IR) and its substrate IRS-1 [571, 572] while in the leptin pathway, it binds and dephosphorylates tyrosine kinase downstream of the Janus-Activated Kinase 2 (JAK2) leptin receptor [573]. In cell cultures, overexpression of PTP1B resulting to a decrease in the insulin-stimulated phosphorylation of IR and IRS-1, although PTP1B raises insulin-initiated signaling level reduction [574, 575]. Quantitative analysis of trait loci and mutations in the human PTP1B gene propped up the hypothesis that the expression of PTP1B might contribute to diabetes and obesity [576]. PTP1B knockout mice, in *in vivo* studies, displayed elevated resistance to insulin sensitivity and obesity induced by high-fat diet [577, 578].

Furthermore, other studies revealed that the tissue-specific PTP1B knockout mice that neuronal PTP1B controlled the leptin action, adiposity as well as body weight [579]. Thus, PTP1B inhibitors may be a highly promising approach in T2DM management and obesity amelioration. To conquer the lack of cellular activity of highly charged phosphonates, aryl carboxylic acids, including isoxazole [580], hydroxylpropionic [581], 2-oxalylamino benzoic (OBA) acids [582] and thiophene diacid [583] have been acknowledged as an alternative phosphotyrosine (pTyr) surrogates. Moreover, benzyl aryl α-ketoacid derivatives revealed, in a non-competitive pattern, significant PTP1B inhibitory effects and that was targeted to conserved protein loop (WPD loop) open conformation [584]. The presence of a benzyl group in these bioactive molecules may increase PTP1B binding affinity and

increases the cell membrane permeability also due to hydrophobic nature. It is also suggested that might become an oncogene in breast cancer [569]. Thus, PTP1B inhibitors may be a therapeutic target for T2DM, obesity and cancer. The search for novel and promising natural inhibitors of PTP1B is getting much attention.

Nearly 300 natural products from different natural sources having PTP1B inhibitory capacity were isolated and characterized and many of them were of marine origin [585]. The first documented marine metabolite having PTP1B inhibitory activity was the sulfircin, a sesterterpene sulfate, isolated from deep-water sponge Ircinia (unknown species) [585]. Afterwards, marine sponges with diverse structures including polybromodiphenyl ether [586], sesquiterpenoids and sesquiterpene quinones [587] got consideration as precious sources of PTP1B inhibitors. Even so, the novel screening models of marine resource has persuaded the new studies with the aim to find potential of these marine resources as forthcoming anti-diabetic agents. Marine algae, seaweeds, soft corals, sponges and lichens exhibited PTP1B inhibitory effects among these models.

***In Vitro* and *In Vivo* Concerns**

The mechanism of regulation of insulin signaling regulation is the action of PTPs on IR themselves or their substrates. Studies, which were made to find the role of PTPs in insulin signaling pathways and diabetes using vanadium compounds, revealed reduction in serum glucose levels in both T1 and T2 diabetic animal models [588, 589]. Since, vanadium compounds show fundamental *in vitro* and *in vivo* insulinomimetic effects therefore these compounds on oral administration promote the normalization of serum glucose levels in T2DM rats, increasing glucose uptake [590]. Following the insulin and vanadate treatment the increased levels of hepatic cytosolic PTP activity in these rats were decreased resulting to normalization of serum glucose levels. Such findings may be explained through the inhibition of PTPs, which as a consequence improves cellular tyrosine phosphorylation [591]. Studies based on the structure of PTP1B enzyme in addition to IR recognition, identified JAK2 and tyrosine kinase 2 (TYK2) as potential PTP1B substrates. In PTP1B null fibroblasts, both kinases showed hyperphosphorylation, upon stimulation through

interferons [592].

Further, this finding was attested in *in vivo* models where the negative regulation produced by PTP1B of leptin-stimulated JAK2 phosphorylation reduced leptin signaling. In PTP1B-deficient ob/ob mice, significant decrease in weight gain and increase in resting metabolic rates was observed on introducing null PTP1B mutation into leptin-deficient obese ob/ob mice. Fat pad analysis, further, suggested that variations in weight are due to decrease in adipose tissue. Therefore, loss in PTP1B in the absence of leptin may reduce weight gain with no any modification in food intake [593, 594]. Furthermore, due to leptin and feeding suppression, PTP1B-deficient mice had shown an increased response to weight loss. A noticeable improvement in leptin-induced transcription factor STAT3 phosphorylation evinced in the hypothalami of these mice, hints that in PTP1B deficiency introduction of exogenous leptin would result in to increase in leptin sensitivity [593, 594]. It was confirmed actually, in substrate trapping trials using catalytically inactive PTP1B D181A, that when leptin-activated JAK2 is considered as a PTP1B substrate in PTP1B null mice, the reduction in leptin signaling is an obesity resistance mechanism.

Human Concerns

Weight loss and improved insulin sensitivity in humans are closely related to decrease in PTP activity together with LAR and expression of PTP1B in adipose tissue [595]. It is noteworthy that PTP1B activity might not always related to its level of expression. In obese and diabetic subjects, the levels of PTP1B protein in abdominal adipose tissue showed a 3- to 5-fold increase and observed a remarkable decrease in the PTP1B activity per unit of PTP1B protein [596]. It is further observed that the marked rise in adipose tissue in obese individuals was not due to increased PTP1B activity but total cellular PTP. Additionally, the reduced insulin-stimulated glucose transport is due to increased in PTP activity, not by PTP1B activity, suggesting a tissue-specific role of PTP1B in glucose homeostasis [597]. Moreover, the mapping of the PTP1B locus to chromosome 20 in the region q13.1–q13.2 [598] provided genetic evidence linking PTP1B to diabetes and obesity in humans, because

this region is recognized as a quantitative trait locus linked to insulin and obesity. The role of PTP1B in insulin resistance has also been correlated to various polymorphisms. Thus, identification of new PTP inhibitors for diabetes and obesity control is demand of time.

3.9. THE DIPEPTIDYL PEPTIDASE IV (DPP-IV) INHIBITORS

To maintain euglycaemia, the proper regulation of insulin secretion is must. The deteriotaed insulin secretion and the developed peripheral insulin resistance in T2D may result to development of hyperglycaemia. During the fasting state, the extent of secretaion of insulin is physiologically small in order to enhance the uptake of glucose by the peripheral tissues. The secretaion of insulin is get stimulated quickly and considerably after meal so that plasma glucose concentrations be maintained within a narrow physiological range [599]. The promotion in the post-prandial stimulation of glucose is due not only to the post-prandial rise in glucose concentrations but also to the glucagon-like peptide-1 (GLP-1) and the gastric inhibitory polypeptide (GIP), which are gastrointestinal hormones. The stimulation of insulin secretion by GLP-1 and GIP is under hyperglycemic conditions and their contribuition in post-prandial insulin secretion is nearly 70%.

These are known as incretin hormones because of their importance physiologically, in the stimulation of post-prandial insulin secretion [600-602]. The incretin effect illustrates the phenomenon that orally ingested glucose, rather than administered intravenously, results to a much higher response to insulin. The incretin effect is weakened and the post-prandial insulin secretion is also deteriorating in type 2 diabetes [600, 603] and thus elevation of GLP-1 pharmacologically may restore insulin secretion [604]. GLP-1 stimulates the insulin secretion only under hyperglycaemic conditions therefore, there exits a negligible intrinsic risk of hypoglycaemia. The excessive stimulation of glucagon secretion in turn stimulates production of hepatic glucose production and therefore in T2D, GLP-1 is beneficial contributing to maintain euglycaemia. The inhibition of glucagon secretion under hyperglycaemic conditions by GLP-1 improves glycaemia. The biological or plasma half-life of this peptide hormone, GLP-1, is only a few minutes [602, 605] is because

of rapid enzymatic degradation of GLP-1 by the enzyme dipeptidyl peptidase IV (DPP-4) [606]. Orally active small molecules can inhibit DPP-4.

An increase in the concentration of endogeneous GLP-1 observed upon admistration of orally active small molecules as DPP-4 inhibitors [607]. Thus, for DPP-4, GLP-1 acts as a high affinity substrate ,i.e., "direct target" and the elevation of other substrates, besides GLP-1 by DPP-4 inhibition may also be contributory to the normalization of glycaemia in T2D and these substrates are considered as "indirect or off target" [608]. The physiology of the incretin hormones after food intake and the mode of action of the DPP-4 inhibitors are depicted in Figure 2.8 [605].

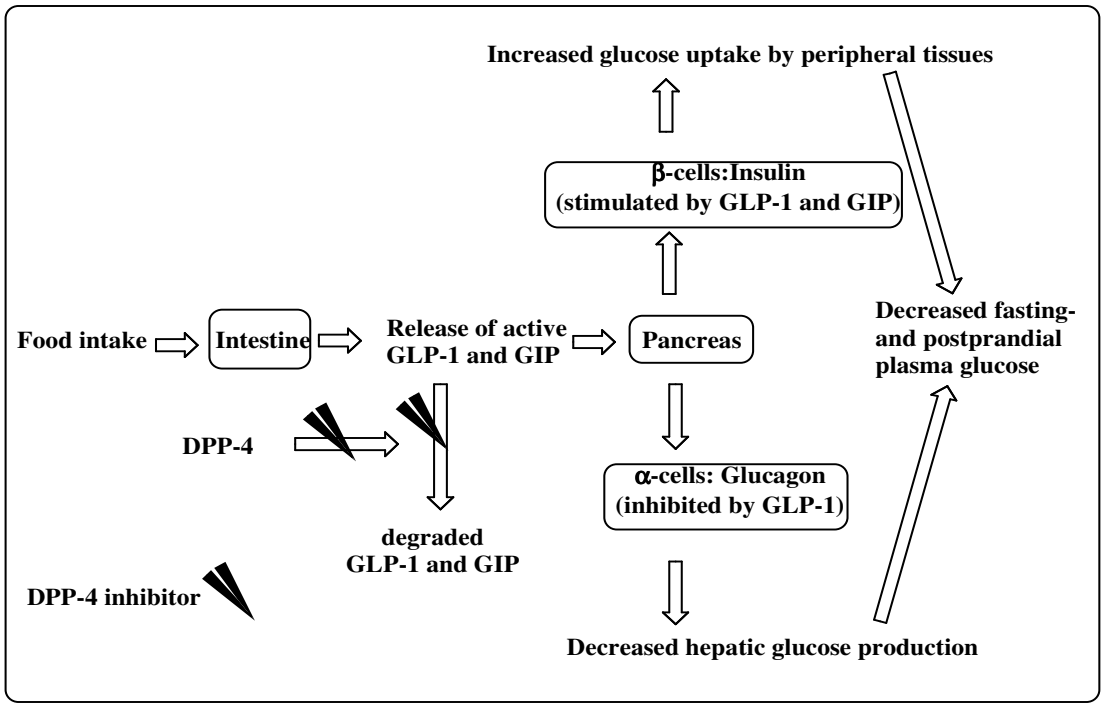


Figure 2.8: Action of DPP-4 inhibitors and physiology of the post-prandial regulation of glucose homoeostasis by the incretin system.

At present, for the treatment of T2D, DPP-4 inhibitors are the established class of oral antidiabetic agents. The first agent sitagliptin was introduced in 2006 [609] followed by linagliptin, vildagliptin, saxagliptin, and allogliptin. Agents including anagliptin, gemigliptin, and teneligliptin are in use in Asian countries. DPP-

4 inhibitors are being implemented in many national and international guidelines into the treatment algorithms of type 2 diabetes [610].

A homogenous class of molecules could not be assigned to the various DPP-4 inhibitors as they interact differently with the active site of the enzyme molecule. Based on the findings of the characterization of the binding modes of most widely clinically used DPP-4 inhibitors, three different classes of these have been proposed. Classes of the various commonly used DPP-4 inhibitors and the binding domains of the various classes to specific areas of the DPP-4 molecule according to Tomovic *et al.* [611] and Nabeno *et al.* [612] have been depicted in Figure 2.9.

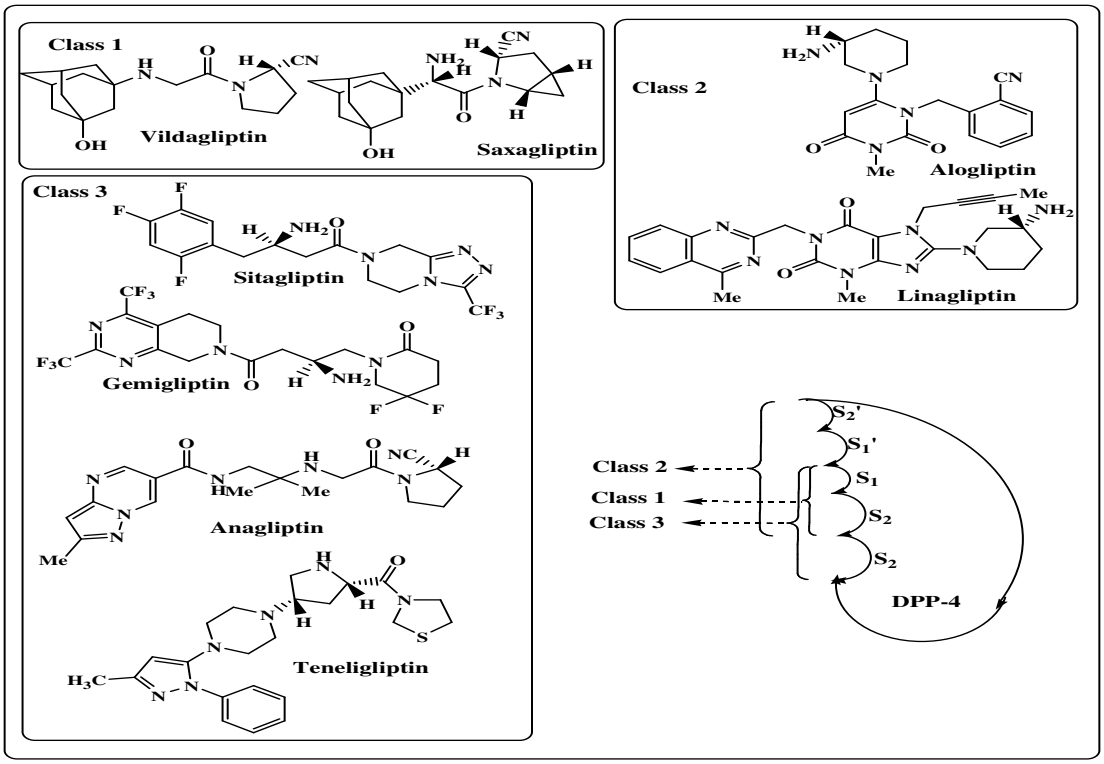


Figure 2.9: Classes of DPP-4 inhibitors with the numerous commonly used DPP-4 inhibitors and the binding domains of the various classes to specific areas of the DPP-4 molecule.

Class 1 comprises the molecules having interactions with the S1- and S2 subsites of the active center and covalently binding with Ser630 of the DPP-4

molecule e.g. saxagliptin and vildagliptin. Binding with S1 and S2 but also interact with S1' and/or S2', such as alogliptin and linagliptin belong to class 2. Sitagliptin, anagliptin, gemigliptin, and teneligliptin constitute the class 3 of the DPP-4 inhibitors [611, 612].

The above mentioned orally active, rapidly absorbed DPP-4 inhibitors are suitable for once daily or twice daily administration, resulting to inhibition of 70-90% over 24 h. These are eliminated renally after little metabolization but linagliptin. Saxagliptin gets metabolized generating an active metabolite. Linagliptin is eliminated through a biliary route [608, 611, 613, 614].

Cardiovascular safety studies

Adverse safety signals of rosiglitazone raised the concern for novel diabetes medications regarding proven cardiovascular safety of these as compared to standard therapy under glycaemic equipoise. In 2008, The FDA established the "Clinical Guidance for Pharmaceutical Industry–Diabetes Mellitus—Evaluating Cardiovascular Risk in New Antidiabetic Therapies to Treat Type 2 Diabetes". Till now, four studies for cardiovascular safety studies on DPP-4 inhibitors have been completed and published and these are the Examination of Cardiovascular Outcomes with Alogliptin vs. Standard of Care (EXAMINE) study for alogliptin [615, 616], the Saxagliptin Assessment of Vascular Outcomes Recorded in Patients with Diabetes Mellitus–Thrombolysis in Myocardial Infarction 53 (Savor-TIMI-53) study for saxagliptin [617, 618], the Trial Evaluating Cardiovascular Outcomes With Sitagliptin (TECOS) study for sitagliptin [619, 620] and the last Cardiovascular safety and Clinical Outcome with Linagliptin (CARMELINA) study for linagliptin [621, 622]. In all these studies, respective DPP-4 inhibitors have proved cardiovascular safety. These studies have shown a very homogenous result if the effects on the primary endpoint are compared. The results of the studies are heterogenous with respect to the secondary endpoint hospitalization due to heart failure as saxagliptin therapy was associated with a significant increase in the rate of hospitalization due to heart failure as compared to standard therapy [618].

However, this imbalance did not affect the primary endpoint and the

occurrence rate was higher in those patients who have a previous history of heart failure and independent of renal function at study baseline [623-625]. A similar but not so significant indication was observed for alogliptin but for the other DPP-4 inhibitors not as in the TECOS and CARMELINA study for sitagliptin and linagliptin, respectively [616, 620, 622, 623, 626]. This may be ascribed to that the differed cardiovascular disease status of the patients in the studies might have have influenced the outcome of heart failure outcome [623]. Thus, saxaglitpin treatment should be avoided consequently in patients with heart failure. The cardiovascular safety study, CAROLINA (CARdiovascular Outcome Trial of LINAgliptin vs. Glimepiride in Type 2 Diabetes) with linagliptin, comparing linagliptin treatment as add on therapy, to metformin directly and with a therapy with the sulfonylurea glimepiride will bring results later and may provide additional insights into the association of and mechanisms that links hypoglycaemic- with cardiovascular events [627, 628]. The demonstrated cardiovascular safety of the DPP-4 inhibitors in multiple studies is the basis for a positioning of the DPP-4 inhibitors as second-line therapy for the treatment of type 2 diabetes particularly when hypoglycaemia should strictly be avoided.

Positioning in the treatment algorithm of type 2 Diabetes

The American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD) have published a new joint position statement for the treatment of T2D in 2018 [629, 630]. In 2019, the ADA has adopted thses statements in annual recommendations “Standards of Medical Care in Diabetes” [631]. A patient centered and individualized treatment approach is used with the aim to prevent diabetes-related complications and to optimize quality of life in patients with T2D, is recommended in the treatment algorithm.

In these recommendations, the intervention in the life-style such as patient education and motivation, increase of physical activity and healthy eating, is still in the beginning and center of therapy as recommended previously by the ADA and EASD [632, 633] followed by with metformin therapy pharmacologically [629-633]. If therapeutic goals are not achieved with these measures, in that case certain patient

characteristics will determine the further recommended treatment options. Patients having established cardiovascular disease should receive an intensified pharmacological treatment with such an agent which has demonstrated benefit in earlier cardiovascular safety studies. A GLP-1 receptor agonist or SGLT-2 inhibitor with well characterized cardiovascular safety should be used in the patients with pre-existing atherosclerotic cardiovascular disease (ASCVD). Patients with prevalent heart failure (HF) or chronic kidney disease (CKD) would receive an intensified treatment with an SGLT-2 inhibitor with the respective evidence revealed by clinical studies due to the respective data [629-631].

DPP-4 inhibitors are recommended as one possible third-line therapy in addition to SGLT-2 inhibitors, thiazolidinediones or insulin in patients populated with both HF and CKD and with pre-existing cardiovascular disease if the therapeutic goals are unmet with the previous dual combination. The combination of DPP-4 inhibitors and GLP-1 receptor agonists is restricted not because of safety issues, but due to lack of significant additional benefit expected clinically [629-631]. Saxagliptin as DPP-4 inhibitor contraindicated, in patients with pre-existing HF, on cardiovascular safety grounds in respective study because there was a significant increase of hospitalization for HF as secondary endpoint in the saxagliptin arm [618, 624]. The summary of the ADA and EASD recommended treatment algorithm is given in Figure 2.10.

In conclusion, DPP-4 inhibitors are important oral antidiabetic agents placed as second-line therapy after failure of metformin as insulinotropic agents with minimum or no intrinsic risk of hypoglycaemia and body weight gain. Additionally, they inhibit glucagon secretion under hyperglycaemic conditions. Thus, these should be used mainly as a second-line therapy as add on to metformin in T2D patients with no pre-existing cardiovascular disease and as a goal to avoid hypoglycaemic events therapeutically.

There are only a few reported treatment-limiting adverse effects of DPP-4 inhibitors and these have shown cardiovascular safety. In impaired renal function patients, DPP-4 inhibitors have shown efficacy and safety profile.

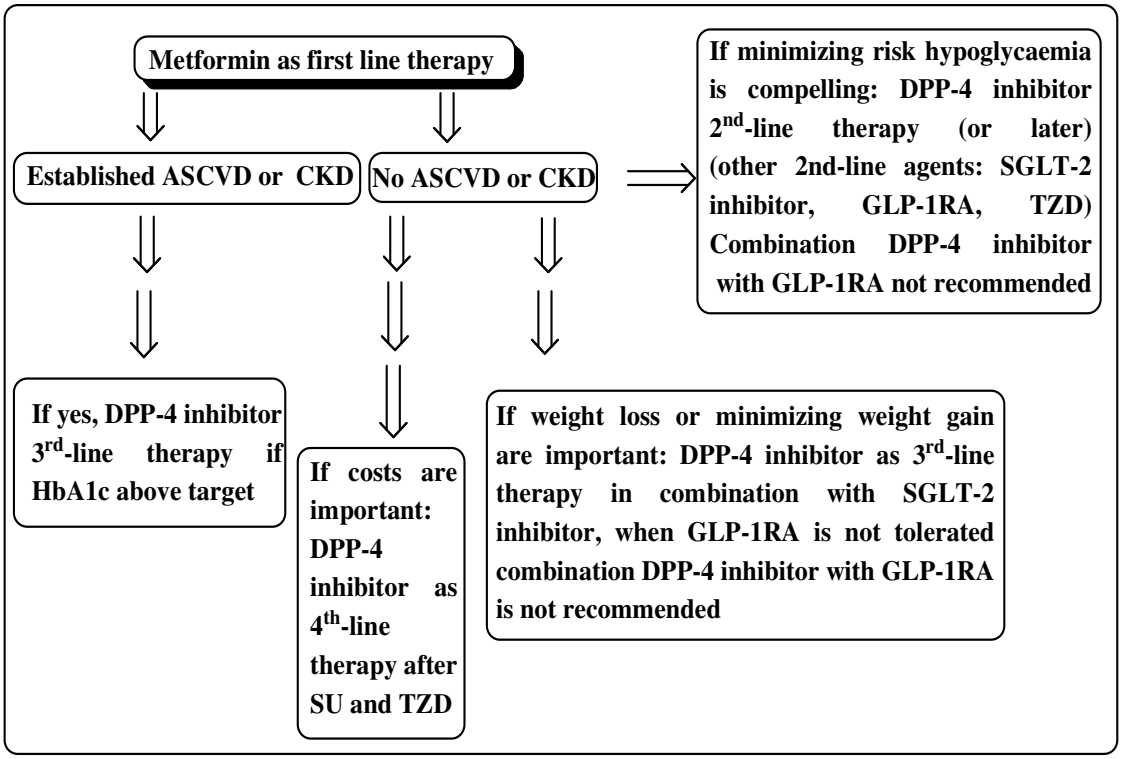


Figure 2.10: Placement of DPP-4 inhibitors into the treatment algorithm according to the recommendations of the ADA and EASD.

Additonally, DPP-4 inhibitors may also be used in triple combination treatment such as either with metformin and SGLT-2 inhibitors or with metformin and insulin. A reduction in hypoglycaemic episodes, in combination with insulin, was shown in some studies because of a reduction in the insulin dose. As both the DPP-4 inhibitors and GLP-1 receptor agonists, elevate "GLP-1" plasma concentrations therefore this combination is not recommended.

GLP-1 receptor agonists increase 8-10-fold whereas DPP-4 inhibitors lead to 2-3 fold endogenous GLP-1 concentrations. An explorative study has not showed additive effects of sitagliptin and liraglutide as the GLP-1 receptor agonist and studies showing additional effects using shorter acting GLP-1 receptor agonists and DPP-4 inhibitors are still needed [629-460, 634]. Not only the DPP-4 inhibitors as insulinotropic agents are replacing increasingly sulfonylureas but these might serve a

good therapeutic alternative to other treatment options including glitazones or glucosidase inhibitors as well.

3.10. GLUCOKINASE ACTIVATORS (GKAs)

It is imperative in the context of globally rising prevalence of T2DM and a low proportion of patients achieving optimal glycaemic control, there exist a need for novel target ant treatment strategies. As a result, novel therapies targeting pathogenetic pathways associated with the gut, brain and kidney, which are researched more recently, has been introduced. At present, except metformin, there is no available class of glucose-lowering medications targeting directly or indirectly the enhanced hepatic glucose ouput caused by primary dysregulation of liver that is associated with T2DM [635]. This unmet need may be served by the activation of glucokinase (GK) which belongs to the hexokinase family.

Glucose Homeostasis and GK

The nature of glucose homeostasis is albeit complex, it may be understood as a result of net effect of a two competitive hormones, they are the insulin and glucagon. Glucagon secures energy in the fasting state through maintaining euglycaemia and is secreted by α islet cells of the pancreas. It is achieved by two promoting two processes namely, gluconeogenesis ,i.e, which is the de novo glucose production from amino acids and fat and glycogenolysis ,i.e., glycogen breakdown and glucose release from the liver resulting to increase in the hepatic glucose output. The similar purpose may also be achieved by free fatty acids released from adipose tissue.

Contrary to this, a glucose-lowering effect in the fed state is produced by insulin which is secreted by β islet cells by utilizing glucose in the periphery including skeletal muscle and adipose tissue and uptaking of hepatic glucose through switching liver into a “glycogen synthesis mode”. In parallel, the elevated glucose inhibits gluconeogenesis and glycogenolysis in the fed state [636]. GK acts as a “glucose-sensor” [637] in pancreatic cells reducing secretion of glucose-stimulated insulin and as a “gate-keeper” for glucose in hepatocytes through promoting uptake of hepatic glucose and glycogen synthesis and storage. GK on activation phosphorylates

glucose through magnesium adenosine triphosphate to glucose-6-phosphate (G6P) which in turn activates glycogen synthase and serves as a substrate for glycogen synthesis [638]. The biochemical properties and kinetics made GK to serve this dual role [639] because when glucose is in the physiological range the activity of GK is restrained due to the low affinity for glucose ($K_{0.5}$ ~ 7-8 mmol/L) [640]. Additionally, GK is not inhibited by its end product secretion from the β-cells, the G6P, and having a sigmoidal saturation curve with glucose, following non-Michaelis–Menten kinetics, showed an inflection point close to the threshold of insulin secretion of 4-5 mmol/L.

Thus, graded responses for fluctuating glucose levels are ensured and glucokinase activity enters a plateau phase while glucose is near to the physiological threshold (5 mM) for glucose-stimulated insulin secretion [640]. In the liver, when glucose concentration is less than ~ 10 mM, GK remains as an inactive complex with the glucokinase regulatory protein (GKRP) which is its endogenous inhibitor. It is conferred that GK has lower affinity for hepatic glucose than to pancreatic β-cells and it gets activated only during the postprandial state to increase hepatic glucose uptake [641, 642]. In this way, at the hepatic cell, GKRP acts as a competitive inhibitor of glucose as it sequesters GK at low concentrations and and dissociaties from GK at increased glucose concentrations.

However, GK is also expressed in entero-endocrine cells, neurons, pancreatic α- and δ-cells, and cells in the anterior pituitary [643] but its major role in glucose homeostasis is in pancreatic β-cells and in hepatic cells. Based on the GK-mediated blood glucose lowering pathways it may be hypothesized that controlling of GK activity might represent a novel way to intervene in the glucose homeostasis as activation of GK results in to glucose lowering and its activity is low in T2DM patients [644]. Studies of a genetically discrete subgroup of diabetes and known as maturity-onset diabetes of the young type 2 (MODY2) has further reinforced this hypothesis. In MODY2, the people carrying inactivating heterozygous mutations in the glucokinase gene [645] manifested a benign form of hyperglycaemia usually and with low risk of associated microvascular complications [646] revealing a defective “glucose-sensing” ability [645]. On the other hand, presence of compound

heterozygosity or homozygosity may develop a severe form of permanent neonatal diabetes [647]. However, activating mutations in the glucokinase gene, leading to congenital hyperinsulinaemic hypoglycaemia with heterogenous phenotypic manifestations, do rarely occur [648].

The physiological role of GK in glucose homeostasis based on the *in vivo* evidence and experiments on mice and human cells [649-653], in addition to its feasibility of activation, corroborates glucokinase as target in patients with T2DM. The gradual decline in β-cell mass in T2D patients leading to defective insulin secretion in addition to elevated hepatic glucose output represent two discrete and interrelated derangements pathophysiologically in T2DM and that might be addressed by pharmacological upregulation of GK activity potentially.

GK Activators (GKAs)

After the introduction of the first agent of GKA in 2003 [654], various glucokinase activators have been designed and evaluated [655, 656]. These molecules facilitate activation of GK through binding to an allosteric site in the enzyme and stabilize a high-affinity conformation. This region of the enzyme is enriched with the majority of activating mutations. These may be classified according the chemical structure such as carbon-, urea-, 1,2,4-substituted aryl-, 1,3,5-substituted aryl-centred or other [640]. Another classification can be referred to the site of action such as hepatoselective and systemic. The hepatoselective GKAs act with or without disrupting the GK-GKRP interaction in hepatic cells [657, 658], which is contrary to systemic GKAs including piragliatin or dorzagliatin. In addition to another classification may be full or partial GKAs. Unfortunately, only a few GKAs have reached the phase of clinical trials.

Serious concerns about the efficacy and safety issues on use of older generation GKAs were raised such as risk of hypoglycaemia, induction of fatty liver, dyslipidaemia and diminished long term efficacy. In the early phases of GKAs development indeed, the over-stimulation of pancreatic GK and hepatic GK resulted into hypoglycaemia and dyslipidaemia, respectively were recognized as potential risks [639, 659]. A naturally occurring consequence of GK activation is the acute

insulin release, resulted from the exaggerated response to glucose, was also a reasonable risk.

The hypoglycaemic episodes were particularly occurred with piragliatin and MK-0941. This risk was addressed by the use of partial activators, which maintained a higer degree of dependency on glucose levels, so that the risk of activation at low glucose concentrations [655] might be minimized. Partial GKA PF-04937319 had shown lower hypoglycaemic risk [660]. As far as the pathophysiological processes associated to GK activation leading to dyslipidaemia is concerned, it is hypothesized that hepatic GK overstimulation resulting to excessive G6-P accumulation and that activates fructose 2,6-bisphosphate mediated glycolysis which demonstrate a parallel to G6-P increase and in this the underlying mechanism is feedforward allosteric activation. This activation of glycolysis finally results into the accumulation of acetyl- CoA converted from pyruvate which it turn results in increased influx to fatty acids via malonyl-CoA and triglycerides or enhanced de-novo hepatic lipogenesis [639]. It is in consistenacy with the first stage reported for nonalcoholic fatty liver disease (NAFLD) and that might range from simple steatosis to steatohepatitis [661]. Steatohepatitis on chronic exposure may develop the acute effect of hypertriglyceridemia as was noted with MK-0941 especially.

The increase in triglyceride levels are not so pronounced than the induced by a high-carbohydrate low-fat diet [639] and it remained still unwanted in patients with type 2 diabetes which are already prone to develop dyslipidaemia, hypertension and NAFLD [662]. It is interesting that the promising glucose-lowering efficly of GKA, which was noticed within the first few weeks of the introduction, was not prolonged over the course of the clinical studies. The earlier efficacy was declined rapidly through chronic exposure to the GKA agent and it was seen with both MK-0941 and AZD1656, GKA PF-04937319 and AMG 151 nearly four, three and one month, respectively. The reasons of this secondary failure are scantily understood. The study population with MK-0941 possibly may not have minimum critical β-cell mass left which is necessary for it to act as they have long-standing diabetes and were already following insulin treatment but this not in the case of AZD1656. It was suggested that

the failure of GKA was possibly caused by the hepatic de novo lipogenesis or plasma hyperlipidaemia, but it was not accepted universally [639].

In the case of hepatoselective TTP399, the glucose-lowering efficacy was marked only after 3 months of treatment trail but no failure in earlier phase trails. Based on the histological reports in mouse models where double-strand breaks in the deoxyribonucleic acid were seen, hypothesized the toxicity of GKAs on β-cells that it is due to activation of the p53 tumour suppressor and leadind β-cell death, followed by genetic activation of β-cell glucokinase [640, 663].

Indeed, this activation resulted in a rapid unsustainable decrease in blood glucose that is in conformity to observed trajectory in human clinical studies. Agius [640] proposed another hypothesis focusing on the two opposite effects of GKA on hepatic GK. The first stage occurs when GK-induced insulin secretion mends any abnormality observed in insulin/glucagon ratio and promotes GK/GKRP dissociation and hepatic GK activation. The second effect is operative when G6P and down-stream phosphate-ester intermediates of the glucose metabolism get accumulated to such an extent that repress GK gene [664] and thus offsetting any early stimulatory effect which may be induced by GKA [665] and renders their effect as clinically unimportant thus practical neutralization of them. Whatever be the causes for diminishing efficacy of GKA eventually it may inhibit development, maturation and approval of future GKA agents further.

3.11. THE GLUCAGON-LIKE PEPTIDE 1 (GLP-1) RECEPTOR AGONISTS

The incretin system, in recent years, has become an an important and useful target in the treatment of T2DM [666]. These are the hormones produced in a response to oral intake of nutrients by the intestinal mucosa which increase the glucose-stimulated insulin secretion and lower levels of blood glucose. Futhrmore, when the glucose levels are near to normal, they reduce insulin release also. It has been revealed that secretion of insulin is more in response to oral glucose ingestion as compared to an isoglycemic intravenous glucose infusion and this phenomenon is known as “the incretin effect” [667, 668]. The glucose-dependent insulinotropic polypeptide and glucagon-like peptide 1 (GLP-1) are the two identified incretin

hormones [669]. The interst in GLP-1 particularly is owed to its glucose-lowering effects [670] and its ability to slow gastric emptying and suppress secretion of glucagon as well [666]. T2D patients have reduced incretin effect [671].

The latest understanding of this deficit suggested that it is related to deterioration of the GLP-1 effect having impaired capacity of insulin secretion and enhanced insulin resistance, and hyperglycemia, that may resulting in to a decrease in GLP-1 receptor expression and lead to GLP-1 resistance [672]. The stimulation of GLP-1 receptors on administration of GLP-1 receptor agonists enhances the insulin secretion in response to oral and intravenous glucose. The similar extents of insulin secretion in both the cases means there is no change in the magnitude of the incretin effect [673]. Numerous GLP-1 receptor agonists have now been approved for treatment of type 2 diabetes.

Pharmacological Effects of GLP-1

In the setting of T2DM, GLP-1 has shown a number of potentially beneficial effects [674]. These are shown in Figure 5.11. In T2D patients, the intravenously administration of exogenous GLP-1 have shown a reduction in plasma glucose concentrations to the normal fasting range even in the patients having not enough response to oral antihyperglycemic drugs [675]. The effects of exogenous GLP-1 seen after the administration to T2DM patients [676] include decreased glucagon concentrations, improved insulin sensitivity, decreased A1C, slowed gastric emptying, increased satiety, decreased free fatty acid concentrations and decreased body weight. Due to its very short half-life and rapid degradation, native GLP-1s have limited therapeutical utility and it was overcome by the further development of degradation resistant GLP-1 receptor agonists [677].

Overview of GLP-1 Receptor Agonists

More than 10 years ago, the first agent in the class exenatide was approved and thereafter a number of GLP-1 receptor agonists are available including the exenatide (BID) (short-acting agents, twice daily) [678], liraglutide (intermediate-acting agents, administered once daily) [679] and exenatide QW (long-acting agents, administered once weekly) [680], albiglutide [681] and dulaglutide [682].

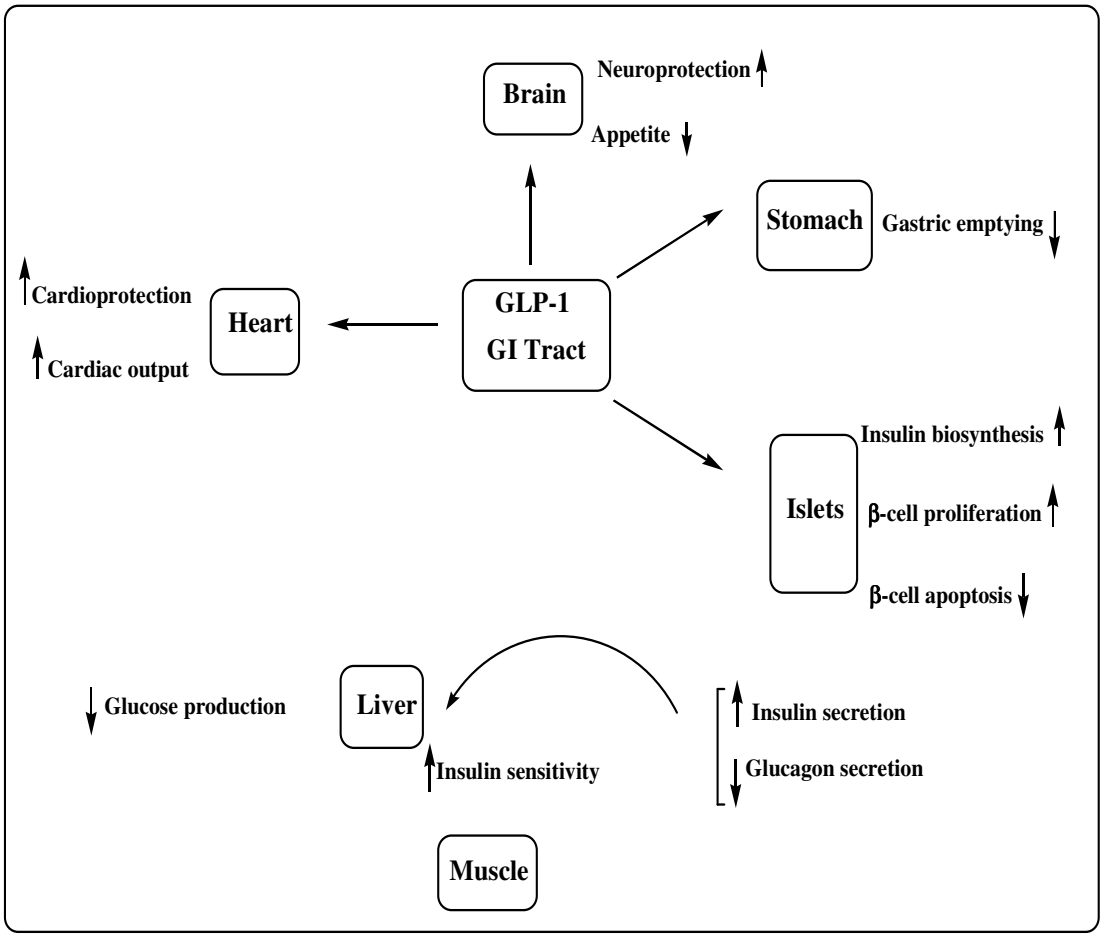


Figure 2.11: Various actions of GLP-1 in target tissues.

A recently approved and administered once daily is lixisenatide [683]. Each of these agents, have a certain half life to show its effects in plasma concentrations. The extended-release formulation of exenatide contained the same active compound as was in exenatide BID, which is encapsulated in microspheres to degrade slowly and continuous release of the drug is provided by them. The the primary goal of antihyperglycemic therapy is the glycemic control. It was indicated from the results of a meta-analysis of clinical studies that GLP-1 receptor agonists reduced the A1C from baseline versus placebo [684]. Studies also reported that mean changes in A1C

are pronounced on using the GLP-1 receptor agonists alone or in combination with oral antihyperglycemic therapies [678-683].

Acoording to the recent update to the American Heart Association/American Diabetes Association (AHA/ADA) guidelines regarding prevention of cardiovascular disease (CVD) in T2DM adults, weight management is a key component and suggested that the health care providers must consider using antihyperglycemic drugs which produce weight loss such as the GLP-1 receptor agonists [685]. In 2015, the clinical practice guidelines of American Association of Clinical Endocrinologists (AACE) and American College of Endocrinology (ACE) also focused the importance of weight management in T2D patients and urged the use of antihyperglycemic agents considering weight loss or at least with a neutral effect on body weight [686]. In clinical trials meant to evaluate GLP-1 receptor agonists, body weight reduction was commonly observed in T2D patients. Mean reductions in body weight in randomized controlled trials of GLP-1 receptor agonists (exenatide BID and QW and liraglutide) in overweight or obese T2D patients in a mixed-treatment comparison meta-analysis were greater than with placebo [680-682]. Weight loss on using GLP-1 receptor agonists is thought to be arisen as a result of slowed gastric emptying and increased satiety. Obese patients with accelerated gastric emptying following 30 days exenatide BID treatment led to slowed gastric emptying and a modest reduction in caloric intake than to placebo [687].

Cardiovascular Effects

The cardiovascular effects of antihyperglycemic treatments are of particular interest as the T2D peoples are at increased risk for cardiovascular complications. Morever, in 2008, FDA recommendations broadcasted a call for evidence that T2D therapies do not increase the risk of cardiovascular events like myocardial infarction [688] therefore, it is necessitated to assess cardiovascular outcomes in clinical trials of new antihyperglycemic agents and at present the intensified multifactorial treatment to therapeutic goals such as targeting glycemic control, blood pressure, lipid levels and renal function are associated with reduced cardiovascular and microvascular complications [689]. It is indicated by a number of findings that GLP-1

receptor agonists do not worsen CVD and may serve potential cardiovascular benefits in T2D patients [690].

A meta-analysis of 25 studies pertaining to GLP-1 receptor agonists has shown that there was no increase in major adverse cardiovascular events such as cardiovascular death, nonfatal myocardial infarction, stroke, and acute coronary syndromes and/or heart failure with a significant reduction versus placebo [691]. A retrospective analysis revealed that T2D patients receiving exenatide were less likely to have a CVD event than patients receiving other glucose-lowering treatments even though there was a prevalent previous ischemic heart disease, obesity, hyperlipidemia, hypertension, and other comorbidities at baseline [692]. The findings of ELIXA (Evaluation of Lixisenatide in Acute Coronary Syndrome) outcomes study were neutral with regard to cardiovascular outcomes [693]. On the other hand, liraglutide demonstrated advantages versus placebo in the LEADER (Liraglutide Effect and Action in Diabetes: Evaluation of cardiovascular outcome Results) trial [694]. Other intensified cardiovascular outcomes trials are also being conducted to evaluate the cardiovascular safety of exenatide QW, (EXenatide Study of Cardiovascular Event Lowering, EXSCEL) [695], albiglutide (HARMONY Outcomes) and dulaglutide (REsearching Cardiovascular Events with a Weekly INcretin in Diabetes, REWIND).

Place in Therapy

According to the AACE/ACE and ADA algorithms for the treatment of diabetes for glycemic control [696, 697] in the patients who are unable to achieve their A1C target following 3 months of metformin therapy, GLP-1 receptor agonists are recommended as add-on therapy. In the patients who cannot tolerate or are contraindicated for metformin, GLP-1 receptor agonists are the recommended first-line therapy also, as a substitute to metformin. GLP-1 receptor agonists stimulate release of insulin and suppress glucagon secretion only at elevated blood glucose concentrations, thus these are well suited for early use in T2D with low risk of hypoglycemia [698]. As dual therapy for patients who do not achieve A1C goals with metformin alone these are recommended in combination with metformin.

Patients with persistent hyperglycemia and overweight patients trying to control their weight as a triple therapy GLP-1 receptor agonists might be combined with metformin and a sodium-glucose cotransporter 2 inhibitor. In addition to this, the use of bolus (mealtime) insulin may be delayed by the incretin use with basal insulin leads to reduced risk of hypoglycemia. This simplified regimen is beneficiary in reducing the need for matching mealtime insulin to specific carbohydrate ratios and thus helps to mitigate the weight gain often noted using insulin, also.

Since the beginning of the GLP-1 therapeutics, the use of of GLP-1R agonists for the treatment of T2D and obesity is emergent. The demonstration that these agents reduce myocardial infarction, stroke and cardiovascular death unlike insulin with a favorable benefit/risk profile has broadened their clinical use and more enhanced interest in mechanisms of GLP-1 action. The declined long-term safety in GLP-1-treated subjects with obesity at high risk for cardiovascular disease remained an important face up to for expansion of GLP-1 therapeutics in non-diabetic peoples.

A better understanding of how nutrients, bacterial metabolites, and microbial populations control GLP-1 secretion might enable the development of GLP-1 secretagogues which are more potent and better tolerated than metformin. Furthermore, multiple new GLP-1-based therapeutic agents, either they be small molecules and peptides or larger hybrid proteins, keep on to be developed. Hence, for scientists and healthcare providers which are focused on the treatment of diabetes, obesity, and related complications further delineation of the mechanisms of action of GLP-1 be continued to have immediate translational relevance.

The world's most social health shocking disease is the diabetes mellitus (DM), not only due to its high occurrence but for its chronic sequels derived at inadequate glycemic control. Almost ninety percent of the diabetic population is together with this type 2 diabetes mellitus (T2DM) and keeps on rising all around the occidental world prevalently as an upshot of population aging and increasing of obesity and sedentary life styles. The most frequent endocrine-metabolic diseases, T2DM and obesity, are characterized through insulin resistance, defects in insulin secretion

which ultimately lead to a high hepatic production of glucose. The management of these patients therapeutically must include alteration in their life styles, adequate diet and exercise and if not controlled properly in that case pharmacological methods might be applied. A wide range of numerous therapeutic approaches are being employed to tackle the multifactorial nature of T2DM. The inclusion of cardiovascular parameters in these approaches in addition to the growing awareness of the importance of dealing with cardiometabolic disease including allied liver diseases, in toto, means that the near future holds great potential promise for the identification of novel therapeutic strategies for T2DM and its associated diseases.

REFERENCES

172. Al-Lawati JA. Om Med J (2017) 32:177-179.

173. Xiao JB and Högger P. Curr Med Chem (2015) 22:23-38.

174. Cho NH, Shaw JE, Karuranga S, Huang Y, da Rocha Fernandes JD, Ohlrogge AW and Malanda B. Diabetes Res Clin Pract (2018) 138:271-281.

175. American Diabetes Association. Diabetes Care (2015) 40:S11-S24.

176. Volpe CMO, Villar-Delfino PH, Dos Anjos PMF and Nogueira-Machado JA. Cell Death Dis (2018) 9:119.

177. Harrison, L.C. In Clinical Immunology (Robert RR, William TS, Anthony JF, Thomas AF, Harry WS and Cornelia MW, Eds.). Elsevier (2019) pp. 957-966.

178. Goedeke L, Perry RJ and Shulman GI. Ann Rev Pharmacol Toxicol (2019) 59:65-87.

179. Chung YR, Ha KH, Kim HC, Park SJ, Lee K and Kim DJ. Diabetes Metab J (2019) 43:640-648.

180. American Diabetes Association. Diabetes Care (2010) 33:S62-S69.

181. Reddy VS, Sahay RK and Bhada SK. J Indian Acad Clin Med (2000) 1:245-251.

182. Moller DE. Nature (2001) 414:821-827.

183. McKinlay J and Marceau L. Lancet (2000) 356:757-761.

184. American Diabetes Association. Diabetes Care (2000) 23:77-79.

185. Definition, Diagnosis and Classification of Diabetes Mellitus and Its Complications, World Health Organization, Department of Non Communicable Disease Surveillance, Geneva, Switzerland, 1999.

186. Jovanovic L and Pettitt DJ. J Am Med Assoc (2001) 286:2516-2518.

187. Coustan D. In Diabetes in America. (Harris MI, Cowie CC, Stern MP, Boyko EJ, Reiber GE and Bennett PH, Eds.), 2nd edition (1995) pp. 703-717. US Government Printing Office, Washington, DC, USA.

188. Kronenberg HM, Melmed S, Polonsky KS and Larsen PR. Williams Textbook of Endocrinology, 11th edition (2008) Saunders, Philadelphia, Pa, USA.

189. Ovalle F, Vaughan III TB, Sohn JE and Gower B. Am J Med Sci (2008)

335:298-303.

190. Trout KK and Scheiner G. Endocrinol Rev (2008) pp. 27-29.

191. National Institute for Clinical Excellence, Technology Appraisal Guidance 60. Guidance on the Use of Patient Education Models for Diabetes, NICE, 2003.

192. Grundy SM. Nat Rev Drug Discov (2006) 5:295-309.

193. Haffner SM, Lehto S, Rönnemaa T, Pyörälä K, Laakso M. N Engl J Med (1998) 339:229-234.

194. Meier M and Hummel M. Vasc Health Risk Manag (2009) 5:859-871.

195. Best JH, Hoogwerf BJ, Herman WH, Pelletier EM, Smith DB, Wenten M and Hussein MA. Diabetes Care (2011) 34:90-95.

196. Bjorntorp P. Nutrition (1997) 13:795-803.

197. Overman RA, Yeh JY and Deal CL. Arthritis Care Res (2013) 65:294-298.

198. Seckl JR and Walker BR. Endocrinology (2001) 142:1371-1376.

199. Stewart PM and Krozowski ZS. Vitam Horm (1999) 57:249-324.

200. Chapman KE and Seckl JR. Neurochem Res (2008) 33:624-636.

201. Albiston AL, Obeyesekere VR, Smith RE and Krozowski ZS. Mol Cell Endocrinol (1994) 105:R11-R17.

202. Masuzaki H, et al. Science (2001) 294:166-170.

203. Masuzaki H, Yamamoto H, Kenyon CJ, Elmquist JK, Morton NM, Paterson JM, Shinyama H, Sharp MG, Fleming S, Mullins JJ, Seckl JR and Flier JS. J Clin Invest (2003) 112:83-90.

204. Morgan SA, McCabe EL, Gathercole LL, Hassan-Smith ZK, Larner DP, Bujalska IJ, Stewart PM, Tomlinson JW and Lavery GG. Proc Natl Acad Sci USA (2014) 111:E2482-E2491.

205. Wake DJ, Rask E, Livingstone DE, Söderberg S, Olsson T and Walker BR. J Clin Endocrinol Metab (2003) 88:3983-3988.

206. Tomlinson JW, Moore J, Cooper MS, Bujalska I, Shahmanesh M, Burt C, Strain A, Hewison M, Stewart PM. Endocrinology (2001) 142:1982-1989.

207. Hirosumi J, Tuncman G, Chang L, Görgün CZ, Uysal KT, Maeda K, Karin M and Hotamisligil GS. Nature (2002) 420:333-336.

208. Han MS, Jung DY, Morel C, Lakhani SA, Kim JK, Flavell RA and Davis RJ. Science (2013) 339:218-222.

209. Kaneto H Nakatani Y, Miyatsuka T, Kawamori D, Matsuoka TA, Matsuhisa M, Kajimoto Y, Ichijo H, Yamasaki Y and Hori M. Nat Med (2004) 10:1128-1132.

210. Ijaz A, Tejada T, Catanuto P, Xia X, Elliot SJ, Lenz O, Jauregui A, Saenz MO, Molano RD, Pileggi A, Ricordi C and Fornoni A. Kidney Int (2009) 75:381-388.

211. Murtagh J, McArdle E, Gilligan E, Thornton L, Furlong F and Martin F. J Cell Biol (2004) 166:133-143.

212. Qi AQ, Qiu J, Xiao L and Chen YZ. J Neurosci Res (2005) 80:510-517.

213. Furst R, Zahler S and Vollmar AM. Endocrinology (2008) 149:3635-3642.

214. Peng K, Pan Y, Li J, Khan Z, Fan M, Yin H, Tong C, Zhao Y, Liang G and Zheng C. Sci Rep (2016) 6:37160.

215. Alberti L, Girola A, Gilardini L, Conti A, Cattaldo S, Micheletto G and Invitti C. Int J Obes (London) (2007) 31:1826-1831.

216. Desbriere,R, Vuaroqueaux V, Achard V, Boullu-Ciocca S, Labuhn M, Dutour A and Grino M. Obesity (2006) 14:794-798.

217. Michailidou Z, Jensen MD, Dumesic DA, Chapman KE, Seckl JR, Walker BR and Morton NM. Obesity (2007) 15:1155-1163.

218. Munoz R, Carvajal C, Escalona A, Boza C, Perez G, Ibanez L and Fardella C. Obes Surg (2009) 19:764-770.

219. Paulsen SK, Pedersen SB, Fisker S, and Richelsen, B. Obesity (2007) 15:1954-1960.

220. Tomlinson JW, Finney J, Gay C, Hughes BA, Hughes SV and Stewart PM. Diabetes (2008) 57:2652-2660.

221. Wamil M and Seckl JR. Drug Discov Today (2007) 12:504-520.

222. Masuzaki H, Paterson J, Shinyama H, Morton NM, Mullins JJ, Seckl JR and Flier JS. Science (2001) 294:2166-2170.

223. Morton NM, Paterson JM, Masuzaki H, Holmes MC, Staels B, Fievet C,

Walker BR, Flier JS, Mullins JJ and Seckl JR. Diabetes (2004) 53:931-938.

224. Edgerton DS, Basu R, Ramnanan CJ, Farmer TD, Neal D, Scott M, Jacobson P, Rizza RA and Cherrington AD. Am J Physiol Endocrinol Metab (2010) 298:E1019-E1026.

225. Winnick JJ, Ramnanan CJ, Saraswathi V, Roop J, Scott M, Jacobson P, Jung P, Basu R, Cherrington AD and Edgerton DS. Am J Physiol Endocrinol Metab (2013) 304:E747-E756.

226. Sundbom M, Kaiser C, Bjorkstrand E, Castro VM, Larsson C, Selen G, Nyhem CS and James SR. BMC Pharmacol (2008) 8:3.

227. Walker BR, Connacher AA, Lindsay RM, Webb DJ and Edwards CR. J Clin Endocrinol Metab (1995) 80:3155-3159.

228. Taylor A, Irwin N, McKillop AM, Flatt PR and Gault VA. Biol Chem (2008) 389:441-445.

229. Stewart PM, Wallace AM, Valentino R, Burt D, Shackleton CH and Edwards CR. Lancet (1987) 2:821-824.

230. Anderson A and Walker BR. Drugs (2013) 73:1385-1393.

231. Hermanowski-Vosatka A, Balkovec JM, Cheng K, Chen HY, Hernandez M, Koo GC, Le Grand CB, Li Z, Metzger JM, Mundt SS, Noonan H, Nunes CN, Olson SH, Pikounis B, Ren N, Robertson N, Schaeffer JM, Shah K, Springer MS, Strack AM, Strowski M, Wu K, Wu T, Xiao J, Zhang BB, Wright SD and Thieringer R. J Exp Med (2005) 202:517-527.

232. Rosenstock J, Banarer S, Fonseca VA, Inzucchi SE, Sun W, Yao W, Hollis G, Flores R, Levy R, Williams WV, Seckl JR and Huber R. Diabetes Care (2010) 33:1516-1522.

233. Feig PU, Shah S, Hermanowski-Vosatka A, Plotkin D, Springer MS, Donahue S, Thach C, Klein EJ, Lai E and Kaufman KD. Diabetes Obes Metab (2011) 13:498-504.

234. Shah S, Hermanowski-Vosatka A, Gibson K, Ruck RA, Jia G, Zhang J, Hwang PM, Ryan NW, Langdon RB and Feig PU. J Am Soc Hypertens (2011) 5:166-176.

235. Xu Z, Tice CM, Zhao W, Cacatian S, Ye YJ, Singh SB, Lindblom P, McKeever BM, Krosky PM, Kruk BA, Berbaum J, Harrison RK, Johnson JA, Bukhtiyarov Y, Panemangalore R, Scott BB, Zhao Y, Bruno JG, Togias J, Guo J, Guo R, Carroll PJ, McGeehan GM, Zhuang L, He W and Claremon DA. J Med Chem (2011) 54:6050-6062.
236. Hamilton BS, Himmelsbach F, Nar H, Schuler-Metz A, Krosky P, Guo J, Guo R, Meng S, Zhao Y, Lala DS. Zhuang L, Claremon DA and McGeehan GM. Eur J Pharmacol (2015) 746:50-55.
237. Ahren B. Nat Rev Drug Discov (2009) 8:369-385.
238. Mohler ML, He Y, Wu Z, Hwang DJ and Miller DD. Med Res Rev (2009) 29:125-195.
239. Fredriksson R, Hoglund PJ, Gloriam DE, Lagerstrom MC and Schioth HB. FEBS Lett (2003) 554:381-388.
240. Shah U and Kowalski TJ. Vitam Horm (2010) 84:415-448.
241. Bonini JA and Borowsky BE. (2001) p. US6221660.
242. Bonini JA and Borowsky BE. (2002) p. US6468756.
243. Jones RM. (2004) p. WO2004065380.
244. Takeda S, Kadowaki S, Haga T, Takaesu H and Mitaku S. FEBS Lett (2002) 520:97-101.
245. Davey J. Expert Opin Ther Targets (2004) 8:165-170.
246. Griffin G. (2006) p.US7083933.
247. Fyfe MC. Expert Opin Drug Discov (2008) 3:403-413.
248. Fu J. Nature (2003) 425:90-93.
249. Ohishi T. (2003) p. EP1338651.
250. Soga T, Ohishi T, Matsui T, Saito T, Matsumoto M, Takasaki J, Matsumoto S, Kamohara M, Hiyama H, Yoshida S, Momose K, Ueda Y, Matsushime H, Kobori M and Furuichi K. Biochem Biophys Res Commun (2005) 326:744-751.
251. Chu ZL, Jones RM, He H and Carroll C. Endocrinology (2007) 148:2601-2609.

252. Ohishi T and Yoshida S. Expert Opin Investig Drugs (2012) 21:321-328.
253. Chu Z-L. (2006) p. WO2006076231
254. Drucker DJ, Jin T, Asa SL, Young TA and Brubaker PL. Mol Endocrinol (1994) 8:1646-1655.
255. Dhayal S and Morgan NG. Drug News Perspect (2010) 23:418-424.
256. Overton HA, Babbs AJ, Doel SM, Fyfe MC, Gardner LS, Griffin G, Jackson HC, Procter MJ, Rasamison CM, Tang-Christensen M, Widdowson PS, Williams GM and Reynet C. Cell Metab (2006) 3:167-175.
257. Overton HA, Fyfe MC, Reynet C. Br J Pharmacol (2008) 153(Suppl 1):S76-S81.
258. Rodriguez de Fonseca F, Navarro M, Gomez R, Escuredo L, Nava F, Fu J, Murillo-Rodriguez E, Giuffrida A, LoVerme J, Gaetani S, Kathuria S, Gall C and Piomelli D. Nature (2001) 414:209-212.
259. Yang Y, Chen M, Georgeson KE and Harmon CM. Am J Physiol Regul Integr Comp Physiol (2007) 292:R235-241.
260. Proulx K, Cota D, Castaneda TR, Tschop MH, D'Alessio DA, Tso P, Woods SC and Seeley RJ. Am J Physiol Regul Integr Comp Physiol (2005) 89:R729-737.
261. Campbell JE and Drucker DJ. Cell Metab (2013) 17:819-37.
262. International Textbook of Diabetes Mellitus. (DeFronzo RA, Ferrannini E, Zimmet P and Alberti KGMM. Eds.). (2015) Oxford:Wiley-Blackwell Ltd.
263. Li Y, Li L and Hölscher C. Rev Neurosci (2016) 27:689-711.
264. João AL, Reis F and Fernandes R. Obes Rev (2016) 17:553-72.
265. Holst JJ, Vilsbøll T and Deacon CF. Mol Cell Endocrinol (2009) 297:127-136.
266. Watterson KR, Hudson BD, Ulven T and Milligan G. Front Endocrinol (2014) 5:137.
267. Poulsen Ll, Siersbæk M and Mandrup S. Semin Cell Dev Biol (2012) 23:631-639.
268. Kebede MA, Alquier T, Latour MG and Poitout V. Diabetes Obes Metab

(2009) 11:10-20.

269. Fyfe MC, McCormack JG, Overton HA, Procter MJ and Reynet C. Expert Opin Drug Discov (2008) 3:403-413.

270. Chu ZL, Carroll C, Alfonso J, Gutierrez V, He H, Lucman A, Pedraza M, Mondala H, Gao H, Bagnol D, Chen R, Jones RM, Behan DP and Leonard J. Endocrinology (2008) 149:2038-2047.

271. Hothersall JD, Bussey CE, Brown AJ, Scott JS, Dale I and Rawlins P. Eur J Pharmacol (2015) 762:430-442.

272. Ritter K, Buning C, Halland N, Pöverlein C and Schwink L. J Med Chem (2016) 59:3579-3592.

273. Nunez DJ, Bush MA, Collins DA, McMullen SL, Gillmor D, Apseloff G, Atiee G, Corsino L, Morrow L and Feldman PL. PLoS One (2014) 9:e92494.

274. Saraiva FK and Sposito AC. Cardiovasc Diabetol (2014) 13:142.

275. Candeias EM, Sebastião IC, Cardoso SM, Correia SC, Carvalho CI, Plácido AI, Santos MS, Oliveira CR, Moreira PI and Duarte AI. World J Diabetes (2015) 6:807-827.

276. Jones RM, Leonard JN, Buzard DJ and Lehmann J. Expert Opin Ther Pat (2009) 19:1339-1359.

277. Ritter K, Buning C, Halland N, Pöverlein C and Schwink L. J Med Chem (2016) 59:3579-3592.

278. Semple G, Ren A, Fioravanti B, Pereira G, Calderon I, Choi K, Xiong Y, Shin YJ, Gharbaoui T, Sage CR, Morgan M, Xing C, Chu ZL, Leonard JN, Grottick AJ, Al-Shamma H, Liang Y, Demarest KT and Jones RM. Bioorg Med Chem Lett (2011) 21:3134-3141.

279. Polli JW, Hussey E, Bush M, Generaux G, Smith G, Collins D, McMullen S, Turner N and Nunez DJ. Xenobiotica. (2013) 43:498-508.

280. http://clinicaltrials.gov/ct2/show/NCT01035879.

281. Matsuda D, Kobashi Y, Mikami A, Kawamura M, Shiozawa F, Kawabe K, Hamada M, Oda K, Nishimoto S, Kimura K, Miyoshi M, Takayama N, Kakinuma H and Ohtake N. Bioorg Med Chem Lett (2016) 26:3441-3446.

282. Alelyunas YW, Empfield JR, McCarthy D, Spreen RC, Bui K, Pelosi-Kilby L and Shen C. Bioorg Med Chem Lett (2010) 20:7312-7316.
283. Makishima M, Okamoto AY, Repa JJ, Tu H, Learned RM, Luk A, Hull MV, Lustig KD, Mangelsdorf DJ and Shan B. Science (1999) 284:1362-1365.
284. Lefebvre P, Cariou B, Lien F, Kuipers F and Staels B. Physiol Rev (2009) 89:147-191.
285. Ahlberg J, Angelin B, Bjorkhem I and Einarsson K. Gastroenterology (1977) 73:1377-1382.
286. Guo C, Xie S, Chi Z, Zhang J, Liu Y, Zhang L, Zheng M, Zhang X, Xia D, Ke Y, Lu L and Wang D. Immunity (2016) 45:802-816.
287. Benoit B, Meugnier E, Castelli M, Chanon S, Vieille-Marchiset A, Durand C, Bendridi N, Pesenti S, Monternier PA, Durieux AC, Freyssenet D, Rieusset J, Lefai E, Vidal H and Ruzzin J. Nat Med (2017) 23:990-996.
288. Li T and Chiang JY. Pharmacol Rev (2014) 66:948-983.
289. Shapiro H, Kolodziejczyk AA, Halstuch D and Elinav E. J Exp Med (2018) 215:383-396.
290. Katsuma S, Hirasawa A and Tsujimoto G. Biochem Biophys Res Commun (2005) 329:386-390.
291. Kawamata Y, Fujii R, Hosoya M, Harada M, Yoshida H, Miwa M, Fukusumi S, Habata Y, Itoh T, Shintani Y, Hinuma S, Fujisawa Y and Fujino M. J Biol Chem (2003) 278:9435-9440.
292. Wang YD, Chen WD and Huang W. Histol Histopathol (2008) 23:621-627.
293. Sato H, Genet C, Strehle A, Thomas C, Lobstein A, Wagner A, Mioskowski C, Auwerx J and Saladin R. Biochem Biophys Res Commun (2007) 362:793-798.
294. Lo SH, Li Y, Cheng KC, Niu CS, Cheng JT and Niu HS. Naunyn Schmiedebergs Arch Pharmacol (2017) 390:1097-1104.
295. Wang LY, Cheng KC, Li Y, Niu CS, Cheng JT and Niu HS. Biomed Pharmacother (2017) 95:599-604.
296. Li Y, Cheng KC, Niu CS, Lo SH, Cheng JT and Niu HS. Drug Des Devel

Ther (2017) 11:1127-1134.

297. Baggio LL and Drucker DJ. Gastroenterology (2007) 132:2131-2157.

298. Trabelsi MS, Daoudi M, Prawitt J, Ducastel S, Touche V, Sayin SI, Perino A, Brighton CA, Sebti Y, Kluza J, Briand O, Dehondt H, Vallez E, Dorchies E, Baud G, Spinelli V, Hennuyer N, Caron S and Bantubungi K. Nat Commun (2015) 6:7629.

299. Kumar DP, Rajagopal S, Mahavadi S, Mirshahi F, Grider JR, Murthy KS and Sanyal AJ. Biochem Biophys Res Commun (2012) 427:600-605.

300. Renga B, Mencarelli A, Vavassori P, Brancaleone V and Fiorucci S. Biochim Biophys Acta (2010) 1802:363-372.

301. Kumar DP, Asgharpour A, Mirshahi F, Park SH, Liu S, Imai Y, Nadler JL, Grider JR, Murthy KS and Sanyal AJ. J Biol Chem (2016) 291:6626-6640.

302. Alejandro EU, Gregg B, Blandino-Rosano M, Cras-Meneur C, Bernal-Mizrachi E. Mol Aspects Med (2015) 42:19-41.

303. Zarrinpar A and Loomba R. Aliment Pharmacol Ther (2012) 36:909-921.

304. Watanabe M, Houten SM, Mataki C, Christoffolete MA, Kim BW, Sato H, Messaddeq N, Harney JW, Ezaki O, Kodama T, Schoonjans K, Bianco AC and Auwerx J. Nature (2006) 439:484-489.

305. Perino A, Pols TW, Nomura M, Stein S, Pellicciari R and Schoonjans K. J Clin Invest (2014) 124:5424-5436.

306. Worthmann A, John C, Ruhlemann MC, Baguhl M, Heinsen FA, Schaltenberg N, Heine M, Schlein C, Evangelakos I, Mineo C, Fischer M, Dandri M, Kremoser C, Scheja L, Franke A, Shaul PW and Heeren J. Nat Med (2017) 23:839-849.

307. Wang XX, Edelstein MH, Gafter U, Qiu L, Luo Y, Dobrinskikh E, Lucia S, Adorini L, D'Agati VD, Levi J, Rosenberg A, Kopp JB, Gius DR, Saleem MA and Levi M. J Am Soc Nephrol (2016), 27:1362-1378.

308. Pierre JF, Martinez KB, Ye H, Nadimpalli A, Morton TC, Yang J, Wang Q, Patno N, Chang EB and Yin DP. Am J Physiol Gastrointest Liver Physiol (2016) 311:G286-G304.

309. Briere DA, Ruan X, Cheng CC, Siesky AM, Fitch TE, Dominguez C, Sanfeliciano SG, Montero C, Suen CS, Xu Y, Coskun T and Michael MD. PLoS One (2015) 10:e0136873.

310. Bhat BG, LeeAnne M, Dean PP, Judit M, Shufang Z, et al. J Pharmacol Clin Res (2017) 3:555623.

311. Lasalle M, Hoguet V, Hennuyer N, Leroux F, Piveteau C, Belloy L, Lestavel S, Vallez E, Dorchies E, Duplan I, Sevin E, Culot M, Gosselet F, Boulahjar R, Herledan A, Staels B, Deprez B, Tailleux A and Charton J. J Med Chem (2017) 60:4185-4211.

312. Murphy MP. Biochem J (2009) 417:1-13.

313. McBride HM, Neuspiel M and Wasiak S. Curr Biol. (2006) 16:R551-560.

314. Balaban RS, Nemoto S and Finkel T. Cell (2005) 120:483-95.

315. Raha S and Robinson BH. Trends Biochem Sci (2000) 25:502-508.

316. Garg AK and Aggarwal BB. Mol Immunol (2002) 39:509-517.

317. Janssen-Heininger YM, Poynter ME and Baeuerle PA. Free Radic Biol Med (2000) 28:1317-1327.

318. Fontana L, Partridge L and Longo VD. Science (2010) 328:321-326.

319. Colman RJ, Anderson RM, Johnson SC, Kastman EK, Kosmatka KJ, Beasley TM, Allison DB, Cruzen C, Simmons HA, Kemnitz JW and Weindruch R. Science (2009) 325:201-214.

320. Fontana L, Meyer TE, Klein S and Holloszy JO. Proc Natl Acad Sci USA (2004) 101:6659-6663.

321. Meyer TE, Kovács SJ, Ehsani AA, Klein S, Holloszy JO and Fontana L. J Am Coll Cardiol (2006) 47:398-402.

322. Imai S, Armstrong CM, KaeberleinM and Guarente L. Nature (2000) 403:795-800.

323. Cohen HY, Miller C, Bitterman KJ, Wall NR, Hekking B, Kessler B, Howitz KT, Gorospe M, de Cabo R and Sinclair DA. Science (2004) 305:390-392.

324. Bordone L, Cohen D, Robinson A, Motta MC, van Veen E, Czopik A, Steele AD, Crowe H, Marmor S, Luo J, Gu W and Guarente L. Aging Cell (2007)

6:759-767.

325. Boily G, Seifert EL, Bevilacqua L, He XH, Sabourin G, Estey C, Moffat C, Crawford S, Saliba S, Jardine K, Xuan J, Evans M, Harper ME, McBurney MW. PLoS One (2008) 3:e1759.
326. Michan S and Sinclair D. Biochem J (2007) 404:1-13.
327. Guarente L and Franklin H. N Engl J Med (2011) 364:2235-2244.
328. Kitada M, Kume S, Kanasaki K, Takeda-Watanabe A and Koya D. Curr Drug Targets (2013) 14:622-636.
329. Gomes P, Fleming Outeiro T and Cavadas C. Trends Pharmacol Sci (2015) 36:756-768.
330. McDonnell E, Peterson BS, Bomze HM and Hirschey MD. Trends Endocrinol Metab (2015) 26:486-492.
331. Kuang J, Chen L, Tang Q, Zhang J, Li Y and He J. Front Physiol (2018) 9:135.
332. Schenk S, Saberi M and Olefsky JM. J Clin Invest (2008) 118:2992-3002.
333. de Luca C and Olefsky JM. FEBS Lett (2008) 582:97-105.
334. Gerber PA and Rutter GA. Antioxid Redox Signal (2017) 26:501-518.
335. Morgan D, Oliveira-Emilio HR, Keane D, Hirata AE, Santos da Rocha M, Bordin S, Curi R, Newsholme P and Carpinelli AR. Diabetologia (2007) 50:359-369.
336. Arkan MC, Hevener AL, Greten FR, Maeda S, Li ZW, Long JM, Wynshaw-Boris A, Poli G, Olefsky J and Karin M. Nat Med (2005) 11:191-198.
337. Olefsky JM and Glass CK. Annu Rev Physiol (2010) 72:219-46.
338. Solinas G and Karin M. FASEB J (2010) 24:2596-2611.
339. Archuleta TL, Lemieux AM, Saengsirisuwan V, Teachey MK, Lindborg KA, Kim JS and Henriksen EJ. Free Radic Biol Med (2009) 47:1486-1493.
340. Evans JL, Maddux BA and Goldfine ID. Antioxid Redox Signal (2005) 7: 1040–1052.
341. Evans JL, Goldfine ID, Maddux BA and Grodsky GM. Endocr Rev (2002) 23: 599-622.

342. Bloch-Damti A and Bashan N. Antioxid Redox Signal (2005) 7:1553-1567.
343. Petersen KF, Dufour S, Befroy D, Garcia R and Shulman GI. N Engl J Med (2004) 50:664-671.
344. Morino K, Petersen KF, Dufour S, Befroy D, Frattini J, Shatzkes N, Neschen S, White MF, Bilz S, Sono S, Pypaert M and Shulman GI. J Clin Invest (2005) 115:3587-3593.
345. Morino K, Petersen KF and Shulman GI. Diabetes (2006) 55(Suppl. 2):S9-15.
346. Petersen KF, Befroy D, Dufour S, Dziura J, Ariyan C, Rothman DL, Rothman DL, DiPietro L, Cline GW and Shulman GI. Science (2003) 300:1140-1142.
347. Kitada M and Koya D. Diabetes Metab J (2013) 37:315-325.
348. Milne JC, Lambert PD, Schenk S, Carney DP, Smith JJ, Gagne DJ, Jin L, Boss O, Perni RB, Vu CB, Bemis JE, Xie R, Disch JS, Ng PY, Nunes JJ, Lynch AV, Yang H, Galonek H, Israelian K, Choy W, Iffland A, Lavu S, Medvedik O, Sinclair DA, Olefsky JM, Jirousek MR, Elliott PJ, Westphal CH. Nature (2007) 450:712-716.
349. Smoliga JM, Baur JA and Hausenblas HA. Mol Nutr Food Res (2011) 55:1129-1141.
350. Tao Y, Huang C, Huang Y, Hong L, Wang H, Zhou Z and Qiu Y. Cardiovasc Toxicol (2015) 15:217-223.
351. Lang A, Anand R, Altinoluk-Hambüchen S, Ezzahoini H, Stefanski A, Iram A, Bergmann L, Urbach J, Böhler P, Hänsel J, Franke M, Stühler K, Krutmann J, Scheller J, Stork B, Reichert AS and Piekorz RP. Aging (2017) 9:2163-2189.
352. Singh CK, Chhabra G, Ndiaye MA, Garcia-Peterson LM, Mack NJ and Ahmad N. Antioxid Redox Signal (2018) 28:643-661.
353. Zhou L, Wang F, Sun R, Chen X, Zhang M, Xu Q, Wang Y, Wang S, Xiong Y, Guan KL, Yang P, Yu H, Ye D EMBO Rep (2016) 17:811-822.
354. Wang Y, Liu Q, Huan Y, Li R, Li C, Sun S, Guo N, Yang M, Liu S and Shen Z. Exp Cell Res (2018) 371:205-213.
355. Ma Y and Fei X. Exp Ther Med (2018) 16:1417-1425.

356. Heerspink HJL, Perkins BA, Fitchett DH, Husain M and Cherney DZ. Circulation (2016) 134:752-772.
357. Vallon V and Thomson SC. Diabetologia (2017) 60:215-225.
358. Norton L, Shannon CE, Fourcaudot M, Hu C, Wang N, Ren W, Song J, Abdul-Ghani M, DeFronzo RA, Ren J and Jia W.Diabetes Obes Metab (2017) 19:1322-1326.
359. Solini A, Rossi C, Mazzanti CM, Proietti A, Koepsell H and Ferrannini E. Diabetes Obes Metab (2017) 19:1289-1294.
360. Wang XX, Levi J, Luo Y, Myakala K, Herman-Edelstein M, Qiu L, Wang D, Peng Y, Grenz A, Lucia S, Dobrinskikh E, D'Agati VD, Koepsell H, Kopp JB, Rosenberg AZ and Levi M. J Biol Chem (2017) 292:5335-5348.
361. Song P, Onishi A, Koepsell H and Vallon V. Exp Opin Ther Targets (2016) 20:1109-1125.
362. Scheen AJ. Drugs (2015) 75:33-59.
363. UK Prospective Diabetes Study (UKPDS) Group. Lancet (1998) 352:854-865.
364. Bailey C and Day C. Br J Diabetes Vasc Dis (2010) 10:193-199.
365. Day C and Bailey AV. Br J Diabetes Vasc Dis (2013) 13:2-6.
366. Vallon V and Thomson SC. Nat Rev Nephrol (2017) 13:517-518.
367. Wanner C, Inzucchi SE, Lachin JM, Fitchett D, von Eynatten M, Mattheus M, Johansen OE, Woerle HJ, Broedl UC and Zinman B. N Engl J Med (2016) 375:323-334.
368. Zinman B, Wanner C, Lachin JM, Fitchett D, Bluhmki E, Hantel S, Mattheus M, Devins T, Johansen OE, Woerle HJ, Broedl UC and Inzucchi SE. N Engl J Med (2015) 373:2117-2128.
369. Neal B, Perkovic V, Mahaffey KW, de Zeeuw D, Fulcher G, Erondu N, Shaw W, Law G, Desai M and Matthews DR. N Engl J Med (2017) 377:644-657.
370. American Diabetes Association. Diabetes Care (2018) 41:S55-S64.
371. Khunti K, Davies M, Majeed A, Thorsted BL, Wolden ML and Paul SK. Diabetes Care (2015) 38:316-322.
372. American Diabetes Association. Diabetes Care (2018) 41:S73-S85.

373. De Ferranti SD, De Boer IH, Fonseca V, Fox CS, Golden SH, Lavie CJ, Magge SN, Marx N, McGuire DK, Orchard TJ, Zinman B and Eckel RH. Diabetes Care (2014) 37:2843-2863.
374. Henry RR, Thakkar P, Tong C, Polidori D and Alba M. Diabetes Care (2015) 38:2258-2265.
375. Dandona P, Mathieu C, Phillip M, Hansen L, Griffen SC, Tschöpe D, Thorén F, Xu J and Langkilde AM. Lancet Diabetes Endocrinol (2017) 5:864-876.
376. Garg SK, Henry RR, Banks P, Buse JB, Davies MJ, Fulcher GR, Pozzilli P, Gesty-Palmer D, Lapuerta P, Simó R, Danne T, McGuire DK, Kushner JA, Peters A and Strumph P. N Engl J Med (2017) 377:2337-2348.
377. Sands AT, Zambrowicz BP, Rosenstock J, Lapuerta P, Bode BW, Garg SK, Buse JB, Banks P, Heptulla R, Rendell M, Cefalu WT and Strumph P. Diabetes Care (2015) 38:1181-1188.
378. Grempler R, Thomas L, Eckhardt M, Himmelsbach F, Sauer A, Sharp DE, Bakker RA, Mark M, Klein T and Eickelmann P. Diabetes Obes Metab (2012) 14:83-90.
379. Ogawa W and Sakaguchi K. J Diabetes Investig (2016) 7:135-138.
380. Qiu H, Novikov A and Vallon V. Diabetes Metab Res Rev (2017) 33:e2886.
381. Peters AL, Buschur EO, Buse JB, Cohan P, Diner JC and Hirsch IB. Diabetes Care (2015) 38:1687-1693.
382. US Food and Drug Administration. FDA Drug Saf Commun (2015) 2-4.
383. Fadini GP, Bonora BM and Avogaro A. Diabetologia (2017) 60:1385-1389.
384. Blau JE, Tella SH, Taylor SI and Rother KI. Diabetes Metab Res Rev (2017) 33:e2924.
385. Jerums G, Premaratne E, Panagiotopoulos S and MacIsaac RJ. Diabetologia (2010) 53:2093-2104.
386. Magee GM, Bilous RW, Cardwell CR, Hunter SJ, Kee F and Fogarty DG. Diabetologia (2009) 52:691-697.
387. Vallon V, Blantz RC and Thomson S. J Am Soc Nephrol (2003) 14:530-537.
388. Vallon V, Richter K, Blantz RC, Thomson S and Osswald H. J Am Soc

Nephrol (1999) 10:2569-2576.

389. Thomson SC, Rieg T, Miracle C, Mansoury H, Whaley J, Vallon V and Singh P. AJP Regul Integr Comp Physiol (2012) 302:R75-83.

390. Vallon V, Rose M, Gerasimova M, Satriano J, Platt KA, Koepsell H, Cunard R, Sharma K, Thomson SC and Rieg T. AJP Ren Physiol (2013) 304:F156-167.

391. Vallon V and Thomson SC. Annu Rev Physiol (2012) 74:351-375.

392. Vallon V, Gerasimova M, Rose MA, Masuda T, Satriano J, Mayoux E, Koepsell H, Thomson SC and Rieg T.AJP Ren Physiol (2014) 306:F194-204.

393. Cherney DZI, Perkins BA, Soleymanlou N, Maione M, Lai V, Lee A, Fagan NM, Woerle HJ, Johansen OE, Broedl UC and von Eynatten M. Circulation (2014) 129:587-597.

394. Scheen AJ and Delanaye P. Diabetes Metab (2017) 43:99-109.

395. Scheen AJ. Diabetes Res Clin Pract (2016) 121:204-214.

396. Bolinder J, Ljunggren O, Kullberg J, Johansson L, Wilding J, Langkilde AM, Sugg J and Parikh S. J Clin Endocrinol Metab (2012) 97:1020-1031.

397. Yokono M, Takasu T, Hayashizaki Y, Mitsuoka K, Kihara R, Muramatsu Y, Miyoshi S, Tahara A, Kurosaki E, Li Q, Tomiyama H, Sasamata M, Shibasaki M, Uchiyama Y. Eur J Pharmacol (2014) 727:66-74.

398. Pessoa TD, Campos LCG, Carraro-Lacroix L, Girardi AC and Malnic G. J Am Soc Nephrol (2014) 25:2028-2039.

399. Layton AT and Vallon V. Am J Physiol Renal Physiol (2018) 314:F969-F984.

400. Ferrannini E, Mark M and Mayoux E. Diabetes Care (2016) 39:1108-1114.

401. Morgan NG and Dhayal S. Biochem Pharmacol (2009) 78:1419-1427.

402. Civelli O. Trends Pharmacol Sci (2005) 26:15-19.

403. Oh DY and Lagakos WS. Curr Opin Clin Nutr Metab Care (2011) 14: 322-327.

404. Yonezawa T, Kurata R, Yoshida K, Murayama MA, Cui X and Hasegawa A.

Curr Med Chem (2013) 20:3855-3871.

405. Talukdar S, Olefsky JM and Osborn O. Trends Pharmacol Sci (2011) 32:543-550.

406. Mohammad S. Curr Drug Targets (2015) 16:771-775.

407. Negoro N, Sasaki S, Mikami S, Ito M, Suzuki M, Tsujihata Y, Ito R, Harada A, Takeuchi K, Suzuki N, Miyazaki J, Santou T, Odani T, Kanzaki N, Funami M, Tanaka T, Kogame A, Matsunaga S, Yasuma T and Momose Y. ACS Med Chem Lett (2010) 1:290-294.

408. Vettor R, Granzotto M, De Stefani D, Trevellin E, Rossato M, Farina MG, Milan G, Pilon C, Nigro A, Federspil G, Vigneri R, Vitiello L, Rizzuto R, Baratta R and Frittitta L. J Clin Endocrinol Metab (2008) 93:3541-3550.

409. Kebede M, Alquier T, Latour MG, Semache M, Tremblay C and Poitout V. Diabetes (2008) 57:2432-2437.

410. Edfalk S, Steneberg P and Edlund H. Diabetes (2008) 57:2280-2287.

411. Schnell S, Schaefer M and Schofl C. Mol Cell Endocrinol (2007) 263:173-180.

412. Tomita T, Masuzaki H, Noguchi M, Iwakura H, Fujikura J, Tanaka T, Ebihara K, Kawamura J, Komoto I, Kawaguchi Y, Fujimoto K, Doi R, Shimada Y, Hosoda K, Imamura M and Nakao K. Biochem Biophys Res Commun (2005) 338:1788-1790.

413. Steneberg P, Rubins N, Bartoov-Shifman R, Walker MD and Edlund H. Cell Metab (2005) 1:245-258.

414. Fujiwara K, Maekawa F and Yada T. Am J Physiol Endocrinol Metab (2005) 289:E670-677.

415. Briscoe CP, Tadayyon M, Andrews JL, Benson WG, Chambers JK, Eilert MM, Ellis C, Elshourbagy NA, Goetz AS, Minnick DT, Murdock PR, Sauls HR Jr, Shabon U, Spinage LD, Strum JC, Szekeres PG, Tan KB, Way JM, Ignar DM, Wilson S and Muir AI. J Biol Chem (2003) 278:11303-11311.

416. Itoh Y, Kawamata Y, Harada M, Kobayashi M, Fujii R, Fukusumi S, Ogi K, Hosoya M, Tanaka Y, Uejima H, Tanaka H, Maruyama M, Satoh R, Okubo S,

Kizawa H, Komatsu H, Matsumura F, Noguchi Y, Shinohara T, Hinuma S, Fujisawa Y and Fujino M. Nature (2003) 422:173-176.

417. Nagasumi K, Esaki R, Iwachidow K, Yasuhara Y, Ogi K, Tanaka H, Nakata M, Yano T, Shimakawa K, Taketomi S, Takeuchi K, Odaka H and Kaisho Y. Diabetes (2009) 58:1067-1076.
418. Briscoe CP, Peat AJ, McKeown SC, Corbett DF, Goetz AS, Littleton TR, McCoy DC, Kenakin TP, Andrews JL, Ammala C, Fornwald JA, Ignar DM, Jenkinson S. Br J Pharmacol (2006) 148:619-628.
419. Tomita T, Hosoda K, Fujikura J, Inagaki N and Nakao K. Front Endocrinol (Lausanne) (2014) 5:152.
420. Xiong Y, Swaminath G, Cao Q, Yang L, Guo Q, Salomonis H, Lu J, Houze JB, Dransfield PJ, Wang Y, Liu JJ, Wong S, Schwandner R, Steger F, Baribault H, Liu L, Coberly S, Miao L, Zhang J, Lin DC and Schwarz M. Mol Cell Endocrinol (2013) 369:119-129.
421. Parker HE, Habib AM, Rogers GJ, Gribble FM and Reimann F. Diabetologia (2009) 52:289-298.
422. Habib AM, Richards P, Rogers GJ, Reimann F and Gribble FM. Diabetologia (2013) 56:1413-1416.
423. Yamashima T. Prog Lipid Res (2012) 51:221-231.
424. Boneva NB, Kikuchi M, Minabe Y and Yamashima T. J Pharmacol Sci (2011) 116:163-172.
425. Boneva NB and Yamashima T. Hippocampus (2012) 22:896-905.
426. Ma D, Zhang M, Larsen CP, Xu F, Hua W, Yamashima T, Mao Y and Zhou L. Brain Res (2010) 1330:1-8.
427. Ma D, Tao B, Warashina S, Kotani S, Lu L, Kaplamadzhiev DB, Mori Y, Tonchev AB and Yamashima T. Neurosci Res (2007) 58:394-401.
428. Khan MZ and He L. Neuropharmacology 2015.
429. Yamashima T. Prog Lipid Res (2015) 58:40-50.
430. Nakamoto K, Nishinaka T, Sato N, Mankura M, Koyama Y and Tokuyama S. Yakugaku Zasshi (2014) 134:397-403.

431. Nakamoto K and Tokuyama S. Nihon Shinkei Seishin Yakurigaku Zasshi (2012) 32:233-237.

432. Honore JC, Kooli A, Hamel D, Alquier T, Rivera JC, Quiniou C, Hou X, Kermorvant-Duchemin E, Hardy P, Poitout V and Chemtob S. Arterioscler Thromb Vasc Biol (2013) 33:954-961.

433. Skov V, Glintborg D, Knudsen S, Tan Q, Jensen T, Kruse TA, Beck-Nielsen H and Højlund K. PLoS One (2008) 3:e2466.

434. Skov V, Glintborg D, Knudsen S, Jensen T, Kruse TA, Tan Q, Brusgaard K, Beck-Nielsen H and Højlund K. Diabetes (2007) 56:2349-2355.

435. Timmers S, Konings E, Bilet L, Houtkooper RH, van de Weijer T, Goossens GH, Hoeks J, van der Krieken S, Ryu D, Kersten S, Moonen-Kornips E, Hesselink MKC, Kunz I, Schrauwen-Hinderling VB, Blaak E, Auwerx J and Schrauwen P. Cell Metab (2011) 14:612-622.

436. Tateno C, Miya F, Wake K, Kataoka M, Ishida Y, Yamasaki C, Yanagi A, Kakuni M, Wisse E, Verheyen F, Inoue K, Sato K, Kudo A, Arii S, Itamoto T, Asahara T, Tsunoda T and Yoshizato K. Lab Invest (2013) 93:54-71.

437. Affo S, Dominguez M, Lozano JJ, Sancho-Bru P, Rodrigo-Torres D, Morales-Ibanez O, Moreno M, Millán C, Loaeza-del-Castillo A, Altamirano J, García-Pagán JC, Arroyo V, Ginès P, Caballería J, Schwabe RF and Bataller R. Gut (2013) 62:452-460.

438. van Dijk SJ, Feskens EJ, Bos MB, Hoelen DW, Heijligenberg R, Bromhaar MG, de Groot LC, de Vries JH, Müller M and Afman LA. Am J Clin Nutr (2009) 90:1656-1664.

439. She X, Rohl CA, Castle JC, Kulkarni AV, Johnson JM and Chen R. BMC Genomics (2009) 10:269.

440. Hardy OT, Perugini RA, Nicoloro SM, Gallagher-Dorval K, Puri V, Straubhaar J and Czech MP. Surg Obes Relat Dis (2011) 7:60-67.

441. Keller P, Gburcik V, Petrovic N, Gallagher IJ, Nedergaard J, Cannon B and Timmons JA. BMC Endocr Disord (2011) 11:7.

442. Wu HT, Chen W, Cheng KC, Ku PM, Yeh CH and Cheng JT. J Nutr Biochem

(2012) 23:1264-1270.

443. Suh HN, Huong HT, Song CH, Lee JH, Han HJ. Am J Physiol Cell Physiol (2008) 295:C1518-1527.

444. Ou HY, Wu HT, Hung HC, Yang YC, Wu JS and Chang CJ. Am J Physiol Endocrinol Metab (2013) 304:E668-676.

445. Lin DC, Zhang J, Zhuang R, Li F, Nguyen K, Chen M, Tran T, Lopez E, Lu JY, Li XN, Tang L, Tonn GR, Swaminath G, Reagan JD, Chen JL, Tian H, Lin YJ, Houze JB and Luo J. PLoS One (2011) 6:e27270.

446. Houze JB, Zhu L, Sun Y, Akerman M, Qiu W, Zhang AJ, Sharma R, Schmitt M, Wang Y, Liu J, Liu J, Medina JC, Reagan JD, Luo J, Tonn G, Zhang J, Lu JY, Chen M, Lopez E, Nguyen K, Yang L, Tang L, Tian H, Shuttleworth SJ and Lin DC. Bioorg Med Chem Lett (2012) 22:1267-1270.

447. Yazaki R, Kumagai N and Shibasaki M. Org Lett (2011) 13:952-955.

448. Tan CP, Feng Y, Zhou YP, Eiermann GJ, Petrov A, Zhou C, Lin S, Salituro G, Meinke P, Mosley R, Akiyama TE, Einstein M, Kumar S, Berger JP, Mills SG, Thornberry NA, Yang L and Howard AD. Diabetes (2008) 57:2211-2219.

449. Song F, Lu S, Gunnet J, Xu JZ, Wines P, Proost J, Liang Y, Baumann C, Lenhard J, Murray WV, Demarest KT and Kuo GH. J Med Chem (2007) 50:2807-2817.

450. Stoddart LA, Brown AJ and Milligan G. Mol Pharmacol (2007) 71:994-1005.

451. Tanaka H, Yoshida S, Minoura H, Negoro K, Shimaya A, Shimokawa T and Shibasaki M. Life Sci (2014) 94:115-121.

452. Fujita T, Matsuoka T, Honda T, Kabashima K, Hirata T and Narumiya S. J Invest Dermatol (2011) 131:1660-1667.

453. Christiansen E, Hansen SV, Urban C, Hudson BD, Wargent ET, Grundmann M, Jenkins L, Zaibi M, Stocker CJ, Ullrich S, Kostenis E, Kassack MU, Milligan G, Cawthorne MA and Ulven T. ACS Med Chem Lett (2013) 4: 441-445.

454. Brown SP, Dransfield PJ, Vimolratana M, Jiao X, Zhu L, Pattaropong V, Sun Y, Liu J, Luo J, Zhang J, Wong S, Zhuang R, Guo Q, Li F, Medina JC,

Swaminath G, Lin DC and Houze JB. ACS Med Chem Lett (2012) 3:726-730.

455. Gowda N, Dandu A, Singh J, Biswas S, Raghav V, Lakshmi MN, Shilpa PC, Sunil V, Reddy A, Sadasivuni M, Aparna K, Verma MK, Moolemath Y, Anup MO, Venkataranganna MV, Somesh BP and Jagannath MR. BMC Pharmacol Toxicol (2013) 14:28.

456. Doshi LS, Brahma MK, Sayyed SG, Dixit AV, Chandak PG, Pamidiboina V, Motiwala HF, Sharma SD and Nemmani KV. Metabolism 2009; 58(3):333-343.

457. Ito R, Tsujihata Y, Matsuda-Nagasumi K, Mori I, Negoro N and Takeuchi K. Br J Pharmacol (2013) 170:568-580.

458. Lead GPR40 agonist bites the dust. Nat Rev Drug Discov [News and Analysis] 2014; 13(2): 91-.

459. Srivastava A, Yano J, Hirozane Y, Kefala G, Gruswitz F, Snell G, Lane W, Ivetac A, Aertgeerts K, Nguyen J, Jennings A and Okada K. Nature (2014) 513:124-127.

460. Luo J, Swaminath G, Brown SP, Zhang J, Guo Q, Chen M, Nguyen K, Tran T, Miao L, Dransfield PJ, Vimolratana M, Houze JB, Wong S, Toteva M, Shan B, Li F, Zhuang R and Lin DC. PLoS One (2012) 7:e46300.

461. Tanaka H, Yoshida S, Oshima H, Minoura H, Negoro K, Yamazaki T, Sakuda S, Iwasaki F, Matsui T and Shibasaki M. J Pharmacol Exp Ther (2013) 346:443-452.

462. Zhao YF, Wang L, Zha D, Qiao L, Lu L, Yu J, Qu P, Sun Q, Qiu J and Chen C. J Mol Endocrinol (2013) 51:69-77.

463. Philippe C, Wauquier F, Leotoing L, Coxam V and Wittrant Y. Exp Cell Res (2013) 319:3035-3041.

464. Christiansen E, Urban C, Merten N, Liebscher K, Karlsen KK, Hamacher A, Spinrath A, Bond AD, Drewke C, Ullrich S, Kassack MU, Kostenis E and Ulven T. J Med Chem (2008) 51:7061-7064.

465. Christiansen E, Due-Hansen ME, Urban C, Grundmann M, Schröder R,

Hudson BD, Milligan G, Cawthorne MA, Kostenis E, Kassack MU and Ulven T. J Med Chem (2012) 55:6624-6628.

466. Urban C, Hamacher A, Partke HJ, Roden M, Schinner S, Christiansen E, Due-Hansen ME, Ulven T, Gohlke H and Kassack MU. Naunyn Schmiedebergs Arch Pharmacol (2013) 386:1021-1030.

467. Wagner R, Kaiser G, Gerst F, Christiansen E, Due-Hansen ME, Grundmann M, Machicao F, Peter A, Kostenis E, Ulven T, Fritsche A, Häring HU, Ullrich S. Diabetes (2013) 62:2106-2111.

468. Naik H, Vakilynejad M, Wu J, Viswanathan P, Dote N, Higuchi T and Leifke E. J Clin Pharmacol (2012) 52:1007-1016.

469. Leifke E, Naik H, Wu J, Viswanathan P, Demanno D, Kipnes M, Vakilynejad M. Clin Pharmacol Ther (2012) 92:29-39.

470. Kaku K, Araki T and Yoshinaka R. Diabetes Care (2013) 36:245-250.

471. Burant CF, Viswanathan P, Marcinak J, Cao C, Vakilynejad M, Xie B, Leifke E. Lancet (2012) 379:1403-1411.

472. Araki T, Hirayama M, Hiroi S and Kaku K. Diabetes Obes Metab (2012) 14:271-278.

473. Molecule of the Month. TAK-875. Drug News Perspect (2010) 23:667.

474. Frumento D, Ben Nasr M, El Essawy B, D'Addio F, Zuccotti GV and Fiorina P. J Endocr Invest (2017) 40:803-814.

475. van Belle TL, Coppieters KT, and von Herrath MG. Physiol Rev (2011) 91:79-118.

476. Pociot F and Lernmark A. Lancet (2106), 387:2331-2339.

477. Rewers M and Ludvigsson L. Lancet (2016) 387:2340-2348.

478. Leete P, Willcox A, Krogvold L, Dahl-Jørgensen K, Foulis AK, Richardson SJ and Morgan NG. Diabetes (2016) 65:1362-1369.

479. Krogvold L, Wiberg A, Edwin B, Buanes T, Jahnsen FL, Hanssen KF, Larsson E, Korsgren O, Skog O and Dahl-Jørgensen K. Diabetologia (2016) 59:492–501.

480. Coppieters KT, Dotta F, Amirian N, Campbell PD, Kay TW, Atkinson MA,

Roep BO and von Herrath MG. J Exp Med (2012) 209:51–60.

481. Krogvold L, Skog O, Sundstrom G, Edwin B, Buanes T, Hanssen KF, Ludvigsson J, Grabherr M, Korsgren O and Dahl-Jørgensen K. Diabetes (2015) 64:2506-2512.

482. Cobo-Vuilleumier N, Lorenzo PI, Rodriguez NJ, Herrera Gómez IG, Fuente-Martin E, López-Noriega L, Mellado-Gil JM, Romero-Zerbo SY, Baquié M, Lachaud CC, Stifter K, Perdomo G, Bugliani M, De Tata V, Bosco D, Parnaud G, Pozo D, Hmadcha A, Florido JP, Toscano MG, de Haan P, Schoonjans K, Sánchez Palazón L, Marchetti P, Schirmbeck R, Martín-Montalvo A, Meda P, Soria B, Bermúdez-Silva FJ, St-Onge L and Gauthier BR. Nature Commun (2018) 9:1488.

483. Atkinson MA, Roep BO, Posgai A, Wheeler DCS and Peakman M. Lancet Diab Endocr (2019) 7:52-64. .

484. Menendez-Gutierrez MP, Roszer T and Ricote M. Curr Topics Med Chem (2012) 12:548-584.

485. Tyagi S, Gupta P, Saini AS, Kaushal C and Sharma S. J Adv Pharm Technol Res (2011) 2:236–240.

486. Umemoto T and Fujiki Y. Genes to Cells (2012) 17:576–596.

487. Le Menn G and Neels JG. Int J Mol Sci (2018) 19:1575. .

488. Capelli D, Cerchia C, Montanari R, Loiodice F, Tortorella P, Laghezza A, Cervoni L, Pochetti G and Lavecchia A. Sci Rep (2016) 6:34792.

489. Poulsen L, Siersbaek M and Mandrup S. Sem Cell Dev Biol (2012) 23:631-639.

490. Zoete V, Grosdidier A and Michielin O. Biochimica Biophysica Acta (2007) 1771:915-925.

491. Nettles KW and Greene GL. Ann Rev Physiol (2005) 67:309-333.

492. Ricote M and Glass CK. Biochim Biophys Acta (2007) 1771: 926-935.

493. Torchia J, Glass C and Rosenfeld MG. Curr Opin Cell Biol (1998) 10:373-383.

494. Yessoufou A and Wahli W. Swiss Medical Weekly (2010) 140:w13071.

495. Mandard S, Muller M and Kersten S. Cell Mol Life Sci (2004) 61:393-416.
496. Zingarelli B, Piraino G, Hake PW, O'Connor M, Denenberg A, Fan H and Cook JA. Am J Pathol (2010) 177:1834-1847.
497. Fajas L, Auboeuf D, Raspe E Schoonjans K, Lefebvre AM, Saladin R, Najib J, Laville M, Fruchart JC, Deeb S, Vidal-Puig A, Flier J, Briggs MR, Staels B, Vidal H and Auwerx J. J Biol Chem (1997) 272:18779-18789.
498. Ricote M, Huang JT, Welch JS and Glass CK. J Leukoc Biol (1999) 66:733-739.
499. Odegaard JI, Ricardo-Gonzalez RR, Goforth MH, Morel CR, Subramanian V, Mukundan L, Red Eagle A, Vats D, Brombacher F, Ferrante AW, Chawla A. Nature (2007) 447:1116-1120.
500. Willson TM, Brown PJ, Sternbach DD and Henke BR. J Med Chem (2000) 43:527-550.
501. Racke MK, Gocke AR, Muir M, Diab A, Drew PD and Lovett-Racke AE. J Nutr (2006) 136:700-703.
502. Aprahamian TR, Bonegio RG, Weitzner Z, Gharakhanian R and Rifkin IR. Immunology (2014) 142:363-373.
503. Fallahi P, Ferrari SM, Elia G, Nasini F, Colaci M, Giuggioli D, Vita R, Benvenga S, Ferri C and Antonelli A. Expert Rev Clin Pharmacol (2016) 9:853-861.
504. Pawlak-Adamska E, Daroszewski J, Bolanowski M, Oficjalska J, Janusz P, Szalinski M and Frydecka I. Immunogenetics (2013) 65:493-500.
505. Okamoto H, Iwamoto T, Kotake S, Momohara S, Yamanaka H and Kamatani N. Clin Exp Rheumatol (2005) 23:323-330.
506. Lima EA, de Andrade Lima MMD, Marques CDL, Duarte ALBP, da Rocha Pita I and da Rocha Pita MG. An Bras Dermatol (2013) 88:1029-1035.
507. Ramkalawan H, Wang YZ, Hurbungs A, Yang YF, Tian FF, Zhou WB, Li J, Yang H, Xiao B and Zhang W. Inflammation (2012) 35:1338-1347.
508. Varga T, Czimmerer Z and Nagy L. BBA-Mol Basis Dis (2011) 1812:1007-1022.

509. Beeson PB. Am J Med (1994) 96:457-462.
510. Dunn SE, Ousman SS, Sobel RA, Zuniga L, Baranzini SE, Youssef S, Crowell A, Loh J, Oksenberg J and Steinman L. J Exp Med (2007) 204:321-330.
511. Holm LJ, Krogvold K, Hasselby JP, Kaur S, Claessens LA, Russell MA, Mathews CE, Hanssen KF, Morgan NG, Koeleman BPC, Roep BO, Gerling IC, Pociot F, Dahl-Jørgensen K and Buschard K. Diabetologia (2018) 61:1650-1661.
512. Janani C and Ranjitha Kumari BD, Diab Metab Syndr (2015) 9: 46-50.
513. Hull CM, Peakman M and Tree TI. Diabetologia (2017) 60:1839-1850.
514. Szablewski L. Int Immunopharmacol (2014) 22:182-191.
515. Bloem SJ and Roep BO. Diabetologia (2017) 60:1185-1189.
516. Willcox A, Richardson SJ, Bone AJ, Foulis AK and Morgan NG. Clin Exp Immunol (2009) 155:173-181.
517. Choi JM and Bothwell AL. Mol Cell (2012) 33:217-222.
518. Eizirik DL, Sammeth M, Bouckenooghe T, Bottu G, Sisino G, Igoillo-Esteve M, Ortis F, Santin I, Colli ML, Barthson J, Bouwens L, Hughes L, Gregory L, Lunter G, Marselli L, Marchetti P, McCarthy MI and Cnop M. PLoS Genet (2012) 8:e1002552.
519. Ravnskjaer K, Frigerio F, Boergesen M, Nielsen T, Maechler P and Mandrup S. J Lipid Res (2010) 51:1370-1379.
520. Dillon JS, Yaney GC, Zhou Y, Voilley N, Bowen S, Chipkin S, Bliss CR, Schultz V, Schuit FC, Prentki M, Waxman DJ and Corkey BE. Diabetes (2000) 49:2012-2020.
521. Gremlich S, Nolan C, Roduit R, Burcelin R, Peyot ML, Delghingaro-Augusto V, Desvergne B, Michalik L, Prentki M and Wahli W. Endocrinology (2005) 146:375-382.
522. Ravnskjaer K, Boergesen M, Dalgaard LT and Mandrup S. J Mol Endocrinol (2006) 36:289-299.
523. Bihan H, Rouault C, Reach G, Poitout V, Staels B and Guerre-Millo M. FEBS

Lett (2005) 579:2284-2288.

524. Tang T, Abbott MJ, Ahmadian M, Lopes AB, Wang Y and Sul HS. Cell Metabol (2013) 18:883-895.

525. Iglesias J, Barg S, Vallois D, Lahiri S, Roger C, Yessoufou A, Pradevand S, McDonald A, Bonal C, Reimann F, Gribble F, Debril MB, Metzger D, Chambon P, Herrera P, Rutter GA, Prentki M, Thorens B and Wahli W. J Clin Investig (2012) 122:4105-4117.

526. Nakamichi Y, Kikuta T, Ito E, Ohara-Imaizumi M, Nishiwaki C, Ishida H and Nagamatsu S. Biochem Biophys Res Commun (2003) 306:832-836.

527. Bollheimer LC, Troll S, Landauer H, Wrede CE, Scholmerich J and Buettner R. J Mol Endocrinol (2003) 31:61-69.

528. Ito E, Ozawa S, Takahashi K, Tanaka T, Katsuta H, Yamaguchi S, Maruyama M, Takizawa M, Katahira H, Yoshimoto K, Nagamatsu S and Ishida H. Biochem Biophys Res Commun (2004) 324:810-814.

529. Ravnskjaer K, Boergesen M, Rubi B, Larsen JK, Nielsen T, Fridriksson J, Maechler P and Mandrup S. Endocrinol (2005) 146:3266-3276.

530. Wang X, Zhou L, Shao L, Qian L, Fu X, Li G, Luo T, Gu Y, Li F, Li J, Zheng S and Luo M. Life Sci (2007) 81:160-165.

531. Gupta D, Jetton TL, Mortensen RM, Duan SZ, Peshavaria M and Leahy JL. J Biol Chem (2008) 283:32462-32470.

532. Shimabukuro M, Zhou YT, Lee Y and Unger RH. J Biol Chem (1998) 273:3547-3550.

533. Higa M, Zhou YT, Ravazzola M, Baetens D, Orci L and Unger RH. Proc Natl Acad Sci (1999) 96:11513-11518.

534. Finegood DT, McArthur MD, Kojwang D, Thomas MJ, Topp BG, Leonard T and Buckingham RE. Diabetes (2001) 50:1021-1029.

535. Jia DM, Tabaru A, Nakamura H, Fukumitsu KI, Akiyama T and Otsuki M. Metabolism (2000) 49:1167-1175.

536. Jia DM and Otsuki M. Pancreas (2002) 24:303-312.

537. Matsui J, Terauchi Y, Kubota N, Takamoto I, Eto K, Yamashita T, Komeda

K, Yamauchi T, Kamon J, Kita S, Noda M and Kadowaki T. Diabetes (2004) 53:2844-2854.

538. Kim EK, Kwon KB, Koo BS, Han MJ, Song MY, Song EK, Han MK, Park JW, Ryu DG and Park BH. Int J Biochem Cell Biol (2007) 39:1260-1275.

539. Shen X, Yang L, Yan S, Zheng H, Liang L, Cai X and Liao M. Mol Cell Endocrinol (2015) 412:1-11.

540. Lin CY, Gurlo T, Haataja L, Hsueh WA and Butler PC. J Endocrinol Metab (2005) 90:6678-6686.

541. Hull RL, Shen Z-P, Watts MR, Kodama K, Carr DB, Utzschneider KM, Zraika S, Wang F and Kahn SE. Diabetes (2005) 54:2235-2244.

542. Akiyama M, Hatanaka M, Ohta Y, Ueda K, Yanai A, Uehara Y, Tanabe K, Tsuru M, Miyazaki M, Saeki S, Saito T, Shinoda K, Oka Y and Tanizawa Y. Diabetologia (2009) 52:653-663.

543. Evans-Molina C, Robbins RD, Kono T, Tersey SA, Vestermark GL, Nunemaker CS, Garmey JC, Deering TG, Keller SR, Maier B and Mirmira RG. Mol Cell Biol (2009) 29:2053-2067.

544. Pearson JA, Wong FS and Wen L. J Autoimmun (2016) 66:76-88.

545. Reed JC and Herold KC. Nature Rev Endocrinol (2015) 11:308-314.

546. Yaacob NS, Goh KS and Norazmi MN. Exp Toxic Pathol (2012) 64:127-131.

547. Yaacob NS, Kaderi MA and Norazmi MN. J Clin Immunol (2009) 29:595-602.

548. Beales PE and Pozzilli P. Diab Metab Res Rev (2002) 18:114-117.

549. Ogawa J, Takahashi S, Fujiwara T, Fukushige J, Hosokawa T, Izumi T, Kurakata S and Horikoshi H. Life Sci (1999) 65:1287-1296.

550. Fu Z, Zhen W, Yuskavage J and Liu D. Br J Nutr (2011) 105:1218-1225.

551. Zhang S, Yang X, Luo J, Ge X, Sun W, Zhu H, Zhang W, Cao J and Hou Y. Nutr Cancer (2014) 66:315-324.

552. Castro CN, Barcala Tabarrozzi AL, Winnewisser J, Gimeno ML, Antunica Noguerol M, Liberman AC, Paz DA, Dewey RA and Perone MJ. Clin Exp Immunol (2014) 177:149-160.

553. Mazidi M, Karimi E, Meydani M, Ghayour-Mobarhan M and Ferns GA. World J Methodol (2016) 6:112-117.
554. Weiss L, Zeira M, Reich S, Har-Noy M, Mechoulam R, Slavin S and Gallily R. Autoimmun (2006) 39:143-151.
555. O'Sullivan SE. Br J Pharmacol (2016) 173:1899-1910.
556. Bi X, Li F, Liu S, Jin Y, Zhang X, Yang T, Dai Y, Li X and Zhao AZ. J Endocrinol Invest (2017) 127:1757-1771.
557. Nevius E, Srivastava PK and Basu S. Mucosal Immunol (2012) 5:76-86.
558. Kang JH, Goto T, Han IS, Kawada T, Kim YM and Yu R. Obesity (2010) 18:780-787.
559. Arany E, Strutt B, Romanus P, Remacle C, Reusens B and Hill DJ. Diabetologia (2004) 47:1831-1837.
560. Bonfleur ML, Borck PC, Ribeiro RA, Caetano LC, Soares GM, Carneiro EM and Balbo SL. Life Sci (2015) 135:15-21.
561. Soares FL, de Oliveira ML, Teixeira LG, Menezes Z, Pereira SS, Alves AC, Batista NV, de Faria AM, Cara DC, Ferreira AV and Alvarez-Leite JI. J Nutr Biochem (2013) 24:1105-1111.
562. Funda DP, Kaas A, Bock T, Tlaskalova-Hogenova H and Buschard K. Diab Metabol Res Rev (1999) 15:323-327.
563. Antvorskov JC, Josefsen K, Haupt-Jorgensen M, Fundova P, Funda DP and Buschard K. J Diabet Res (2016) 2016:7.
564. Haupt-Jorgensen M, Larsen J, Josefsen K, Jørgensen TZ, Antvorskov JC, Hansen AK and Buschard K. Diab Metab Res Rev (2018) 34:e2987.
565. Scheen AJ and Lefebvre PJ. Diabetes Care (1999) 22:1568-1577.
566. Ahmad F, Li PM, Meyerovitch J and Goldstein BJ. J Biol Chem (1995) 270:20503-20508.
567. Zinker BA, Rondinone CM, Trevillyan JM, Gum RJ, Clampit JE, Waring JF, Xie N, Wilcox D, Jacobson P, Frost L, Kroeger PE, Reilly RM, Koterski S, Opgenorth TJ, Ulrich RG, Crosby S, Butler M, Murray SF, McKay RA, Bhanot S, Monia BP and Jirousek MR. Proc Nat Acad Sci USA (2002)

99:11357-11362.

568. Johnson TO, Ermolieff J, and Jirousek MR. Nat Rev Drug Discov (2002) 1:696-709.

569. Zhang S and Zhang ZY. Drug Discov Today (2007) 12:373-381.

570. Zhang ZY and Lee SY. Expert Opin Investig Drugs (2003) 12:223-233.

571. Bandyopadhyay D, Kusari A, Kenner KA, Liu F, Chernoff J, Gustafson TA and Kusari J. J Biol Chem (1997) 272:1639-1645.

572. Goldstein BJ, Bittner-Kowalczyk A, White MF and Harbeck M. J Biol Chem (2000) 275:4283-4289.

573. Lund IK, Hansen JA, Andersen HS, Moller NPH and Billestrup N. J Mol Endocrinol (2005) 34:339-351.

574. Byon JCH, Kusari AB and Kusari J. Mol Cell Biochem (1998) 182:101-108.

575. Kenner KA, Anyanwu E, Olefsky JM and Kusari J. J Biol Chem (1996) 271:19810-19816.

576. Meshkani R, Taghikhani M, Al-Kateb H, Larijani B, Khatami S, Sidiropoulos GK, Hegele RA and Adeli K. Clin Chem (2007) 53:1585-1592.

577. Elchebly M, Payette P, Michaliszy E, Cromlish W, Collins S, Loy AL, Normandin D, Cheng A, Himms-Hagen J, Chan CC, Ramachandran C, Gresser MJ, Tremblay ML and Kennedy BP.Science (1999) 283:1544-1548.

578. Klaman LD, Boss O, Peroni OD, Kim JK, Martino JL, Zabolotny JM, Moghal N, Lubkin M, Kim YB, Sharpe AH, Stricker-Krongrad A, Shulman GI, Neel BG and Kahn BB. Mol Cell Biol (2000) 20:5479-5489.

579. Bence KK, Delibegovic M, Xue B, Gorgun CZ, Hotamisligil GS, Neel BG and Kahn BB. Nat Med (2006) 12:917-924.

580. Zhao H, Liu G, Xin Z, Serby MD, Pei Z, Szczepankiewicz BG, Hajduk PJ, Abad-Zapatero C, Hutchins CW, Lubben TH, Ballaron SJ, Haasch DL, Kaszubska W, Rondinone CM, Trevillyan JM and Jirousek MR. Bioorgan Med Chem Lett (2004) 14:5543-5546.

581. Leung KW, Posner BI and Just G. Bioorg Med Chem Lett (1999) 9:353-356.

582. Andersen HS, Olsen OH, Iversen LF, Sorensen ALP, Mortensen SB,

Christensen M, Branner S, Hansen TK, Lau JF, Jeppesen L, Moran EJ, Su J, Bakir F, Judge L, Shahbaz M, Collins T, Vo T, Newman MJ, Ripka WC and Møller NP. J Med Chem (2002) 45:4443-4459.

583. Wilson DP, Wan ZK, Xu WX, Kirincich SJ, Follows BC, Joseph-McCarthy D, Foreman K, Moretto A, Wu J, Zhu M, Binnun E, Zhang YL, Tam M, Erbe DV, Tobin J, Xu X, Leung L, Shilling A, Tam SY, Mansour TS and Lee J. J Med Chem (2007) 50:4681-4698.

584. Liu S, Zeng LF, Wu L, Yu X, Xue T, Gunawan AM, Long YQ and Zhang ZY. J Am Chem Soc (2008) 130:17075-17084.

585. Jiang CS, Liang LF and Guo YW. Acta Pharmacol Sin (2012) 33:1217-1245.

586. Yamazaki H, Sumilat DA, Kanno S, Ukai K, Rotinsulu H, Wewengkang DS, Ishikawa M, Mangindaan REP and Namikoshi M. J Nat Med (2013) 67:730-735.

587. Li Y, Zhang Y, Shen X and Guo YW. Bioorg Med Chem Lett (2009) 19:390-392.

588. Fantus IG, Deragon G, Lai R and Tang S. Mol Cell Biochem (1995) 153:103-112.

589. Fantus IG and Tsiani E. Mol Cell Biochem (1998) 182:109-119.

590. Meyerovitch J, Farfel Z, Sack J and Shechter Y. J Biol Chem (1987) 262:6658-6662.

591. Meyerovitch J, Backer J and Kahn CR. J Clin Investig (1989) 84:976-983.

592. Myers MP, Andersen JN, Cheng A, Tremblay ML, Horvath CM, Parisien JP, Salmeen A, Barford D and Tonks NK. J Biol Chem (2001) 276:47771-47774.

593. Cheng A, Uetani N, Simoncic PD, Chaubey VP, Lee-Loy A, McGlade CJ, Kennedy BP and Tremblay ML. Dev Cell (2002) 2:497-503.

594. Zabolotny JM, Bence-Hanulec KK, Stricker-Krongrad A, Haj F, Wang Y, Minokoshi Y, Kim YB, Elmquist JK, Tartaglia LA, Kahn BB and Neel BG. Dev Cell (2002) 2:489-495.

595. Ahmad F, Considine RV, Bauer TL, Ohannesian JP, Marco CC and Goldstein

BJ. Metabolism (1997) 46:1140-1145.

596. Cheung A, Kusari J, Jansen D, Bandyopadhyay D, Kusari A and Bryer-Ash M. J Lab Clin Med (1999) 134:115-123.

597. Wu X, Hardy VE, Joseph JI, Jabbour S, Mahadev K, Zhu L and Goldstein BJ. Metabolism (2003) 52:705-712.

598. Brown-Shimer S, Johnson KA, Lawrence JB, Johnson C, Bruskin A, Green NR and Hill DE. Proc Natl Acad Sci USA (1990) 87:5148-5152.

599. Gerich JE. Mayo Clin Proc (2003) 78:447-456.

600. Creutzfeldt W. Diabetologia (1979) 16:75-85.

601. Nauck MA, Bartels E, Orskov C, Ebert R and Creutzfeldt W. J Clin Endocrinol Metab (1993) 76:912-917.

602. Nauck MA and Meier JJ. Diabetes Obes Metab (2018) 20(Suppl 1):5-21.

603. Nauck M, Stöckmann F, Ebert R and Creutzfeldt W. Diabetologia (1986) 29:46-52.

604. Nauck MA, Kleine N, Orskov C, Holst JJ, Willms B and Creutzfeldt W. Diabetologia (1993) 36:741-744.

605. Drucker DJ and Nauck MA. Lancet (2006) 368:1696-1705.

606. Mentlein R, Gallwitz B and Schmidt WE. Eur J Biochem (1993) 214:829-835.

607. Holst JJ and Deacon CF. Diabetes (1998) 47:1663-1670.

608. Deacon CF. Front Endocrinol (Lausanne) (2019) 10:80.

609. Drucker DJ. Cell Metab (2006) 3:153-165.

610. Sesti G, Avogaro A, Belcastro S, Bonora BM, Croci M and Daniele G. Acta Diabetol (2019) 56:605-617.

611. Tomovic K, Lazarevic J, Kocic G, Deljanin-Ilic M and Anderluh M. Med Res Rev (2019) 39:404-422.

612. Nabeno M, Akahoshi F, Kishida H, Miyaguchi I, Tanaka Y and Ishii S. Biochem Biophys Res Commun (2013) 434:191-196.

613. Patel BD and Ghate MD. Eur J Med Chem (2014) 74:574-605.

614. Kim SH, Yoo JH, LeeWJ and Park CY. Diabetes Metab J (2016) 40:339-353.

615. White WB, Bakris GL, Bergenstal RM, Cannon CP, Cushman WC, Fleck P, Heller S, Mehta C, Nissen SE, Perez A, Wilson C and Zannad F. Am Heart J (2011) 162:620-626.e1.

616. White WB, Cannon CP, Heller SR, Nissen SE, Bergenstal RM, Bakris GL, Perez AT, Fleck PR, Mehta CR, Kupfer S, Wilson C, Cushman WC and Zannad F. N Engl J Med. (2013) 369:1327-1335.

617. Mosenzon O, Raz I, Scirica BM, Hirshberg B, Stahre CI, Steg PG, Davidson J, Ohman P, Price DL, Frederich B, Udell JA, Braunwald E and Bhatt DL. Diabetes Metab Res Rev (2013) 29:417-426.

618. Scirica BM, Bhatt DL, Braunwald E, Steg PG, Davidson J, Hirshberg B Ohman P, Frederich R, Wiviott SD, Hoffman EB, Cavender MA, Udell JA, Desai NR, Mosenzon O, McGuire DK, Ray KK, Leiter LA and Raz I. N Engl J Med (2013) 369:1317-1326.

619. Green JB, Bethel MA, Paul SK, Ring A, Kaufman KD, Shapiro DR, Califf RM and Holman RR. Am Heart J (2013) 166:983-989.e7.

620. Green JB, BethelMA, Armstrong PW, Buse JB, Engel SS, Garg J, Josse R, Kaufman KD, Koglin J, Korn S, Lachin JM, McGuire DK, Pencina MJ, Standl E, Stein PP, Suryawanshi S, Van de Werf F, Peterson ED and Holman RR. N Engl J Med (2015) 373:232-242.

621. Rosenstock J, Perkovic V, Alexander JH, Cooper ME, Marx N, Pencina MJ, Toto RD, Wanner C, Zinman B, Baanstra D, Pfarr E, Mattheus M, Broedl UC, Woerle HJ, George JT, von Eynatten M and McGuire DK. Cardiovasc Diabetol (2018) 17:39.

622. Rosenstock J, Perkovic V, Johansen OE, Cooper ME, Kahn SE, Marx N, Alexander JH, Pencina M, Toto RD, Wanner C, Zinman B, Woerle HJ, Baanstra D, Pfarr E, Schnaidt S, Meinicke T, George JT, von Eynatten M and McGuire DK. JAMA (2019) 321:69-79.

623. Bailey CJ and Marx N. Diabetes Obes Metab (2019) 21:3-14.

624. Scirica BM, Braunwald E, Raz I, Cavender MA, Morrow DA, Jarolim P, Udell JA, Mosenzon O, Im K, Umez-Eronini AA, Pollack PS, Hirshberg B,

Frederich R, Lewis BS, McGuire DK, Davidson J, Steg PG, Bhatt DL. Circulation (2014) 130:1579-1588.

625. Udell JA, Bhatt DL, Braunwald E, Cavender MA, Mosenzon O, Steg PG, Davidson JA, Nicolau JC, Corbalan R, Hirshberg B, Frederich R, Im K, Umez-Eronini AA, He P, McGuire DK, Leiter LA, Raz I and Scirica BM. Diabetes Care (2015) 38:696-705.

626. McGuire DK, Van de Werf F, Armstrong PW, Standl E, Koglin J, Green JB, Bethel MA, Cornel JH, Lopes RD, Halvorsen S, Ambrosio G, Buse JB, Josse RG, Lachin JM, Pencina MJ, Garg J, Lokhnygina Y, Holman RR and Peterson ED. JAMA Cardiol (2016) 1:126-135.

627. Gallwitz B, Rosenstock J, Rauch T, Bhattacharya S, Patel S, von Eynatten M, Dugi KA and Woerle HJ. Lancet (2012) 380:475-483.

628. Marx N, Rosenstock J, Kahn SE, Zinman B, Kastelein JJ, Lachin JM, Espeland MA, Bluhmki E, Mattheus M, Ryckaert B, Patel S, Johansen OE and Woerle HJ. Diab Vasc Dis Res (2015) 12:164-174.

629. Davies MJ, D'Alessio DA, Fradkin J, Kernan WN, Mathieu C, Mingrone G, Rossing P, Tsapas A, Wexler DJ and Buse JB. Diabetes Care (2018) 41:2669-2701.

630. Davies MJ, D'Alessio DA, Fradkin J, Kernan WN, Mathieu C, Mingrone G, Rossing P, Tsapas A, Wexler DJ and Buse JB.Diabetologia (2018) 61:2461-2498.

631. American Diabetes Association. Diabetes Care (2019) 42(Suppl 1):S90-S102.

632. Inzucchi SE, Bergenstal RM, Buse JB, Diamant M, Ferrannini E, Nauck M, Peters AL, Tsapas A, Wender R and Matthews DR. Diabetologia (2012) 55:1577-1596.

633. Inzucchi SE, Bergenstal RM, Buse JB, Diamant M, Ferrannini E, Nauck M, Peters AL, Tsapas A, Wender R and Matthews DR. Diabetologia (2015) 58:429-442.

634. Nauck MA, Kahle M, Baranov O, Deacon CF and Holst JJ. Diabetes Obes Metab (2017) 19:200-207.

635. Tahrani AA, Bailey CJ, Del Prato S and Barnett AH. Lancet (2011) 378:182-197.
636. Rines AK, Sharabi K, Tavares CD and Puigserver P. Nat Rev Drug Discov (2016) 15:786-804.
637. Perseghin G. J Clin Endocrinol Metab (2010) 95:4871-4873.
638. Petersen MC, Vatner DF and Shulman GI. Nat Rev Endocrinol (2017) 13:572-587.
639. Agius L. Expert Opin Ther Pat (2014) 24:1155-1159.
640. Nakamura A and Terauchi Y. J Diabetes Investig (2015) 6:124-132.
641. Choi JM, Seo MH, Kyeong HH, Kim E and Kim HS. Proc Natl Acad Sci USA (2013) 110:10171-10176.
642. Raimondo A, Rees MG and Gloyn AL. Curr Opin Lipidol (2015) 26:88-95.
643. Matschinsky FM and Wilson DF. Front Physiol (2019) 10:148.
644. Agius L. Biochem J (2008) 414:1-18.
645. Chakera AJ, Steele AM, Gloyn AL, Shepherd MH, Shields B, Ellard S and Hattersley AT. Diabetes Care (2015) 38:1383-1392.
646. Steele AM, Shields BM, Wensley KJ, Colclough K, Ellard S and Hattersley AT. JAMA (2014) 311:279-286.
647. Amed S and Oram R. Can J Diabetes (2016) 40:449-454.
648. Ping F, Wang Z and Xiao X. J Diabetes Investig (2019) 10:1454-1462.
649. Doliba NM, Fenner D, Zelent B, Bass J, Sarabu R and Matschinsky FM. Diabetes Obes Metab (2012) 14(Suppl 3):109-119.
650. Lu M, Li P, Bandyopadhyay G, Lagakos W, Dewolf WE Jr, Alford T, Chicarelli MJ, Williams L, Anderson DA, Baer BR, McVean M, Conn M, Véniant MM and Coward P. PLoS One (2014) 9:e88431.
651. Nakamura A, Shimazaki H, Ohyama S, Eiki J and Terauchi Y. J Diabetes Investig (2011) 2:276-279.
652. Nakamura A, Terauchi Y, Ohyama S, Kubota J, Shimazaki H, Nambu T, Takamoto I, Kubota N, Eiki J, Yoshioka N, Kadowaki T and Koike T. Endocrinology (2009) 150:1147-1154.

653. Nakamura A, Togashi Y, Orime K, Sato K, Shirakawa J, Ohsugi M, Kubota N, Kadowaki T and Terauchi Y. Diabetologia (2012) 55:1745-1754.
654. Grimsby J, Sarabu R, Corbett WL, Haynes NE, Bizzarro FT, Coffey JW, Guertin KR, Hilliard DW, Kester RF, Mahaney PE, Marcus L, Qi L, Spence CL, Tengi J, Magnuson MA, Chu CA, Dvorozniak MT, Matschinsky FM and Grippo JF. Science (2003) 301:370-373.
655. Filipski KJ and Pfefferkorn JA. Expert Opin Ther Pat (2014) 24:875-891.
656. Sarabu R, Berthel SJ, Kester RF and Tilley JW. Expert Opin Ther Pat (2011) 21:13-33.
657. Pfefferkorn JA. Expert Opin Drug Discov (2013) 8:319-330.
658. Egan A and Vella A. Expert Opin Investig Drugs (2019) 28:741-747.
659. Matschinsky FM. Trends Pharmacol Sci (2013) 34:90-99.
660. Amin NB, Aggarwal N, Pall D, Paragh G, Denney WS, Le V, Riggs M and Calle RA. Diabetes Obes Metab (2015) 17:751-759.
661. Brouwers M, Jacobs C, Bast A, Stehouwer CDA and Schaper NC. Trends Mol Med (2015) 21:583-594.
662. Scheen AJ. Lancet Diabetes Endocrinol (2018) 6:591-593.
663. Tornovsky-Babeay S, Dadon D, Ziv O, Tzipilevich E, Kadosh T, Schyr-Ben Haroush R, Hija A, Stolovich-Rain M, Furth-Lavi J, Granot Z, Porat S, Philipson LH, Herold KC, Bhatti TR, Stanley C, Ashcroft FM, Veld PI, Saada A, Magnuson MA, Glaser B and Dor Y. Cell Metab (2014) 19:109-121.
664. Arden C, Petrie JL, Tudhope SJ, Al-Oanzi Z, Claydon AJ, Beynon RJ, Towle HC and Agius L. Diabetes (2011) 60:3110-3120.
665. Arden C, Tudhope SJ, Petrie JL, Al-Oanzi ZH, Cullen KS, Lange AJ, Towle HC and Agius L. Biochem J (2012) 443:111-123.
666. Baggio LL and Drucker DJ. Gastroenterology (2007) 132:2131-2157.
667. DeFronzo RA. Diabetes (2009) 58:773-795.
668. Elrick H, Stimmler L, Hlad CJ Jr and Arai Y. J Clin Endocrinol Metab (1964) 24:1076-1082.

669. Drucker DJ and Nauck MA. Lancet (2006) 368:1696-1705.
670. Meier JJ and Nauck MA. Curr Opin Investig Drugs (2004) 5:402-410.
671. Nauck M, Stockmann F, Ebert R and Creutzfeldt W. Diabetologia (1986) 29: 46-52.
672. Calanna S, Christensen M, Holst JJ, Laferrère B, Gluud LL, Vilsbøll T and Knop FK. Diabetes Care (2013) 36:3346-3352.
673. Nauck MA and Meier JJ. Lancet Diabetes Endocrinol (2016) 4:525-536.
674. Drucker DJ. Nat Clin Pract End Met (2005) 1:22-31.
675. Nauck MA, Kleine N, Orskov C, Holst JJ, Willms B and Creutzfeldt W. Diabetologia (1993) 36:741-744.
676. Zander M, Madsbad S, Madsen JL and Holst JJ. Lancet (2002) 359:824-830.
677. Gupta V. Indian J Endocrinol Metab (2013) 17:413-421.
678. Byetta (exenatide) prescribing information. Wilmington, Del., AstraZeneca, February 2015.
679. Victoza (liraglutide) prescribing information. Plainsboro, N.J., Novo Nordisk, March 2015.
680. Bydureon (exenatide extended release) prescribing information. Wilmington, Del., AstraZeneca, September 2015.
681. Tanzeum (albiglutide) prescribing information. Wilmington, Del., GlaxoSmithKline, May 2015.
682. Trulicity (dulaglutide) prescribing information. Indianapolis, Ind., Eli Lilly And Company, March 2015.
683. Adlyxin (lixisenatide) prescribing information. Bridgewater, N.J., Sanofi-Aventis, July 2016.
684. Kayaniyil S, Lozano-Ortega G, Bennett HA, Johnsson K, Shaunik A, Grandy S and Kartman B. Diabetes Ther (2016) 7:27-43.
685. Fox CS, Golden SH, Anderson C, Bray GA, Burke LE, de Boer IH, Deedwania P, Eckel RH, Ershow AG, Fradkin J, Inzucchi SE, Kosiborod M, Nelson RG, Patel MJ, Pignone M, Quinn L, Schauer PR, Selvin E and Vafiadis DK. Diabetes Care (2015) 38:1777-1803.

686. Handelsman Y, Bloomgarden ZT, Grunberger G, Umpierrez G, Zimmerman RS, Bailey TS, Blonde L, Bray GA, Cohen AJ, Dagogo-Jack S, Davidson JA, Einhorn D, Ganda OP, Garber AJ, Garvey WT, Henry RR, Hirsch IB, Horton ES, Hurley Dl, Jellinger PS, Jovanovič L, Lebovitz HE, LeRoith D, Levy P, McGill GB, Mechanick JI, Mestman JH, Moghissi ES, Orzeck EA, Pessah-Pollack R, Rosenblit PD, Vinik AI, Wyne K and Zangeneh F. Endocr Pract (2015) 21(Suppl. 1):1-87.

687. Acosta A, Camilleri M, Burton D, O'Neill J, Eckert D, Carlson P and Zinsmeister AR. Physiol Rep (2015) 3:e12610

688. U.S. Food and Drug Administration. Silver Spring, Md., U.S. Food and Drug Administration, 2008.

689. Gaede P, Vedel P, Larsen N, Jensen GV, Parving HH and Pedersen O. N Engl J Med (2003) 348:383-393.

690. Saraiva FK and Sposito AC. Cardiovasc Diabetol (2014) 13:142-156.

691. Monami M, Dicembrini I, Nardini C, Fiordelli I and Mannucci E. Diabetes Obes Metab (2014) 16:38-47.

692. Best JH, Hoogwerf BJ, Herman WH, Best JH, Hoogwerf Bj, Herman WH, Pelletier EM, Smith DB, Wenten M and Hussein MA. Diabetes Care (2011) 34:90-95.

693. Pfeffer MA, Claggett B, Diaz R, Dickstein K, Gerstein HC, Køber LV, Lawson FC, Ping L, Wei X, Lewis EF, Maggioni AP, McMurray JJV, Probstfield JL, Riddle MC, Solomon SD and Tardif J-C. N Engl J Med (2015) 373:2247-2257.

694. Marso SP, Daniels GH, Brown-Frandsen K, Kristensen P, Mann JFE, Nauck MA, Nissen SE, Pocock S, Poulter NR, Ravn LS, Steinberg WM, Stockner M, Zinman B, Bergenstal RM and Buse JB. N Engl J Med (2016) 375:311-322.

695. Holman RR, Bethel MA, George J, Sourij H, Doran Z, Keenan J, Khurmi NS, Mentz RJ, Oulhaj A, Buse JB, Chan JC, Iqbal N, Kundu S, Maggioni AP, Marso SP, Öhman P, Pencina MJ, Poulter N, Porter LE, Ramachandran A,

Zinman B and Hernandez AF. Am Heart J (2016) 174:103-110.

696. Garber AJ, Abrahamson MJ, Barzilay JI, Blonde L, Bloomgarden ZT, Bush MA, Dagogo-Jack S, DeFronzo RA, Einhorn D, Fonseca VA, Garber JR, Garvey WT, Grunberger G, Handelsman Y, Henry RR, Hirsch IB, Jellinger PS, McGill JB, Mechanick JI, Rosenblit PD and Umpierrez GE. Endocr Pract (2016) 22:84-113.

697. American Diabetes Association. Diabetes Care (2016) 39(Suppl. 1):S1-S112.

698. Cernea S and Raz I. Diabetes Care (2011) 34(Suppl. 2):S264-S271.

Chapter 3

QSAR Studies on DPP-4 Inhibitors

CHAPTER 3

QSAR STUDIES ON DPP-4 INHIBITORS

1. INTRODUCTION

Therapeutics based on Glucogon-like peptide-1 (GLP-1[a]) is among the novel and promising targets to cure type 2 diabetes [699-701]. The active and natural form of GLP-1, the incretin hormone GLP-1 (7-36), is secreted from intestinal L-cells after the intake of meals. The stimulation of insulin secretion, inhibition of glucogon release, delay in gastric emptying and promotion of β-cell trophism in intestinal L-cells are advantageous to glucose homeostasis in both the animal models and human [702, 703]. Studies revealed that GLP-1 levels are noticeably reduced in type 2 diabetics and exogenous infusion of it may lead to normal insulin response to glucose [704-706] and this fact is the basis for GLP-1 and its analogeus as novel treatments of type 2 diabetes. One such example of a GLP-1 analogue is exenatide [707, 708]. Half-maximal effective concentration of 10 pM of the most potent incretin hormone, GLP-1 (7-36), is required to show its effects on pancreatic β-cells [709].

The biological functions of GLP-1 (7-36) are exerted through circulation and binding to the GLP-1 receptor that is highly expressed in pancreatic β-cells. After secretion GLP-1 (7-36) is rapidly degraded by DPP4 (EC 3.4.14.5) to afford inactive GLP-1 (9-36) under normal physiological conditions. The apparent half-life for GLP-1 (7-36) in this quick inactivation process is 60-90s. It is evinced that due to this natural degradation mechanism less than 50% of released active GLP-1 (7-36) can reach circulation [710]. Thus, it is apparent that a DPP-4 inhibitor can prevent degradation of and lead to potentiation of GLP-1 and further improve glucose and insulin homeostasis [711, 712].

DPP-4, ubiquitously expressed throughout the body, is a nonclassical and sequence-specific serine protease. Membrane-bound DPP-4 is highly expressed in the endothelium of the capillary bed in close proximity to intestinal L-cells where secretion of GLP-1 takes place. The other form which circulates in plasma is

soluble form of DPP-4 plays a little role in the cleavage of GLP-1 [713, 714]. Vildagliptin [715], sitagliptin [716], saxagliptin [717] and alogliptin [718] are examples of small molecule DPP-4 inhibitors which have demonstrated ability to lower blood glucose and HbA1c levels and to improve glucose tolerance in type 2 diabetic patients [719].

2. MODELING STUDIES

2.1. THE IMIDAZOLOPYRIMIDINE AMIDES

Several novel series of azolopyrimidine amines, containing an aromatic or heteroaromatic group on the azolo ring, as potent and selective DPP-4 inhibitors were reported in view of medicinal chemistry efforts to discover novel scaffolds [720]. The substitution of aromatic or heteroaromatic group on the azolo ring in these compounds showed enhancement in the binding affinity to DPP-4 but displayed high levels of the human ether-à-go-go related gene (hERG) and sodium channel inhibition. As an attempt to minimize undesired hERG and sodium channel activities a novel series of imidazolopyrimidine amides as a highly potent and selective class of DPP4 inhibitors has been reported by Meng *et al*. [721]. The general structures of these analogues are shown in Figure 3.1.

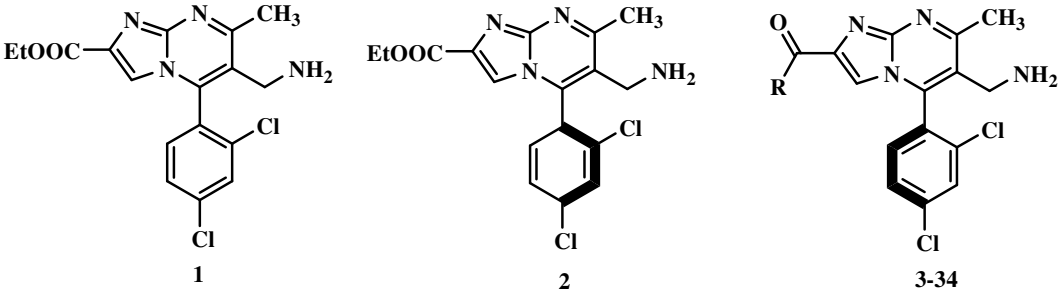
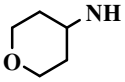
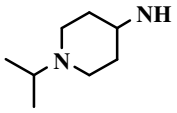
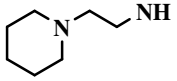
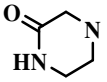
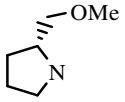
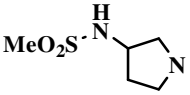
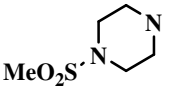
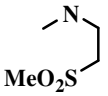
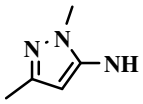
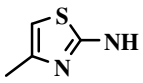
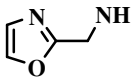
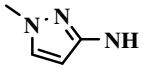
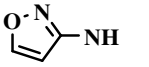
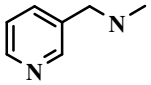
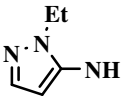
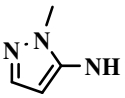
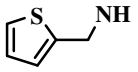
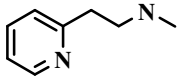
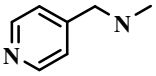
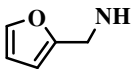
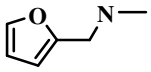
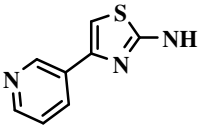


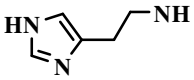
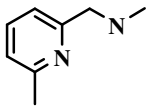
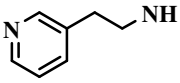
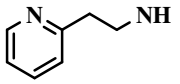
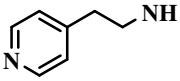
Figure 3.1: General structures of imidazolopyrimidine derivatives

The structural variations along with their reported in vitro human DPP-4 inhibition activity in terms of K_i, and expressed as pK_i on a molar basis are mentioned in Table 3.1.

Table 3.1: Structural variations and DPP-4 binding affinities of imidazolopyrimidine amides.

Cpd.	R	$pK_i(M)$[a]					
			Calculated				
		Obsd[b].	Eq. (7)	Eq. (8)	Eq. (9)	Eq. (10)	PLS
1	Racemic	9.15	8.98	8.94	9.01	9.06	8.99
2	Chiral	7.93	8.08	8.03	7.93	8.10	7.83
3[c]	OEt (chiral)	9.40	8.98	8.94	9.01	9.06	8.99
4		8.62	8.81	8.80	8.77	8.64	8.94
5		8.54	8.45	8.46	8.43	8.68	8.75
6		8.66	8.47	8.68	8.50	8.32	8.50
7[c]		8.51	8.40	8.42	8.52	8.29	8.44
8		8.30	8.26	8.30	8.18	8.26	8.22
9		8.06	8.30	8.33	8.33	8.28	8.14
10		9.00	8.82	8.95	8.95	8.72	8.81
11		8.51	8.52	8.53	8.63	8.32	8.59
12		9.70	9.52	9.55	9.49	9.59	9.56
13		9.30	9.37	9.41	9.33	9.42	9.39

14		9.52	9.70	9.54	9.76	9.56	9.56
15		9.05	8.75	8.73	8.81	8.74	8.87
16		8.18	8.43	8.25	8.43	8.56	8.21
17[c]		8.70	8.54	8.49	8.91	8.65	8.71
18		9.22	9.22	9.19	9.01	8.86	9.10
19		8.74	8.70	8.68	8.88	8.72	8.82
20		8.42	8.73	8.91	8.59	8.58	8.59
21[c]		8.59	8.49	8.67	8.63	8.40	8.58
22		8.96	8.75	8.73	8.62	9.03	8.94
23		8.70	8.74	8.93	8.53	8.97	8.83
24		8.49	8.44	8.27	8.48	8.56	8.51
25[c]		9.15	9.18	8.98	9.04	8.78	8.85
26		8.59	8.46	8.65	8.61	8.47	8.60
27[c]		8.4	8.75	8.56	8.69	8.67	8.50
28		8.52	8.70	8.51	8.86	8.64	8.56
29		8.30	8.56	8.55	8.32	8.67	8.27

30		8.33	8.06	8.22	8.38	8.30	8.48
31		8.27	8.57	8.40	8.61	8.30	8.30
32		8.82	8.68	8.67	8.68	8.83	8.81
33		8.85	8.68	8.50	8.61	8.55	8.56
34[c]		8.89	8.68	8.67	8.68	8.76	8.77

[a]On molar basis; [b]Taken from reference [721]; [c]Compounds in test set.

2.1.1. RESULTS AND DISCUSSION

2.1.1.1. QSAR RESULTS

For the compounds in Table 3.1, a total number of 479 descriptors belonging to 0D- to 2D- classes of DRAGON have been computed and were subjected to CP-MLR analysis. All the 34 compounds of data set were further divided into training-set and test-set. Seven compounds (nearly 20% of total population) have been selected for test-set. The selected test-set was then used for external validation of models derived from remaining twenty seven compounds in the training-set. The squared correlation coefficient between the observed and predicted values of compounds from test-set, r^2_{Test}, was calculated to explain the fraction of explained variance in the test-set which is not part of regression/model derivation. It is a measure of goodness of the derived model equation. A high r^2_{Test} value is always good. But considering the stringency of test-set procedures, often r^2_{Test} values in the range of 0.5 to 0.6 are regarded as logical models. Following the strategy to explore only predictive models, CP-MLR resulted, one model in three descriptors, five models in four descriptors and sixteen models in five descriptors at upper limit of filter-1. The highest significant of them, in statistical sense, are given through Equations (3.1) to (3.10):

pKi = 7.583 + 1.556(0.241)Ms + 0.716(0.215)BELe5 + 0.729 (0.241)MATS2p

n = 27, r = 0.812, s = 0.271, F = 14.904, q^2_{LOO} = 0.515,

q^2_{L5O} = 0.504, r^2_{Test} = 0.125 **(3.1)**

pKi = 7.267 + 1.503(0.229)Ms + 0.300(0.121)PJI2 + 0.985(0.256)BELv3

+ 0.738(0.223)MATS2p

n = 27, r = 0.844, s = 0.254, F = 13.731, q^2_{LOO} = 0.565,

q^2_{L5O} = 0.516, r^2_{Test}= 0.632 **(3.2)**

pKi = 8.043 – 0.677(0.253)Mv + 2.308(0.381)Ms – 0.880(0.255)nDB

+ 0.635(0.248)BELv3

n = 27, r = 0.836, s = 0.260, F = 12.856, q^2_{LOO} = 0.524,

q^2_{L5O} = 0.537, r^2_{Test} = 0.527 **(3.3)**

pKi = 7.935 + 2.678(0.448)Ms – 0.806(0.249)nDB – 0.619 (0.246)IC1

+ 0.575(0.199)BELe5

n = 27, r = 0.832, s = 0.263, F = 12.380, q^2_{LOO} = 0.507,

q^2_{L5O} = 0.594, r^2_{Test} = 0.525 **(3.4)**

pKi = 8.031 + 1.239(0.235)Ms + 0.315(0.127)PJI2 + 0.991(0.272)BELv3

– 0.679(0.232)GATS2v

n = 27, r = 0.831, s = 0.264, F = 12.295, q^2_{LOO} = 0.501,

q^2_{L5O} = 0.513, r^2_{Test} = 0.501 **(3.5)**

pKi = 8.080 + 1.325(0.227)MAXDN – 0.556(0.219)BELm8

+ 0.706(0.345)BELv3 + 0.957(0.256)MATS2p

n = 27, r = 0.821, s = 0.270, F = 11.449, q^2_{LOO} = 0.503,

q^2_{L5O} = 0.525, r^2_{Test} = 0.513 **(3.6)**

pKi = 7.360 + 1.914(0.288)Ms – 0.744(0.185)nDB + 0.214(0.101)PJI2

+ 1.110(0.215)BELv3 + 0.765(0.192)JGI2

n = 27, r = 0.901, s = 0.210, F = 18.170, q^2_{LOO} = 0.651,

q^2_{L5O} = 0.600, r^2_{Test} = 0.548 **(3.7)**

$$pKi = 7.644 + 1.753(0.277)Ms - 0.774(0.188)nDB + 1.059(0.211)BELv3 + 0.729(0.196)JGI2 - 0.357(0.171)C\text{-}028$$
$$n = 27,\ r = 0.900,\ s = 0.211,\ F = 18.042,\ q^2_{LOO} = 0.667,\ q^2_{L5O} = 0.690,\ r^2_{Test} = 0.579 \quad \textbf{(3.8)}$$

$$pKi = 6.629 + 1.573(0.194)Ms + 0.331(0.102)PJI2 + 0.798(0.249)BEHv1 + 0.705(0.232)BELv3 + 0.603(0.191)MATS2p$$
$$n = 27,\ r = 0.898,\ s = 0.213,\ F = 17.656,\ q^2_{LOO} = 0.644,\ q^2_{L5O} = 0.696,\ r^2_{Test} = 0.616 \quad \textbf{(3.9)}$$

$$pKi = 8.310 + 1.327(0.208)Ms + 0.265(0.107)PJI2 + 0.669(0.172)BELe5 - 0.723(0.208)MATS6e - 0.771(0.199)nCrHR$$
$$n = 27,\ r = 0.883,\ s = 0.227,\ F = 14.972,\ q^2_{LOO} = 0.569,\ q^2_{L5O} = 0.550,\ r^2_{Test} = 0.511 \quad \textbf{(3.10)}$$

where n and F represent respectively the number of data points and the F-ratio between the variances of calculated and observed activities. The data within the parentheses are the standard errors associated with regression coefficients. In all above equations, the F-values remained significant at 99% level. The indices q^2_{LOO} and q^2_{L5O} (> 0.5) have accounted for their internal robustness. For all above models except equation (3.1) the r^2_{Test} values, obtained greater than 0.5, specified that the selected test-set is fully accountable for their external validation. The descriptors, in all above models, have been scaled between the intervals 0 to 1 to ensure that a descriptor will not dominate simply because it has larger or smaller pre-scaled value compared to the other descriptors. In this way, the scaled descriptors would have equal potential to influence the QSAR models. The signs of the regression coefficients have indicated the direction of influence of explanatory variables in above models. The positive regression coefficient associated to a descriptor will augment the activity profile of a compound while the negative coefficient will cause detrimental effect to it.

Though Equations (3.1) to (3.10) emerged as significant predictive models but Equations (3.7) to (3.10) remained statistically more efficient. The later four models, involving five descriptors in each, could estimate up to 81.22 percent of variance in observed activity of the compounds. In fact, a total number of sixteen such models, sharing 19 descriptors among them, have been obtained through CP-MLR and the most significant four of them have been documented through Equation (3.7) to (3.10). The shared 19 descriptors along with their brief description, average regression coefficients and total incidences are given in Table 3.2.

Table 3.2: Identified descriptors[a] along with their physical meaning, average regression coefficient and incidence[b], in modeling the DPP-4 binding affinity.

S. No.	Descriptor	Descriptor class	Physical meaning	Average regression coefficient (incidence)
1	Ms	Constitutional	Mean electrotopological state	1.835(15)
2	nDB	Constitutional	Number of double bonds	-0.685(11)
3	HNar	Topological	Narumi harmonic topological index	-1.304(2)
4	PJI2	Topological	2D Petitijean shape index	0.280(7)
5	IC1	Topological	information content index of neighborhood symmetry of 1-order	-0.585(1)
6	BELm8	BCUT	Lowest eigenvalue n.8 of Burden matrix/ weighted by atomic masses	-0.479(1)
7	BEHv1	BCUT	Highest eigenvalue n.1 of Burden matrix/ weighted by atomic van der Waals volumes	0.989(2)
8	BELv3	BCUT	lowest eigenvalue n.3 of Burden matrix/ weighted by atomic van der Waals volumes	0.935(12)
9	BELe5	BCUT	lowest eigenvalue n.5 of Burden matrix/ weighted by atomic Sanderson electronegativities	0.653(6)
10	JGI2	Galvez topological charge indices	Mean topological charge index of order 2	0.747(2)
11	MATS5e	2D autocorrelations	Moran autocorrelation of lag-5/ weighted by atomic Sanderson	-0.597(1)

			electronegativities	
12	MATS6e	2D autocorrelations	Moran autocorrelation of lag-6/ weighted by atomic Sanderson electronegativities	-0.594(2)
13	MATS2p	2D autocorrelations	Moran autocorrelation of lag-2/ weighted by atomic polarizabilities	0.694(5)
14	GATS2v	2D autocorrelations	Geary autocorrelation of lag-2/ weighted by atomic van der Waals volumes	-0.794(1)
15	nCp	Functional	Number of total primary C (sp3)	0.274(1), -0.612(1)
16	nCrHR	Functional	Number of ring tertiary C(sp3)	-0.771(1)
17	nNR2	Functional	Number of tertiary aliphatic amines	-0.702(1)
18	C-028	Atom-centered fragments	R--CR--X	-0.464(6)
19	C-032	Atom-centered fragments	X--CX--X	-0.458(2)

[a]The descriptors are identified from the five parameter models, emerged from CP-MLR protocol with filter-1 as 0.79, filter-2 as 2.0, filter-3 as 0.813, and filter-4 as $0.3 \leq q^2 \leq 1.0$ with a training set of 27 compounds. [b]The average regression coefficient of the descriptor corresponding to all models and the total number of its incidence. The arithmetic sign of the coefficient represents the actual sign of the regression coefficient in the models.

Besides listed descriptors in Table 3.2, the other identified descriptors Mv is from constitutional and MAXDN is from topological class. The Mv represents mean atomic van der waals volume (scaled on carbon atom) (Equation 3.3) and MAXDN is maximal electrotopological negative variation (Equation 3.6). The further discussion is, however, based on the highest significant Equations (3.7)-(3.10). The derived statistical parameters of these four models have shown that these models are significant. These models were, therefore, used to calculate the activity profiles of all the compounds and are included in Table 3.1 for the sake of comparison with observed ones. A close agreement between them has been observed. Additionally, the graphical display, showing the variation of observed versus calculated activities is given in Figure 3.2 to ensure the goodness of fit for each of these four models.

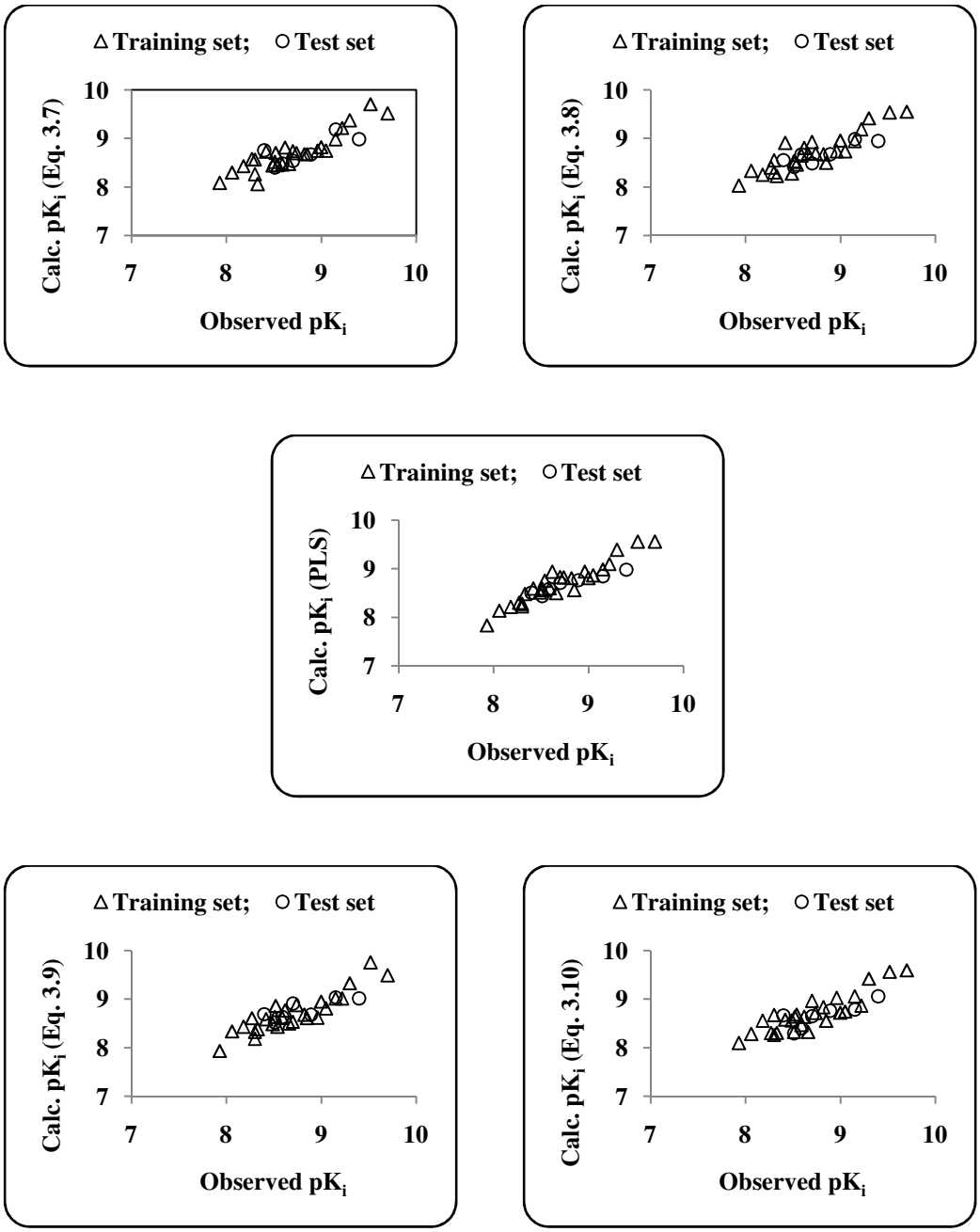


Figure 3.2: Plot of observed versus caculated pK_i values for training- and test-set compounds.

Descriptors Ms (mean electrotopological state) and nDB (number of double bonds in molecular structure) belong to constitutional class. From the sign

of regression coefficients it is evident that higher value of mean electrotopological state (descriptor Ms) and lower number of double bonds (descriptor nDB) are helpful to augment the activity. The descriptor PJI2 participated in these models is topological class descriptor and represents 2D Petitijean shape index. The positive sign of regression coefficient of descriptor PJI2 suggest that a higher value of this descriptor is beneficiary to the DPP-4 inhibition activity.

The descriptors MATS2p (Moran autocorrelation of lag-2/weighted by atomic polarizabilities) and MATS6e (Moran autocorrelation of lag-6/weighted by atomic Sanderson electronegativities) are 2D autocorrelation descriptors. It is evinced from the models mentioned above that the descriptor MATS2p contributed positively and descriptor MATS6e negatively to the activity. Thus a higher value of descriptor MATS2p and a lower value of descriptor MATS6e will be supportive to enhance the inhibition activity. The participated descriptors BELe5 (lowest eigenvalue n.5 of Burden matrix/weighted by atomic Sanderson electronegativities), BELv3 (lowest eigenvalue n.3 of Burden matrix/weighted by atomic van der Waals volumes) and BEHv1 (highest eigenvalue n.1 of Burden matrix/weighted by atomic van der Waals volumes) belong to BCUT class. All these descriptors contributed positively to the activity suggesting that higher value of these will augment the activity.

From Equations (3.7) to (3.10), it appeared that the descriptors nCrHR, a functional group accounting descriptor representing number of ring tertiary C(sp3) functionality in a structure and atom centered fragment accounting descriptor C-028 showing R--CR--X type fragment in a molecular structure make negative contribution to activity and JGI2, mean Galvez topological charge index of order 2 shown positive correlation to the activity. In this way absence of number of ring tertiary C(sp3) functionality along with R--CR--X type fragment in a molecular structure and higher value of mean Galvez topological charge index of order 2 would be advantageous in improving the DPP-4 inhibition activity of a compound.

To corroborate the study further, a PLS analysis has also been carried out on 19 descriptors identified through CP-MLR. For this purpose, the descriptors have been autoscaled (zero mean and unit s.d.) to give each one of them equal weight in the analysis. In the PLS cross-validation, three components have been found to be the optimum for these 19 descriptors and they explained 89.7% variance in the activity (r^2 = 0.897). The results of the PLS analysis and the MLR-like PLS coefficients of these 19 descriptors are given in Table 3.3.

Table 3.3: PLS and MLR-like PLS models from the descriptors of five parameter CP-MLR models for DPP-4 binding affinity.

A: PLS equation			
PLS components		PLS coefficient (s.e.)[a]	
Component-1		0.196(0.015)	
Component-2		-0.113(0.019)	
Component-3		0.078(0.023)	
	Constant	8.693	
B: MLR-like PLS equation			
S. No.	Descriptor	MLR-like coefficient (f.c.)[b]	Order
1	Ms	0.311(0.109)	2
2	nDB	0.045(0.016)	17
3	HNar	-0.157(-0.55)	10
4	PJI2	0.184(0.064)	9
5	IC1	0.115(0.040)	13
6	BELm8	-0.083(-0.029)	14
7	BEHv1	0.318(0.111)	1
8	BELv3	0.215(0.075)	3
9	BELe5	0.210(0.073)	4
10	JGI2	0.130(0.045)	12
11	MATS5e	-0.077(-0.027)	15

12	MATS6e	-0.131(-0.045)	11
13	MATS6p	-0.020(-0.007)	18
14	GATS2v	-0.055(-0.019)	16
15	nCp	0.011(0.004)	19
16	nCrHR	-0.208(-0.072)	5
17	nNR2	-0.184(-0.064)	8
18	C-028	-0.187(-0.065)	7
19	C-032	-0.205(-0.071)	6
	Constant	7.909	

C: PLS regression statistics	Values
n	27
r	0.947
s	0.148
F	67.415
q^2_{LOO}	0.851
q^2_{L5O}	0.860
r^2_{Test}	0.662

[a]Regression coefficient of PLS factor and its standard error. [b]Coefficients of MLR-like PLS equation in terms of descriptors for their original values; f.c. is fraction contribution of regression coefficient, computed from the normalized regression coefficients obtained from the autoscaled (zero mean and unit s.d.) data.

The calculated activity values of training- and test-set compounds are in close agreement to that of the observed ones and are listed in Table 3.1. For the sake of comparison, the plot between observed and calculated activities (through PLS analysis) for the training- and test-set compounds is also given in Figure 3.2. Figure 3.3 shows a plot of the fraction contribution of normalized regression coefficients of these descriptors to the activity (Table 3.3).

The top ten descriptors in decreasing order of significance are BEHv1, Ms, BELv3, BELe5, nCrHR, C-032, C-028, nNR2, PJI2 and HNar (Table 3.3, Figure 3.3).

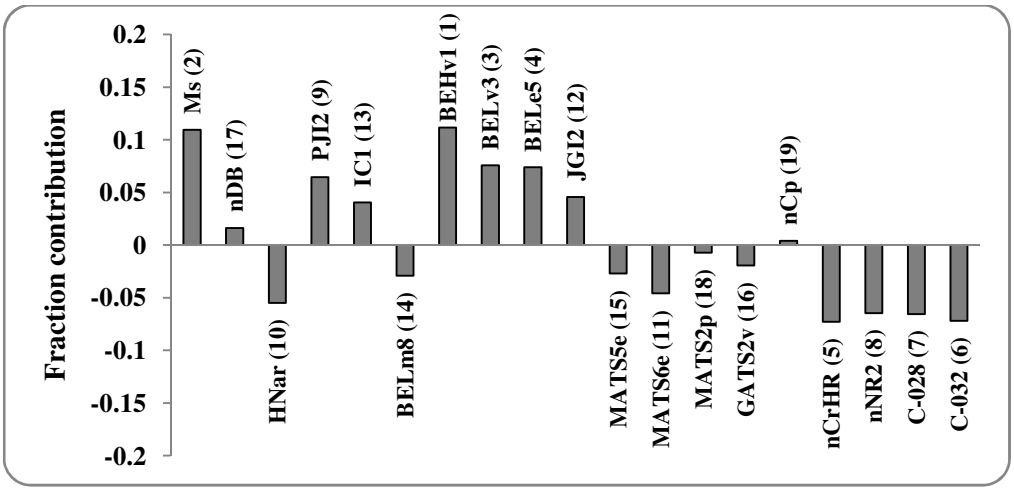


Figure 3.3: Plot of fraction contribution of MLR-like PLS coefficients (normalized) against 19 identified descriptors (Table 3.3) associated with DPP-4 binding affinity of the compounds.

Among these descriptors, BEHv1, Ms, BELv3, BELe5, nCrHR, C-028 and PJI2 are part of Equations discussed above and convey same inferences in PLS analysis. The negative contributions of atom centered fragment descriptor C-032 (X--CX--X type fragment), functional group count descriptor nNR2 (number of tertiary aliphatic amine functionality in a molecule) and toplogical descriptor HNar (Narumi harmonic topological index) advocated lower value of these are helpful in improving the activity profile. It is also observed that PLS model from the dataset devoid of 19 descriptors (Table 3.3) remained inferior in explaining the activity of the analogues.

2.1.1.2. APPLICABILITY DOMAIN (AD)

On analyzing the applicability domain (AD) in the Williams plot (Figure 3.4) of the model based on the whole data set (Table 3.4), no any compound has been identified as an obvious 'outlier' for the DPP-4 inhibitory activity if the limit of normal values for the Y outliers (response outliers) was set as 3×(standard deviation) units. One of the compound (**2**; Table 3.1) was found to have leverage

(h) values greater than the threshold leverage (h*); suggesting it as chemically influential compound.

For both the training-set and test-set, the suggested model matches the high quality parameters with good fitting power and the capability of assessing external data. Furthermore, all of the compounds were within the applicability domain of the proposed model and were evaluated correctly.

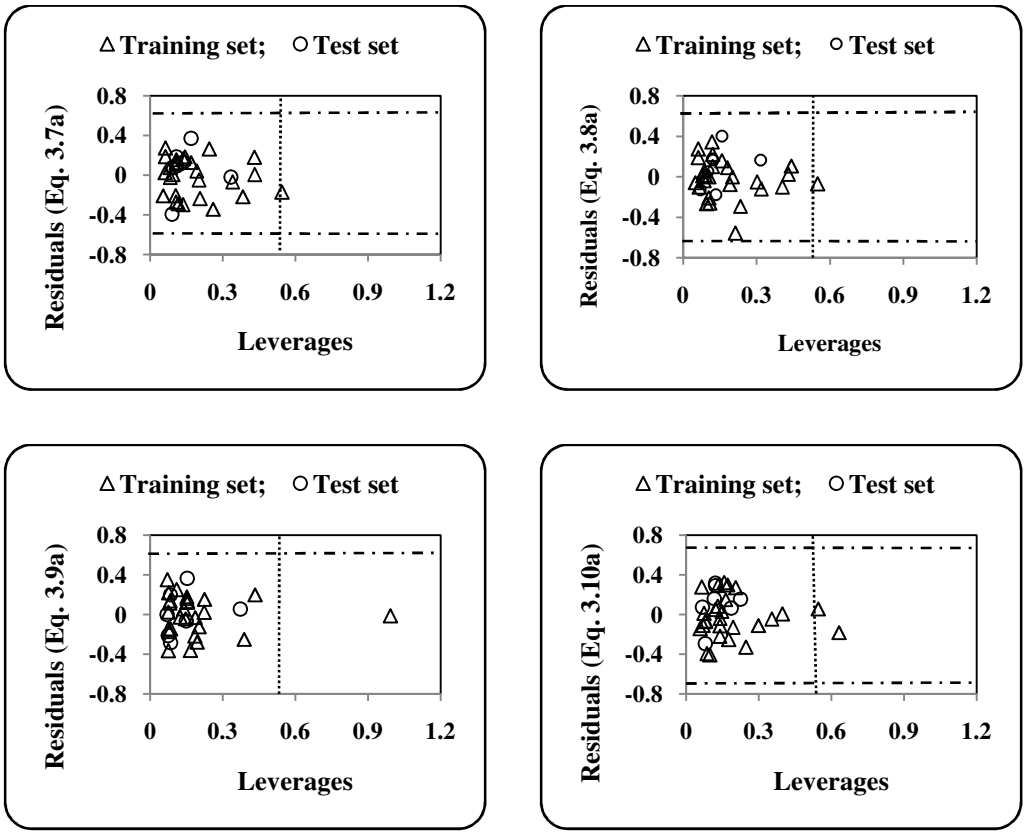


Figure 3.4: Williams plot for the training-set and test- set for inhibition activity of DPP4 for the compounds in Table 3.1. The horizontal dotted line refers to the residual limit (±3×standard deviation) and the vertical dotted line represents threshold leverage h* (= 0.529).

Table 3.4: Models derived for the whole data set (n = 34) for the DPP-4 binding affinity in descriptors identified through CP-MLR.

Model	r	s	F	Eq.
pKi = 7.366 +1.980(0.276)Ms –0.762(0.176)nDB +0.231(0.092)PJI2 +1.073(0.178)BELv3 +0.731(0.181)JGI2	0.880	0.216	19.331	(3.7a)
pKi = 7.653 +1.858(0.265)Ms –0.844(0.176)nDB +1.081(0.175)BELv3 +0.694(0.180)JGI2 –0.426(0.156)C-028	0.884	0.212	20.176	(3.8a)
pKi = 6.571 +1.600(0.184)Ms +0.356(0.092)PJI2 +0.764(0.241)BEHv1 +0.777(0.180)BELv3+0.698(0.171)MATS2p	0.884	0.213	20.026	(3.9a)
pKi = 8.402 + 1.269(0.194)Ms +0.295(0.098)PJI2 +0.685(0.159)BELe5 –0.834(0.187)MATS6e –0.792(0.185)nCrHR	0.865	0.228	16.727	(3.10a)

2.1.2. CONCLUSIONS

The DPP4 inhibition activity of imidazolopyrimidine amides has been quantitatively analyzed in terms of chemometric descriptors. The statistically validated quantitative structure-activity relationship (QSAR) models provided rationales to explain the inhibition activity of these congeners. The descriptors identified through combinatorial protocol in multiple linear regression (CP-MLR) analysis have highlighted the role of mean electrotopological state (Ms), number of double bonds in molecular structure (nDB), 2D Petitijean shape index (PJI2), Moran autocorrelation of lag-2/ weighted by atomic polarizabilities (MATS2p), Moran autocorrelation of lag-6/weighted by atomic Sanderson electronegativities (MATS6e), lowest eigenvalue n.5 of Burden matrix/ weighted by atomic Sanderson electronegativities (BELe5), lowest eigenvalue n.3 of Burden matrix/ weighted by atomic van der Waals volumes (BELv3), highest eigenvalue n.1 of Burden matrix/ weighted by atomic van der Waals volumes (BEHv1). In addition to these 2nd order mean Galvez topological charge index (JGI2), number of ring

tertiary C(sp3) (nCrHR) and R--CR--X type structural fragments (C-028) have also shown prevalence to model the inhibitory activity.

From statistically validated models, it appeared that the descriptors Ms, PJI2, JGI2, MATS2p, BELe5, BELv3 and BEHv1 make positive contribution to activity and their higher values are conducive in improving the DPP4 inhibition activity of a compound. On the other hand, the descriptors nDB, C-028, nCrHR and MATS6e render detrimental effect to activity. Therefore, absence or lower number of double bonds (nDB), R--CR--X type structural fragment (C-028), number of ring tertiary C(sp3) (nCrHR) and lower value of descriptor MATS6e would be advantageous. Such guidelines may be helpful in exploring more potential analogues of the series. The statistics emerged from the test sets have validated the identified significant models. PLS analysis has further confirmed the dominance of the CP-MLR identified descriptors. Applicability domain analysis revealed that the suggested models have acceptable predictability. All the compounds are within the applicability domain of the proposed models and were evaluated correctly.

2.2. THE (2*S*)-CYANOPYRROLIDINE ANALOGUES

DPP-4 is a serine protease, able to cleave the N-terminal dipeptide having preference for L-proline or L-alanine at the penultimate position [722-725]. A large number of DPP-IV inhibitors resemble the P2-P1 dipeptidyl substrate cleavage product. The simplest inhibitors are the compounds which are not having a carbonyl functionality of the proline residue, e.g., aminoacyl pyrrolidines and thiazolidines, possessing moderate inhibition activity for DPP-4. Replacement of hydrogen with an electrophilic nitrile group at the 2-position of the pyrrolidine, in some compounds, elicited a 1000-fold increase in potency compared to the unsubstituted pyrrolidines [726].

One of the potent and stable representatives of the nitrile class is cyclohexylglycine-(2*S*)-cyanopyrrolidine, having a K_i value of 1.4 nM and an excellent chemical stability $t_{1/2}$ ~ 48 h at pH 7.4 [727]. Another class, similar to

proline inhibitors, was synthesized with diverse N-substituted glycines in the P2 site [715]. In this class, the side chain was moved from the α-carbon to the terminal nitrogen, led to two potent derivatives which have showed greater efficacy in clinical trial [728]. From this study, it was concluded that (2*S*)-cyanopyrrolidine derivatives with *N*-substituted glycine in the P2 site are more selective for DPP-IV than α-carbon-substituted glycine.

An interesting study has recently been reported to develop a new pharmacophore in the P2 site with N-substituted glycine [729]. Initially, the P2 site amine extension was designed using β-alanine as building block and it was coupled the C-terminal with various substituted amines to generate a novel pharmacophore in the P2 site. Then, the N-terminal of the β-alanine derivative was combined with the P1 site α-bromoacetyl (2*S*)-cyanopyrrolidide to design 2-[3-[[2-[(2*S*)-2-cyano-1-pyrrolidinyl]-2-oxoethyl]amino]-1-oxopropyl]-based DPP-IV inhibitors. The structure-activity relationships of several series (I–III) of these DPP-4 inhibitors were explored to discover the potent and selective DPP-4 inhibitors. Series I, II, and III, being N-substituted glycine derivatives include, respectively, the bicyclic ring system, monocyclic piperazine ring, and phenylalkyl groups. These compounds were tested for inhibition of DPP-4, DPP-8, and DPP-2. The activity was evaluated in terms of the concentration of a compound required to bring out 50% inhibition of the enzyme concerned.

The reported twenty five (2*S*)-cyanopyrrolidine analogues, belonging to series I, II, and III are considered to formulate the data set for present study [729].

Since the activity variation for DPP-2 is very small, therefore, inhibition profiles for DPP-4 and DPP-8 have only been considered for quantitative analysis. The data set has been further, divided into training and test sets. One fifth of the compounds, from this data set, have been included in the test set for the validation of derived models while remaining compounds were used to derive the model correlating biological activity with descriptors unfolding molecular structures.

The test-set, containing 5 compounds out of the 25 active ones, was generated using in house randomization program. In this way, the identified test set will further ensure the statistical significance and reasonable predictability of derived models. As the leave-one-out (LOO) procedure has been applied to each model, therefore, corresponding to test set the derived model would be validated both internally and externally.

The general structures of series (I–III) are depicted in Figure 3.5. The structural variations, the reported activity values (expressed as IC_{50}(nM), and training and test set compounds are given in Table 3.5.

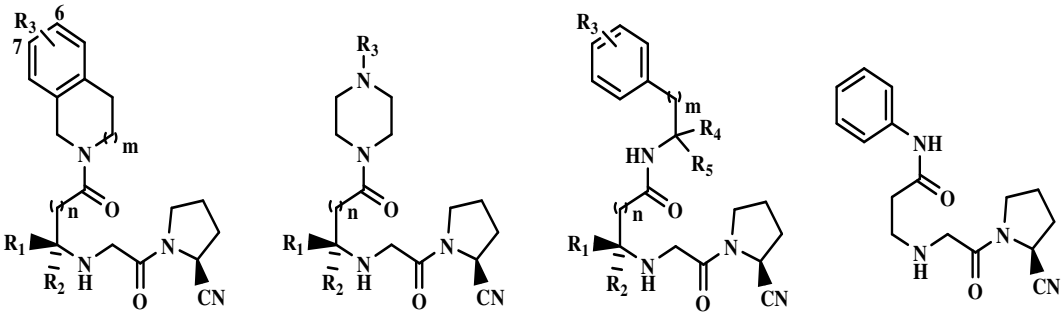


Figure 3.5: General structures of series (I–III)

Table 3.5: Structural variations and reported activities of (2*S*)-Cyanopyrrolidines.

Cpd.	n	m	R_1	R_2	R_3	R_4	R_5	IC_{50}(nM)[a]	
								DPP4	DPP8
1	0	1	H	H	H			3236	4169
2	1	1	H	H	H			116	3583
3	1	1	H	H	6,7-diOMe			651	3340
4[b]	1	1	H	H	6-F			83	1700
5	1	0	H	H	H			132	2121
6	2	1	H	H	H			428	1407
7	1	1	Me	H	H			54	5346
8[b]	1	1	*i*-Pr	H	H			811	41859
9	1	1	Me	Me	H			49	$>10^5$
10	1	1	Me	Me	6-F			30	$>10^5$

11	1	1	Me	Me	6,8-F_2			22	$>10^5$
12	1	0	Me	Me	H			15	$>10^5$
13	1		H	H	CO(3,5-$F_2C_6H_3$)			676	202
14	1		H	H	$SO_2C_6H_4$-4-$NHCOCH_3$			418	3416
15	1		H	H	nicotinonitrile			629	2000
16	1		H	H	benzothiazole			527	2117
17	1	0	H	H	H	H	H	452	10744
18	1	0	H	H	4-NO_2	H	H	317	2387
19[b]	1	0	H	H	H	H	Et[c]	447	21961
20	1	0	H	H	3,5-F_2	H	H	369	5532
21	1	0	H	H	H	H	*i*-Pr[c]	784	12847
22[b]	1	0	H	H	H	Me	Me	119	8338
23	1	0	Me	Me	H	Me	Me	1108	$>10^5$
24	1	1	H	H	H	H	H	564	2592
25[b]								298	855

[a]IC_{50} represents the concentration of a compound required to bring out 50% inhibition of DPP-IV and DPP8, taken from ref [729]; [b]compound of test set; [c]The stereochemistry at the benzylic carbon is not defined (mixture of diastereomers).

Before the application of CP-MLR procedure, all those descriptors which are intercorrelated beyond 0.90 (descriptor vs. descriptor, $r > 0.90$) and showing a correlation of less than 0.1 with the biological endpoints (descriptor vs. activity, $r < 0.1$) were excluded. This has reduced the total dataset of the compounds from 484 to 90 and 84 descriptors as relevant ones for the DPP-IV and DPP8 inhibitory activity, respectively.

2.2.1. RESULTS AND DISCUSSION

2.2.1.1. QSAR RESULTS

Initially, the DPP-IV inhibition activity of titled compounds was investigated with a variety of 0D-, 1D- and 2D-descriptors obtained from DRAGON software. The models, in three parameters of the descriptor pool of 90 descriptors, emerged in CP-MLR for the DPP-IV inhibitory actions are tabulated in Table 3.6 as Equations (3.11) to (3.16). The signs of the regression coefficients have indicated the direction of influence of explanatory variables in above models. The positive regression coefficient associated to a descriptor will

augment the activity profile of a compound while the negative coefficient will cause detrimental effect to it.

Table 3.6: CP-MLR models[a] derived in three parameters for the DPP-IV inhibition activity.

Model	r	s	F	r^2_{Test}	Eq.
pIC_{50} = 5.090 + 1.709(0.280)JGI4 – 1.150(0.331)ATS8p + 2.074(0.297)GATS8p	0.934	0.242	36.897	0.279	(**3.11**)
pIC_{50} = 5.538 + 0.599(0.248)BELm1 – 0.902(0.364)GATS7e + 1.784(0.358)GATS8p	0.854	0.353	14.465	0.331	(**3.12**)
pIC_{50} = 7.174 – 1.493(0.444)Uindex + 1.200(0.347)JGI4 – 1.492(0.442)MATS3e	0.850	0.358	13.927	0.250	(**3.13**)
pIC_{50} = 5.939 – 0.976(0.400)GATS7e + 1.442(0.410)GATS8p + 0.668(0.325)C-002	0.842	0.367	13.021	0.331	(**3.14**)
pIC_{50} = 6.094 – 0.663(0.327)RBN – 0.725(0.361)GATS7e + 1.848(0.374)GATS8p	0.841	0.368	12.913	0.191	(**3.15**)
pIC_{50} = 5.651 + 1.109(0.418)JGI4 + 1.124(0.524)MATS8e – 0.806(0.400)GATS7e	0.804	0.404	9.812	0.233	(**3.16**)

[a]The models, in three parameters, emerged from CP-MLR protocol with filter-1 as 0.79, filter-2 as 2.0, filter-3 as 0.5 and filter-4 as $0.3 \leq q^2 \leq 1.0$ with a training set of 20 compounds.

The maximum number of descriptors, participated in these models ATS8p, GATS8p, GATS7e, MATS3e and MATS8e, belong to 2D-autocorrelations (2D-AUTO) class. Descriptors, GATS8p and MATS8e both added positively to the inhibitory activity whereas ATS8p, GATS7e and MATS3e contributed negatively to the activity advocating that higher values of descriptors GATS8p and MATS8e and lower values of descriptors ATS8p, GATS7e and MATS3e would be beneficiary to the activity. Constitutional class descriptors are dimensionless or 0D descriptors and are independent from molecular connectivity and

conformations. The appeared constitutional class descriptor RBN (number of rotatable bonds) favors the least preference of rotatable bonds.

Descriptor Uindex, corresponds to Balaban U-index, is a topological class descriptor. Topological class descriptors are based on a graph representation of the molecule and are numerical quantifiers of molecular topology obtained by the application of algebraic operators to matrices representing molecular graphs and whose values are independent of vertex numbering or labeling. They can be sensitive to one or more structural features of the molecules such as size, shape, symmetry, branching and cyclicity and can also encode chemical information concering atom type and bond multiplicity. The negative contribution of descriptor Uindex suggested that a lower value of it would be supportive to the activity. The other participated descriptors are JGI4 (from the Galvez topological charge indices), BELm1 (from the modified Burden eigenvalues class, BCUT descriptors) and C-002 (from the atom-centered fragments). The 4^{th} order mean Galvez topological charge index (JGI4), the lowest eigenvalue n.1 of Burden matrix/ weighted by atomic masses (BELm1) and CH2R2 type atom centered fragment (C-002) correlated positively to the activity suggested that a higher value of these will augment the activity.

However, for all the models mentioned in Table 3.6, the r^2_{Test} values (<0.5) are inferior to a specified value. Considering the number of observation in the data set, models with up to four descriptors were explored. A total number of seven models have been obtained through CP-MLR. Following are the selected four-descriptor models, obtained from CP-MLR, for the DPP-4 inhibitory activity.

$pIC_{50} = 4.722 + 1.993(0.276)JGI4 – 1.295(0.300)ATS8p$
$+ 2.163(0.265)GATS8p + 0.405(0.173)C\text{-}024$
$n = 20, r = 0.952, s = 0.214, F = 36.718, q^2_{LOO} = 0.832,$
$q^2_{L5O} = 0.805, r^2_{Test} = 0.576$ (**3.17**)

$pIC_{50} = 5.165 – 0.688(0.336)ATS8p – 0.690(0.267)GATS7e$
$+ 2.112(0.344)GATS8p + 1.331(0.288)MLOGP$

$n = 20, r = 0.921, s = 0.273, F = 21.086, q^2_{LOO} = 0.715,$
$q^2_{L5O} = 0.704, r^2_{Test} = 0.730$ (**3.18**)

$pIC_{50} = 5.910 – 1.302(0.282)RBN + 3.530(0.486)BIC3$
$– 2.310(0.364)BIC5 + 1.897(0.247)H\text{-}052$
$n = 20, r = 0.920, s = 0.274, F = 20.846, q^2_{LOO} = 0.561,$
$q^2_{L5O} = 0.713, r^2_{Test} = 0.801$ (**3.19**)

$pIC_{50} = 3.727 + 3.070(0.467)BIC3 + 1.341(0.226)SRW09$
$+ 0.853(0.243)C\text{-}040 + 2.530(0.335)H\text{-}052$
$n = 20, r = 0.907, s = 0.295, F = 17.516, q^2_{LOO} = 0.558,$
$q^2_{L5O} = 0.582, r^2_{Test} = 0.543$ (**3.20**)

The 18 descriptors, that were shared by these seven models, along with their brief description, average regression coefficients and total incidences are given in Table 3.7.

Table 3.7: Identified descriptors[a] along with their physical meaning, average regression coefficient and incidence[b], in modeling the DPP-4 inhibitory activity.

S. No.	Descriptor	Descriptor class	Physical meaning	Average regression coefficient (incidence)
1	RBN	Constitutional	Number of rotatable bonds	-1.302(1)
2	PJI2	Topological	2D Petitijean shape index	0.360(1)
3	BIC3	Topological	Bond information content (neighborhood symmetry of 3 order)	3.300(2)
4	BIC5	Topological	Bond information content (neighborhood symmetry of 5 order)	-2.310(1)
5	SRW09	Molecular walk counts	Self- returning walk count of order 09	1.341(1)
6	BELm1	BCUT	Lowest eigenvalue n.1 of Burden matrix/ weighted by atomic masses	0.713(1)

7	BEHv1	BCUT	Highest eigenvalue n.1 of Burden matrix/ weighted by atomic van der Waals volumes	0.821(1)
8	BELe1	BCUT	lowest eigenvalue n.1 of Burden matrix/ weighted by atomic Sanderson electronegativities	0.655(1)
9	BELp2	BCUT	lowest eigenvalue n.2 of Burden matrix/ weighted by atomic polarizabilities	-0.675(1)
10	JGI4	Galvez topological charge indices	Mean topological charge index of order 4	1.993(1)
11	ATS8p	2D autocorrelations	Broto-Moreau autocorrelation of a topological structure - lag8/ weighted by atomic polarizabilities	-0.992(2)
12	GATS7e	2D autocorrelations	Geary autocorrelation of lag-7/ weighted by atomic Sanderson electronegativities	-0.901(4)
13	GATS8p	2D autocorrelations	Moran autocorrelation of lag-8/ weighted by atomic polarizabilities	2.020(5)
14	C-024	Atom-centred fragments	R--CH--R	0.405(1)
15	C-040	Atom-centred fragments	R-C(=X)-X/R-C#X/X-=C=X	0.853(1)
16	H-052	Atom-centred fragments	H attached to C0(sp3) with 1X attached to next C	2.214(2)
17	MR	Properties	Ghose-Crippen molecular refractivity	-1.058(1)
18	MLOGP	Properties	Moriguchi octanol-water partition coefficient (logP)	1.331(1)

[a]The descriptors are identified from the four parameter models, emerged from CP-MLR protocol with filter-1 as 0.79, filter-2 as 2.0, filter-3 as 0.809 (r-bar of the three parameter model having the highest r^2_{Test}), and filter-4 as $0.3 \leq q^2 \leq 1.0$ with a training set of 20 compounds. [b]The average regression coefficient of the descriptor corresponding to all models and the total number of its incidence.

The newly emerged descriptors C-024, H-052 and C-040 in these models are atom centered fragments and shown positive correlation to the activity. Therefore, presence of R--CH--R (descriptor C-024), H attached to C0(sp3) with 1X attached to next C (descriptor H-052) and R-C(=X)-X/R-C#X/X-=C=X

(descriptor C-040) type atom centered fragments in a molecular structure would enhance the activity. Topological class descriptor BIC3 (bond information content of 3rd order neighborhood symmetry) contributed positively whereas BIC5 (bond information content of 5th order neighborhood symmetry) contributed negatively to the activity revealed that a higher value of descriptor BIC3 and a lower value of descriptor BIC5 would be beneficiary to the activity. Descriptor SRW09 represents self- returning walk count of order 09 is from molecular walk counts class. Molecular walk counts are 2D-descriptors representing self-returning walk counts of different lengths. The descriptor MLOGP is from properties class representing Moriguchi octanol-water partition coefficient (logP). It is evinced from the models that higher values of both of these descriptors (SRW09 and MLOGP) would augment the activity.

In all above equations, the F-values remained significant at 99% level. The indices q^2_{LOO} and q^2_{L5O} (> 0.5) have accounted for their internal robustness. For all above models the r^2_{Test} values, obtained greater than 0.5, specified that the selected test-set is fully accountable for their external validation. These models are able to estimate up to 90.73 percent of variance in observed activity of the compounds. The derived statistical parameters of these four models have shown that these models are significant. These models were, therefore, used to calculate the activity profiles of all the compounds and are included in Table 3.8 for the sake of comparison with observed ones. A close agreement between them has been observed.

Table 3.8: Observed, calculated and predicted DPP-4 inhibition activities of (2*S*)-Cyanopyrrolidine analogues.

Cpd.	$pIC_{50}(M)$[a]								
	Obsd[b].	Eq. (3.17)		Eq. (3.18)		Eq. (3.19)		Eq. (3.20)	
		Calc.	Pred[c].	Calc.	Pred[c].	Calc.	Pred[c].	Calc.	Pred[c].
1	5.49	5.76	5.90	5.37	5.31	5.75	5.98	5.73	5.79

2	6.94	6.78	6.76	6.75	6.72	6.59	6.52	6.64	6.61
3	6.19	6.42	6.53	6.04	6.00	5.84	5.76	5.86	5.80
4[d]	7.08	7.00	-[d]	6.99	-[d]	7.26	-[d]	7.22	-[d]
5	6.88	6.68	6.65	6.56	6.50	7.03	7.11	6.84	6.82
6	6.37	6.58	6.60	6.29	6.18	6.40	6.41	6.78	6.82
7	7.27	7.09	7.07	7.06	7.04	6.93	6.85	7.27	7.27
8[d]	6.09	6.42	-[d]	6.45	-[d]	6.26	-[d]	6.32	-[d]
9	7.31	7.33	7.34	7.21	7.19	7.13	7.09	7.10	7.06
10	7.52	7.46	7.44	7.42	7.40	7.70	7.77	7.60	7.63
11	7.66	7.44	7.34	7.52	7.47	7.70	7.72	7.60	7.58
12	7.82	7.98	8.07	7.72	7.68	7.45	7.18	7.33	6.99
13	6.17	6.20	6.20	6.23	6.24	6.15	6.14	5.88	5.71
14	6.38	6.72	6.80	6.36	6.34	6.15	6.06	6.57	6.68
15	6.20	6.11	5.99	6.25	6.28	6.21	6.21	6.30	6.34
16	6.28	5.88	5.78	6.31	6.32	6.27	6.27	6.81	7.10
17	6.34	6.33	6.32	6.44	6.45	6.56	6.59	6.27	6.26
18	6.50	6.47	6.47	6.42	6.41	6.67	6.71	6.44	6.43
19[d]	6.35	6.00	-[d]	6.24	-[d]	6.14	-[d]	6.42	-[d]
20	6.43	6.56	6.59	6.93	7.16	6.88	6.99	6.55	6.57
21	6.11	5.97	5.88	5.96	5.88	6.06	6.05	5.87	5.83
22[d]	6.92	6.71	-[d]	6.76	-[d]	6.97	-[d]	6.62	-[d]
23	5.96	6.07	6.20	6.51	6.80	6.32	7.14	6.26	6.87
24	6.25	6.24	6.24	6.75	6.84	6.24	6.23	6.36	6.37
25[d]	6.53	6.51	-[d]	6.60	-[d]	6.68	-[d]	6.16	-[d]

[a]IC_{50} represents the concentration of a compound required to bring out 50% inhibition of DPP-IV and the same is expressed as pIC_{50} on molar basis; [b]Taken from ref. [729]; [c]Leave-one-out (LOO) procedure; [d]Compound included in test set.

Additionally, the graphical display, showing the variation of observed versus calculated activities is given in Figure 3.6 to ensure the goodness of fit for each of these four models.

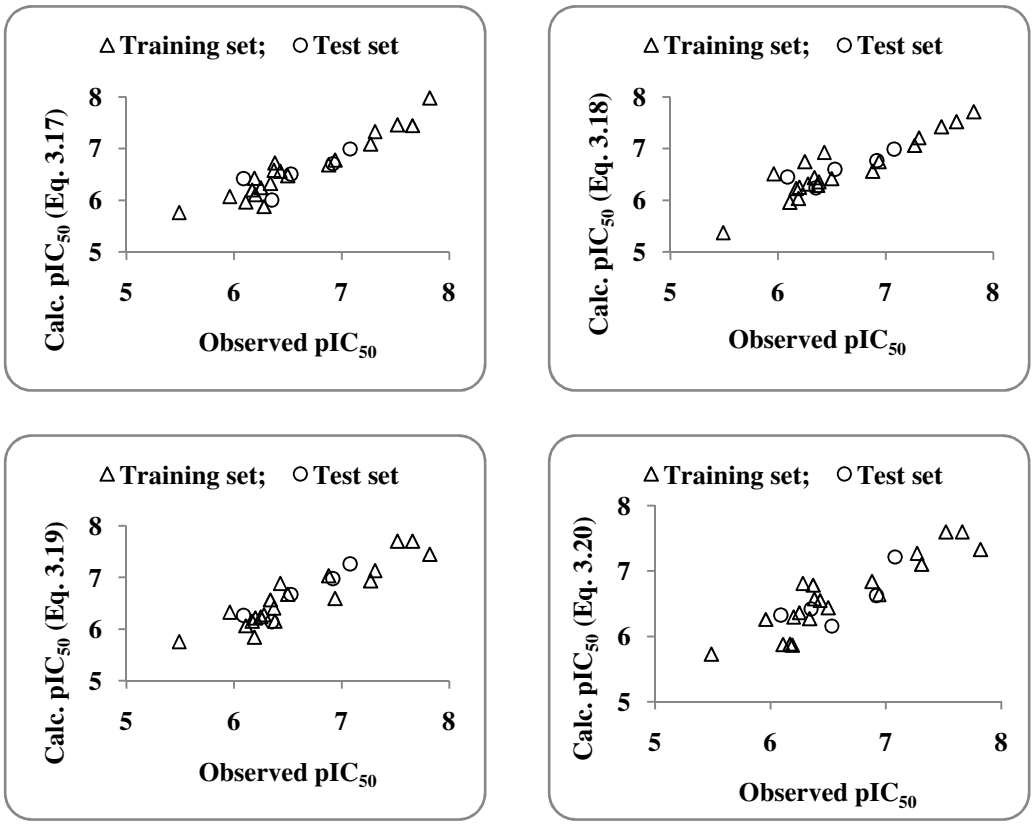


Figure 3.6: Plots showing variations of observed versus calculated pIC_{50} values of training and test set compounds.

CP-MLR analysis has also been performed for the DPP-8 inhibitory activity with the descriptor pool of 84 descriptors with the same test which was used for the DPP-4 inhibitory activity.

All the emerged four models in three descriptors are given below through Equations (3.21) to (3.24).

$pIC_{50} = 5.283 + 0.863(0.234)MW - 0.745(0.207)X1Av + 0.544(0.123)PJI2$

$n = 15, r = 0.885, s = 0.218, F = 13.264, q^2_{LOO} = 0.552,$

$q^2_{L5O} = 0.578, r^2_{Test} = 0.683$ (**3.21**)

$pIC_{50} = 5.419 - 0.536(0.210)X1Av + 0.533(0.129)PJI2 + 0.497(0.148)C\text{-}040$

$n = 15, r = 0.871, s = 0.229, F = 11.623, q^2_{LOO} = 0.507,$

$q^2_{L5O} = 0.612, r^2_{Test} = 0.761$ (**3.22**)

$pIC_{50} = 5.300 - 0.778(0.223)X1Av + 0.575(0.130)PJI2 + 0.771(0.231)BEHm8$

$n = 15, r = 0.871, s = 0.230, F = 11.525, q^2_{LOO} = 0.506,$

$q^2_{L5O} = 0.726, r^2_{Test} = 0.646$ (**3.23**)

$pIC_{50} = 6.037 + 0.609(0.135)PJI2 - 0.601(0.254)GATS1e - 0.691(0.205)C\text{-}024$

$n = 15, r = 0.860, s = 0.238, F = 10.493, q^2_{LOO} = 0.533,$

$q^2_{L5O} = 0.528, r^2_{Test} = 0.511$ (**3.24**)

The derived statistical parameters of these four models have shown that these models are significant and are able to explain up to 78.34 percent of variance in observed DPP-8 activity of the compounds. The activity values, calculated using Eqs. (3.21) to (3.24), are in close agreement to the observed ones and this agreement is given in Table 3.9 for the sake of comparison.

Table 3.9: Observed, calculated and predicted DPP-8 inhibition activities of (2*S*)-Cyanopyrrolidine analogues.

Cpd.	pIC_{50}[a]								
	Obsd[b].	Eq. (3.21)		Eq. (3.22)		Eq. (3.23)		Eq. (3.24)	
		Calc.	Pred[c].	Calc.	Pred[c].	Calc.	Pred[c].	Calc.	Pred[c].
1	5.38	5.72	5.82	5.79	5.89	5.64	5.74	5.81	5.89
2	5.45	5.11	5.04	5.16	5.11	5.20	5.16	5.20	5.16
3	5.48	5.60	5.63	5.34	5.32	5.69	5.75	5.50	5.53
4[d]	5.77	5.88	-[d]	5.80	-[d]	6.01	-[d]	6.20	-[d]
5	5.67	5.72	5.73	5.79	5.82	5.90	5.94	5.81	5.84
6	5.85	5.61	5.55	5.62	5.56	5.79	5.77	5.81	5.81
7	5.27	5.12	5.09	5.13	5.10	5.23	5.23	5.20	5.19
8[d]	4.38	5.11	-[d]	5.02	-[d]	5.17	-[d]	5.20	-[d]
9	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
10	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
11	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
12	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
13	6.69	6.40	6.11	6.43	6.11	6.29	6.02	6.21	5.97
14	5.47	5.55	5.70	5.53	5.62	5.45	5.43	5.55	5.56
15	5.7	5.66	5.66	5.91	6.04	5.64	5.63	5.67	5.65

16	5.67	5.87	5.95	5.57	5.54	5.81	5.85	5.69	5.69
17	4.97	5.01	5.02	5.18	5.21	4.99	5.00	4.88	4.84
18	5.62	5.61	5.61	5.48	5.43	5.41	5.33	5.65	5.67
19[d]	4.66	4.92	-[d]	5.02	-[d]	4.84	-[d]	4.89	-[d]
20	5.26	5.49	5.57	5.41	5.46	5.60	5.72	5.58	5.68
21	4.89	4.97	5.00	5.01	5.05	4.88	4.87	4.89	4.90
22[d]	5.08	5.05	-[d]	5.11	-[d]	4.90	-[d]	4.89	-[d]
23	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
24	5.59	5.51	5.48	5.63	5.65	5.46	5.39	5.50	5.46
25[d]	6.07	5.60	-[d]	5.79	-[d]	5.65	-[d]	5.49	-[d]

[a]IC_{50} represents the concentration of a compound required to bring out 50% inhibition of DPP-4 and the same is expressed as pIC_{50} on molar basis; [b]Taken from ref. [729]; [c]Leave-one-out (LOO) procedure; [d]Compound included in test set.[e]Compound with uncertain activity, not part of data set.

The participated descriptors in above models suggested that higher values of molecular weight (MW, constitutional class descriptor), 2D-Petitijean shape index (PJI2, topological class), highest eigenvalue n.8 of Burden matrix weighted by atomic masses (BEHm8, BCUT descriptor) and presence of R-C(=X)-X/R-C#X/X-=C=X type atom centered fragments (descriptor C-040, atom centered fragment descriptor) would be beneficiary to DPP8 inhibitory activity. Another emerged topological class descriptor X1Av (average valence connectivity index, chi-1), 2D-AUTO class descriptor GATS1e (Geary autocorrelation of lag-1/weighted by atomic Sanderson electronegativities) advocated that a lower value of these descriptors and absence of R--CH--R type fragment (descriptor C-024) would augment the activity.

CP-MLR has also been carried out on DPP-8 inhibitory activity from the pool of 90 descriptors which was used to find rationales for the DPP-4 inhibitory activity. The analysis resulted into 10 models having both the q^2_{LOO} and r^2_{Test} > 0.5 and the highest significant four models are listed in Table 3.10. Models listed in Table 3.10 are able to estimate up 84.82 percent of variance in observed DPP-8 activity of the compounds.

Table 3.10: Three parameter CP-MLR models for the DPP-8 inhibition activity from the descriptor pool of DPP-4.

Model	r	s	F	r^2_{Test}	Eq.
pIC_{50} = 5.218 – 0.999(0.193)X2Av + 0.523(0.103)PJI2 +1.075(0.230)MR	0.920	0.182	20.490	0.545	(**3.25**)
pIC_{50} = 5.510 – 0.737(0.203)X2Av + 0.464(0.118)PJI2 + 0.613(0.161)N-072	0.896	0.208	14.982	0.755	(**3.26**)
pIC_{50} = 5.405 – 0.865(0.217)X2Av + 0.411(0.125)PJI2 + 0.702(0.195)BELp2	0.889	0.214	13.934	0.516	(**3.27**)
pIC_{50} = 5.221 – 1.066(0.276)X0Av + 0.410(0.125)PJI2 + 0.871(0.231)H-047	0.888	0.215	13.730	0.762	(**3.28**)

[a]The models, in three parameters, emerged from CP-MLR protocol with filter-1 as 0.79, filter-2 as 2.0, filter-3 as 0.5 and filter-4 as $0.3 \leq q^2 \leq 1.0$ with a training set of 15 compounds.

The newly appeared descriptors in above models are MR (property class descriptor), N-072 and H-047 (ACF class descriptor), X0Av and X2Av (TOPO class descriptor) and BELp2 (BCUT descriptor). Tabled Equations revealed that lower values of average valence connectivity indices (X0Av and X2Av, chi-0 and chi-2) would be advantageous to enhance the activity. On the other hand, a higher lower value of Ghose-Crippen molecular refractivity (MR) and lowest eigenvalue n.2 of Burden matrix weighted by atomic polarizabilities are incremental to the activity. Counts for certain structural fragments, H attached to C1(sp3) /C0(sp2) (descriptor H-047) and R-CO-N</>N-X=X (descriptor N-072) strongly recommend the presence of such structural features favorable to activity. Thus the descriptors identified for rationalizing the DPP-4 activity give avenues to rationalize the DPP-8 inhibitory activity.

From the different nature of emerged descriptors in final statistically significant models for DPP-4 and DPP-8 inhibition actions, it appeared that the mode of actions of titled compounds are different for DPP-4 and DPP-8 enzyme systems.

2.2.1.2. Applicability domain (AD)

To analyze the applicability domain (AD) a Williams plot of the model based on the whole data set (Table 3.11) has been constructed that is shown in Figure 3.7.

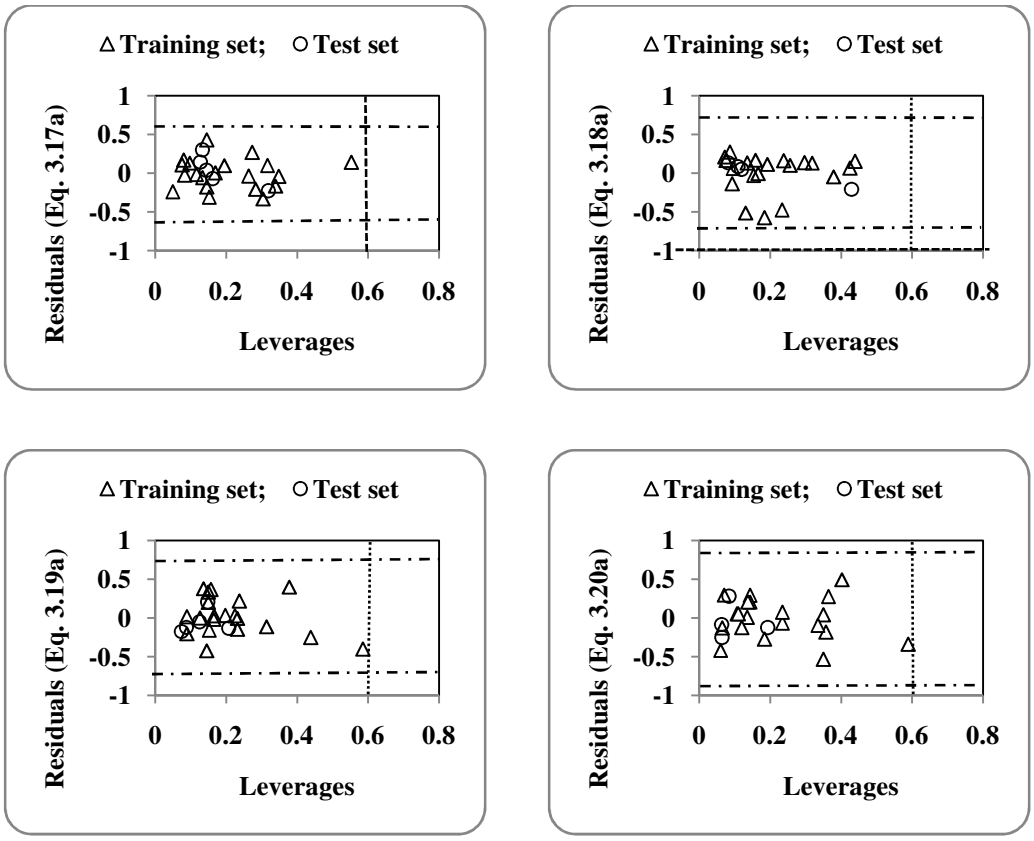


Figure 3.7: Williams plot for the training-set and test- set for inhibition activity of DPP-4 for the compounds in Table 3.5. The horizontal dotted line refers to the residual limit (±3×standard deviation) and the vertical dotted line represents threshold leverage h* (= 0.6).

Table 3.11: Models derived for the whole data set (n = 25) for the DPP-4 inhibitory activity in descriptors identified through CP-MLR.

Model	r	s	F	Eq.
pIC_{50} = 4.707 +2.088(0.260)JGI4 –1.513(0.230)ATS8p +2.197(0.246)GATS8p +0.478(0.158)C-024	0.942	0.212	39.987	(3.17a)
pIC_{50} = 5.216 –0.847(0.249)ATS8p –0.805(0.207)GATS7e +2.143(0.298)GATS8p +1.365(0.254)MLOGP	0.920	0.249	27.603	(3.18a)
pIC_{50} = 5.902 –1.229(0.233)RBN +3.431(0.415)BIC3 –2.257(0.315)BIC5 +1.866(0.217)H-052	0.919	0.250	27.370	(3.19a)
pIC_{50} = 3.827 +2.966(0.412)BIC3 +1.301(0.208)SRW09 +0.809(0.223)C-040 +2.471(0.305)H-052	0.895	0.283	20.219	(3.20a)

The analysis revealed that none of the compound has been identified as an obvious 'outlier' for the DPP-4 inhibitory activity if the limit of normal values for the Y outliers (response outliers) was set as 3×(standard deviation) units. None of the compounds was found to have leverage (h) values greater than the threshold leverage (h*). For both the training-set and test-set, the suggested model matches the high quality parameters with good fitting power and the capability of assessing external data. Furthermore, all of the compounds were within the applicability domain of the proposed model and were evaluated correctly.

2.2.2. CONCLUSIONS

This study has provided a rational approach for the development of (2*S*)-Cyanopyrrolidine analogues as DPP-4 inhibitors. The descriptors identified in CP-MLR analysis have highlighted the role of atomic properties in respective lags of 2D-autocorrelations (ATS8p, GATS8p and GATS7e), 4^{th} order mean Galvez topological charge index (JGI4), 3^{rd} and 5^{th} order bond information content of neighborhood symmetry (BIC3 and BIC5) and 9^{th} order self returning walk-count (SRW09) to explain the biological actions of (2*S*)-Cyanopyrrolidine analogues as

DPP-4 inhibitors. Certain structural features or fragments (RBN, C-024, C-040 and H-052) in molecular structures in addition to hydrophobicity (MLOGP) of a molecule have also shown prevalence to optimize the DPP-4 inhibitory activity of titled compounds. Applicability domain analysis revealed that the suggested model for DPP-4 inhibitory activity matches the high quality parameters with good fitting power and the capability of assessing external data and all of the compounds was within the applicability domain of the proposed model and were evaluated correctly.

2.3. THE β-AMINOAMIDE BEARING TRIAZOLOPIPERAZINES

The GLP-1 therapy is beneficiary due to the regulation of insulin in a strictly glucose-dependent manner. Little or no risk of hypoglycemia, slowing down of gastric emptying and reduction of appetite are the beneficial effects of GLP-1 therapy. As a result of inhibition of DPP-4 the half-life of GLP-1 is increased and thus the beneficial effects of this incretin hormone are prolonged. Sitagliptin [730, 731], LAF-237 [732] and BMS-477118 [733] are examples of DPP-4 inhibitors.

Detailed structure–activity relationships (SARs) of Sitagliptin scaffold as DPP-4 inhibitors are reported in literature with a variety of substituents on the left phenyl and the right triazolopiperazine [734]. Alkyl substitution around the β-aminoamide backbone was found to be detrimental to potency. Other modifications such as lengthening, shortening, or tethering along with alkyl substitution of triazolopiperazine series were discarded due to the similar ineffective SAR trends of corresponding thiazolidine [735] and the piperazine series [736]. A series of β-aminoamides bearing triazolopiperazines having alkyl substitutions around the triazolopiperazine moiety has been reported by Kim *et al*. [737].

The general structure of the reported thirtynine β-aminoamides bearing triazolopiperazine derivatives, which are considered as the data set is given in Figure 3.8.

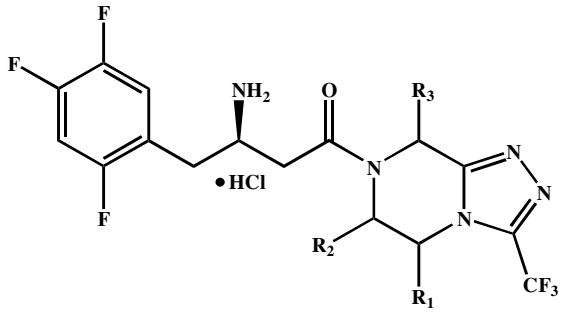


Figure 3.8: The generalized structure of triazolopiperazine derivatives.

The structural variations of these analogues are mentioned in Table 3.12. These compounds were evaluated in vitro for their inhibition of DPP-4 and DPP-8. This table also contains the reported inhibitory activity, in terms of IC_{50}(nM), of these congeners. Additionally, test set compounds, selected through an in-house written randomization program, have also been mentioned in Table 3.12.

Table 3.12: Structural variations and reported DPP-4 and DPP-8 inhibitory activities of triazolopiperazines.

Cpd.	R_1	R_2	R_3	IC_{50}(nM)[a]	
				DPP-4	DPP-8
1[b]	H	H	H	18	48000
2	(*S*)-CH_3	H	H	23	23000
3	(*R*)-CH_3	H	H	14	33000
4[c]	H	(*S*)-CH_3	H	91	$>10^5$
5	H	(*R*)-CH_3	H	42	75000
6[b,c]	H	H	(*S*)-CH_3	88	$>10^5$
7	H	H	(*R*)-CH_3	4.3	17000
8	*di*-CH_3	H	H	92	66000
9	H	H	*di*-CH_3	175	6000
10[b,c]	CH_3	H	CH_3	100	$>10^5$
11[b,c]	CH_3	H	CH_3	209	$>10^5$
12	CH_3	H	CH_3	12	70000
13	CH_3	H	CH_3	11	44000
14[c]	H	H	Et	113	$>10^5$
15	H	H	Et	5	8000
16[c]	H	H	CH_2CF_3	123	$>10^5$
17	H	H	CH_2CF_3	5.7	1600
18	H	H	$CH_2CH{=}CH_2$	1.5	3000
19	H	H	$CH_2CH{=}CH_2$	32	72000
20[c]	H	H	$CH_2CON(CH_3)_2$	377	$>10^5$

21[b]	H	H	$CH_2CON(CH_3)_2$	2.8	30000
22[c]	H	H	CH_2Ph	140	$>10^5$
23	H	H	CH_2Ph	0.66	622
24[c]	H	H	CH_2(4-methoxyphenyl)	320	$>10^5$
25	H	H	CH_2(4-methoxyphenyl)	0.43	367
26[b,c]	H	H	CH_2(2-trifluoromethylphenyl)	438	$>10^5$
27[b]	H	H	CH_2 (2-trifluoromethylphenyl)	0.31	8000
28[b,c]	H	H	CH_2(2-fluorophenyl)	131	$>10^5$
29[b]	H	H	CH_2(2-fluorophenyl)	0.46	1103
30[c]	H	H	CH_2(4-fluorophenyl)	116	$>10^5$
31	H	H	CH_2(4-fluorophenyl)	0.18	332
32[c]	H	H	CH(OH)(4-fluorophenyl)	430	$>10^5$
33	H	H	CH(OH)(4-fluorophenyl)	0.32	326
34	H	H	CH(OH)(4-fluorophenyl)	90	40000
35	H	H	CH(OH)(4-fluorophenyl)	0.5	628
36[c]	H	H	CH_2(3,5-*bis*-trifluoromethylphenyl)	587	$>10^5$
37[b,c]	H	H	CH_2 (3,5-*bis*-trifluoromethylphenyl)	6.3	$>10^5$
38[c]	H	H	CH_2(2-pyridyl)	132	$>10^5$
39	H	H	CH_2(2-pyridyl)	0.4	5000

[a]Concentration of a compound to bring out 50% inhibition (IC_{50}), taken from reference [737]; [b]Compound included in test set; [c]Compound with uncertain activity, not part of data set for DDP-8.

2.3.1. RESULTS AND DISCUSSION

2.3.1.1. QSAR RESULTS

A total number of 158 significant 3D-molecular descriptors have been subjected to CP-MLR analysis. These 158 descriptors were finally obtained after the exclusion of those descriptors which were intercorrelated beyond 0.90 and showing a correlation of <0.1 with the biological endpoints among a total number of 673 descriptors. Statistical models in two, three and four descriptor(s) have been derived successively to achieve the best relationship correlating DPP-4 inhibitory activity. A total number of 10, 20 and 22 models in two, three and four descriptors, respectively, were obtained through CP-MLR. These models (with 158 descriptors) were identified in CP-MLR by successively incrementing the filter-3 with increasing number of descriptors (per equation). For this, the optimum *r*-bar value of the preceding level model has been used as the new

threshold of filter-3 for the next generation. The selected models in two, three and four descriptors are given below.

pIC_{50} = 7.563 – 3.848(0.768)RDF075m + 6.324(1.145)RDF085m
n = 29, r = 0.765, s = 0.722, F = 18.405, Q^2_{LOO} = 0.499, Q^2_{L5O} = 0.503
r^2_{Test} = 0.372, FIT = 1.115, LOF = 0.629, AIC = 0.642 (**3.29**)

pIC_{50} = 7.642 – 3.011(0.737)RDF075m + 2.883(0.568)RDF105p
n = 29, r = 0.740, s = 0.754, F = 15.752, Q^2_{LOO} = 0.469, Q^2_{L5O} = 0.465
r^2_{Test} = 0.165, FIT = 0.954, LOF = 0.687, AIC = 0.701 (**3.30**)

pIC_{50} = 8.702 – 5.173(0.780)RDF075m + 6.241(0.985)RDF085m
– 1.785(0.560)G3e
n = 29, r = 0.840, s = 0.621, F = 19.979, Q^2_{LOO} = 0.615, Q^2_{L5O} = 0.603
r^2_{Test} = 0.646, FIT = 1.577, LOF = 0.528, AIC = 0.509 (**3.31**)

pIC_{50} = 10.050 – 2.942(0.734)DISPv + 3.853(0.976)RDF085m
– 3.848(0.602)RDF110e
n = 29, r = 0.833, s = 0.631, F = 19.029, Q^2_{LOO} = 0.588, Q^2_{L5O} = 0.538
r^2_{Test} = 0.505, FIT = 1.502, LOF = 0.547, AIC = 0.526 (**3.32**)

pIC_{50} = 5.359 –2.502(0.665)RDF075m + 4.134(0.557)RDF085p
+ 1.904(0.474)Mor10m + 2.268(0.534)RTu+
n = 29, r = 0.889, s = 0.534, F = 22.626, Q^2_{LOO} = 0.696, Q^2_{L5O} = 0.615
r^2_{Test} = 0.684, FIT = 2.011, LOF = 0.451, AIC = 0.405 (**3.33**)

pIC_{50} = 6.016 –1.956(0.525)RDF110e + 2.590(0.414)RDF105p
+ 2.139(0.497)Mor10m + 1.442(0.516)RTp+
n = 29, r = 0.877, s = 0.560, F = 20.083, Q^2_{LOO} = 0.621, Q^2_{L5O} = 0.633
r^2_{Test} = 0.525, FIT = 1.785, LOF = 0.495, AIC = 0.444 (**3.34**)

In the randomization study (100 simulations per model), none of the identified models has shown any chance correlation. Most of the participated

descriptors RDF075m, RDF085m, RDF110e, RDF085p and RDF105p in above models belong to RDF descriptors class.

RDF (Radial Distribution Function) descriptors [738] are based on a radial distribution function that may be considered as the probability distribution of finding an atom in a spherical volume of radius (R). The RDF descriptors are represented as RDFkw where k is step size and w is weighting scheme such as the unweighted case (u), atomic mass (m), the van der Waals volume (v), the Sanderson atomic electronegativity (e) and the atomic polarizability (p). The RDF descriptors not only provide information about interatomic distances in the whole molecule but bond distances, atom types, ring types and planar and non-planar systems also. Atomic masses weighted radial distribution function 7.5 (descriptor RDF075m) and atomic Sanderson electronegativities weighted radial distribution functions 11.0 (descriptor RDF110e) contributed negatively to the activity suggesting that a higher values of these radial distribution functions would be detrimental to DPP-4 inhibition actions. On the other hand positive contribution of radial distribution function-8.5/ weighted by atomic masses (descriptor RDF085m) and atomic polarizabilities weighted radial distribution functions 8.5 and 10.5 (RDF085p and RDF105p, respectively) advocated a higher value of these to augment the inhibitory activity.

Descriptor DISPv is representative of geometrical class of descriptors. Geometrical descriptors are derived from the three-dimensional structure of the molecule and calculation of these is based on some optimized molecular geometry obtainable by the methods of the computational chemistry or on crystallographic coordinates. Geometrical descriptors offer more information and discrimination power for similar molecular structures and molecule conformations because a geometrical representation of a molecule involves the knowledge of the relative positions of the atoms in 3D space. Descriptor DISPv is among the COMMA2 descriptors [739]. COMMA2 descriptors are given by moment expansions for which the zero-order moment of a considered property (such as mass (m), van der

Waals volume (v), Sanderson electronegativity (e) and polarizability (p)) field is non-vanishing. The negative contribution of descriptor DISPv (the displacement between the geometric centre and the centre of the van der Waals volume field, calculated with respect to the molecular principal axes) hints that a lower value of it would be beneficiary to the activity.

Descriptor Mor10m is a 3D-MoRSE (3D-Molecule Representation of Structures based on Electron diffraction) descriptor [740]. These descriptors (Mor*sw*) represent the scattered electron intensity (signals). The term *s* represents the scattering in various directions by a collection of atoms and *w* is the atomic property or may be unweighted case. The positive contribution of 3D-MoRSE - signal 10/ weighted by atomic masses (Mor10m) suggests that a higher value of it would be incremental to the activity.

Descriptors RTu+ and RTp+ are from the GETAWAY (GEometry, Topology, and Atom-Weights AssemblY) class of descriptors. GETAWAYs [741] are geometrical descriptors which encode information on the effective position of substituents and fragments in the molecular space. These descriptors are independent of molecule alignment and account for information on molecular size and shape and for specific atomic properties. Both the descriptors RTu+ (unweighted R maximal index) and RTp+ (atomic polarizabilities weighted R maximal index) shown positive correlation to the activity advocating higher values of these for augmented activity.

The remaining descriptor G3e is a Weighted Holistic Invariant Molecular (WHIM) descriptor. These descriptors are geometrical descriptors and are based on statistical indices calculated on the projections of the atoms along principal axes [742]. WHIM descriptors are free from prior alignment of molecules because these are invariant to translation and rotation. WHIM descriptors (categorized as directional and global) furnish relevant molecular 3D information about molecular size, shape, symmetry, and atom distribution with respect to invariant reference frames. The appeared WHIM descriptor, G3e (3rd component symmetry

directional WHIM index/weighted by atomic Sanderson electronegativities) correlated negatively to the activity suggesting lower value of it for elevated DPP-4 activity.

The four descriptor models could estimate nearly 79% in observed activity of the compounds. Considering the number of observations in the dataset, models with up to five descriptors were explored. It has resulted in 4 five-parameter models with test set $r^2 > 0.50$. These models have shared 10 descriptors among them. All these 10 descriptors along with their brief meaning, average regression coefficients, and total incidence are listed in Table 3.13, which will serve as a measure of their estimate across these models.

Table 3.13: Identified descriptors[a] along with their physical meaning, average regression coefficient and incidence[b], in modeling the DPP-4 inhibitory activity of triazolopiperazines.

S. No.	Descriptor class	Descriptor	Physical meaning, average regression coefficient (**incidence**)
1	Geometrical descriptors	DISPv	d COMMA2 value /weighted by atomic van der Waals volumes, -1.712(**1**)
2	RDF descriptors	RDF075m	Radial distribution function at 7.5 Å / weighted by atomic masses, -4.112(**1**)
3		RDF085m	Radial Distribution Function at 8.5 Å / weighted by atomic polarizabilities, 5.814(**1**)
4		RDF110e	Radial Distribution Function at 11.0 Å / weighted by atomic Sanderson electronegativities, -2.674(**3**)
5		RDF155e	Radial Distribution Function at 15.5 Å / weighted by atomic Sanderson electronegativities, -2.016(**1**)
6		RDF085p	Radial Distribution Function at 8.5 Å / weighted by atomic polarizabilities 2.874(**3**)
7	3D-MoRSE descriptors	Mor10m	3D-MoRSE - signal 10 / weighted by atomic masses, 1.977(**4**)
8		Mor12m	3D-MoRSE - signal 12 / weighted by atomic masses, 1.622(**2**)
9	WHIM descriptors	G3e	3rd component symmetry directional WHIM index /weighted by atomic

10	GETAWAY descriptor	RTu+	Sanderson electronegativities, -1.879(**1**) R maximal index / unweighted, 1.613(**3**)

[a]The descriptors are identified from the five parameter models, emerged from CP-MLR protocol with filter-1 as 0.79, filter-2 as 2.0, filter-3 as 0.869, and filter-4 as $0.3 \leq q^2 \leq 1.0$ with a training set of 29 compounds. [b]The average regression coefficient of the descriptor corresponding to all models and the total number of its incidence. The arithmetic sign of the coefficient represents the actual sign of the regression coefficient in the models.

Following are the emerged five-descriptor models for the DPP-4 inhibitory activity of titled compounds.

$pIC_{50} = 7.266 – 1.712(0.685)DISPv – 2.653(0.634)RDF110e$

$+ 3.126(0.567)RDF085p + 1.564(0.507)Mor10m + 1.795(0.570)RTu+$

$n = 29, r = 0.900, s = 0.518, F = 19.774, Q^2_{LOO} = 0.668, Q^2_{L5O} = 0.680$

$r^2_{Test} = 0.606, FIT = 1.830, LOF = 0.496, AIC = 0.409$ (**3.35**)

$pIC_{50} = 5.240 – 2.398(0.595)RDF110e + 3.094(0.581)RDF085p$

$+ 2.392(0.481)Mor10m + 1.436(0.603)Mor12m + 1.740(0.586)RTu+$

$n = 29, r = 0.898, s = 0.523, F = 19.296, Q^2_{LOO} = 0.715, Q^2_{L5O} = 0.726$

$r^2_{Test} = 0.512, FIT = 1.786, LOF = 0.506, AIC = 0.417$ (**3.36**)

$pIC_{50} = 6.452 – 2.969(0.510)RDF110e – 2.016(0.719)RDF155e$

$+ 2.402(0.505)RDF085p + 2.802(0.501)Mor10m + 1.809(0.580)Mor12m$

$n = 29, r = 0.895, s = 0.531, F = 18.573, Q^2_{LOO} = 0.711, Q^2_{L5O} = 0.690$

$r^2_{Test} = 0.735, FIT = 1.719, LOF = 0.522, AIC = 0.430$ (**3.37**)

$pIC_{50} = 7.507 – 4.112(0.760)RDF075m + 5.814(0.901)RDF085m$

$+ 1.151(0.497)Mor10m – 1.879(0.485)G3e + 1.303(0.500)RTu+$

$n = 29, r = 0.894, s = 0.533, F = 18.437, Q^2_{LOO} = 0.611, Q^2_{L5O} = 0.638$

$r^2_{Test} = 0.670, FIT = 1.707, LOF = 0.525, AIC = 0.432$ (**3.38**)

The newly appeared descriptors in above models are RDF155e (a RDF class descriptor) and Mor12m (from 3D-MoRSE class). The signs of regression coefficients of these descriptors suggest that a lower value of radial distribution function – 15.5/weighted by atomic Sanderson electronegativities (descriptor

RDF155e) and a higher value of 3D-MoRSE - signal 12/ weighted by atomic masses (descriptor Mor12m) would be incremental to the activity. In this way the descriptors identified for rationalizing the activity give avenues to modulate the structure to a desirable biological endpoint.

These models have accounted for nearly 81% variance in the observed activities. In the randomization study (100 simulations per model), none of the identified models has shown any chance correlation. The values greater than 0.5 of Q^2 index is in accordance to a reasonable robust QSAR model. The pIC_{50} values of training set compounds calculated using Eqs. (3.35) to (3.38) and predicted from LOO procedure have been included in Table 3.14.

Table 3.14: Observed and modeled DPP-4 inhibitory activities.

S. No.	$pIC_{50}(M)$[a]										
		Eq. (3.35)		Eq. (3.36)		Eq. (3.37)		Eq. (3.38)		PLS	
	Obs[b]	Cal	Pre[c]	Cal	Pre[c]	Calc	Pre[c]	Cal	Pre[c]	Cal	Pre[c]
1[d]	7.74	8.48	-[d]	8.42	-[d]	8.19	-[d]	7.59	-[d]	8.18	-[d]
2	7.64	7.37	7.32	7.60	7.60	7.65	7.65	7.54	7.53	7.50	7.49
3	7.85	8.21	8.29	8.05	8.08	7.96	7.98	7.76	7.75	8.08	8.11
4	7.04	6.96	6.94	6.67	6.52	6.87	6.79	7.27	7.28	6.68	6.64
5	7.38	7.32	7.31	7.65	7.68	7.38	7.38	7.73	7.76	7.43	7.44
6[d]	7.06	7.65	-[d]	7.63	-[d]	6.82	-[d]	7.74	-[d]	7.39	-[d]
7	8.37	8.03	7.98	8.03	7.97	7.99	7.94	8.69	8.75	8.33	8.33
8	7.04	7.11	7.15	6.86	6.73	6.97	6.93	7.47	7.70	7.04	7.04
9	6.76	7.02	7.17	6.89	6.95	7.19	7.31	6.61	6.54	6.86	6.87
10[d]	7.00	6.64	-[d]	7.12	-[d]	7.24	-[d]	7.42	-[d]	6.99	-[d]
11[d]	6.68	7.00	-[d]	6.89	-[d]	6.97	-[d]	7.39	-[d]	6.97	-[d]
12	7.92	7.74	7.72	7.55	7.50	7.60	7.56	7.69	7.67	7.65	7.62
13	7.96	7.90	7.89	7.77	7.75	7.68	7.65	8.00	8.01	7.89	7.87
14	6.95	7.18	7.20	7.53	7.57	7.70	7.75	7.81	7.87	7.59	7.63
15	8.30	7.97	7.94	7.90	7.87	8.08	8.06	7.64	7.60	7.83	7.80
16	6.91	6.92	6.93	6.83	6.81	6.44	6.36	6.89	6.89	6.74	6.72
17	8.24	8.32	8.33	8.16	8.15	8.42	8.44	7.63	7.36	7.92	7.86
18	8.82	8.52	8.42	8.42	8.27	7.82	7.74	8.70	8.66	8.42	8.39
19	7.49	8.27	8.36	8.43	8.56	8.37	8.49	8.12	8.21	8.38	8.45
20	6.42	7.24	7.71	7.88	8.03	7.44	7.68	6.70	6.77	6.89	6.94
21[d]	8.55	7.67	-[d]	7.59	-[d]	7.60	-[d]	7.45	-[d]	7.40	-[d]
22	6.85	7.03	7.06	6.77	6.76	6.64	6.61	6.55	6.50	6.70	6.68
23	9.18	9.60	9.72	9.24	9.26	9.28	9.32	9.33	9.37	9.48	9.54

24	6.49	6.67	6.74	6.75	6.85	6.36	6.27	6.07	5.91	6.39	6.35
25	9.37	9.48	9.54	8.94	8.81	8.93	8.80	8.98	8.87	9.29	9.24
26[d]	6.36	7.83	-[d]	8.07	-[d]	6.61	-[d]	6.86	-[d]	7.06	-[d]
27[d]	9.51	9.43	-[d]	9.17	-[d]	8.74	-[d]	9.13	-[d]	8.94	-[d]
28[d]	6.88	7.51	-[d]	7.60	-[d]	7.76	-[d]	6.60	-[d]	7.26	-[d]
29[d]	9.34	9.40	-[d]	9.49	-[d]	9.17	-[d]	8.98	-[d]	9.44	-[d]
30	6.94	5.96	5.68	6.63	6.44	6.96	6.98	6.15	5.96	6.53	6.48
31	9.74	8.31	8.02	8.56	8.19	9.19	9.04	9.00	8.70	9.07	8.97
32	6.37	6.52	6.54	6.51	6.53	7.14	7.28	6.84	6.96	6.73	6.82
33	9.49	9.57	9.62	9.91	10.11	10.09	10.35	9.34	9.30	9.73	9.81
34	7.05	7.43	7.50	7.16	7.17	7.53	7.58	7.80	7.95	7.69	7.72
35	9.30	8.68	8.24	9.24	9.20	8.98	8.71	8.75	8.36	8.89	8.78
36	6.23	6.85	7.24	6.38	6.66	6.27	6.33	7.30	8.24	6.49	6.60
37[d]	8.20	8.19	-[d]	7.79	-[d]	7.71	-[d]	9.07	-[d]	8.08	-[d]
38	6.88	6.80	6.79	6.78	6.76	6.65	6.61	6.65	6.61	6.75	6.74
39	9.40	9.40	9.40	9.29	9.27	8.80	8.64	9.34	9.32	9.41	9.41

[a]On molar basis; [b]Taken from ref. [737]; [c]Leave-one-out (LOO) procedure; [d]Compound included in test set.

The models (3.35) to (3.38) are validated with an external test set of 10 compounds listed in Table 3.12. The predictions of the test set compounds based on external validation are found to be satisfactory as reflected in the test set r^2 (r^2_{Test}) values and the same is also reported in Table 3.14.

A partial least square (PLS) analysis has been carried out on these 10 CP-MLR identified descriptors (Table 3.13) to facilitate the development of a "single window" structure–activity model. In the PLS cross-validation, three components are found to be the optimum for these 10 descriptors and they explained 93.9% variance in the activity (r^2 = 0.939, Q^2_{LOO} = 0.853, s = 0.391, F = 62.879, r^2_{Test} = 0.763). The MLR-like PLS coefficients of these 10 descriptors are given in Table 3.15. The plot showing goodness of fit between observed and calculated activities (through Eqs. (3.35) to (3.38) and PLS analysis) for the training and test set compounds is given in Figure 3.9. Figure 3.10 shows a plot of the fraction contribution of normalized regression coefficients of these descriptors to the activity.

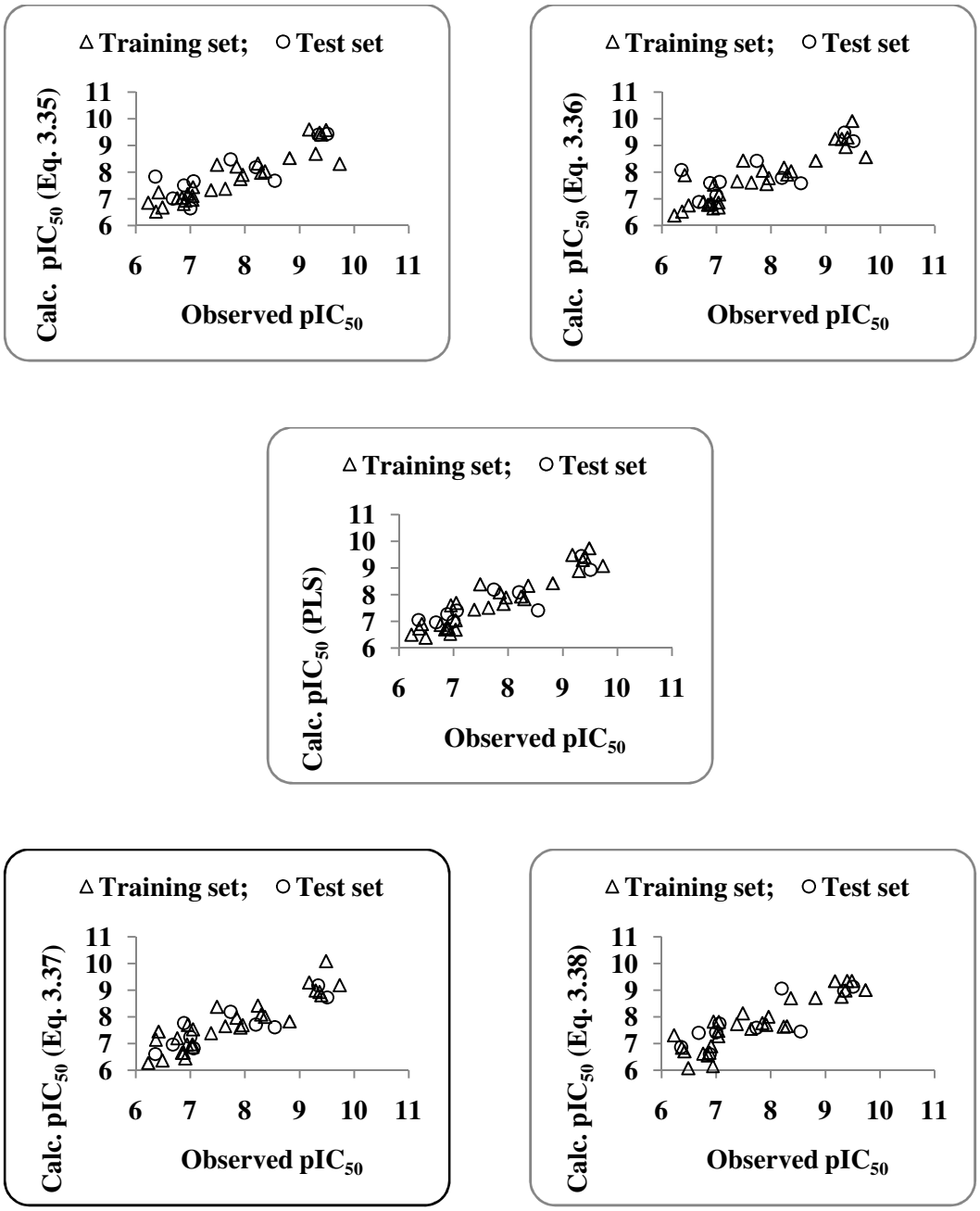


Figure 3.9: Plot of observed and calculated pIC_{50} values of training- and test-set compounds for DPP-4.

Table 3.15: PLS and MLR-like PLS models from the descriptors of five parameter CP-MLR models for DPP-4 inhibitory activity.

A: PLS equation			
PLS components		PLS coefficient (s.e.)[a]	
Component-1		-0.859(0.064)	
Component-2		0.091(0.036)	
Component-3		0.182(0.069)	
Constant		7.737	
B: MLR-like PLS equation			
S. No.	Descriptor	MLR-like coefficient (f.c.)[b]	Order
1	DISPv	-0.159(-0.057)	9
2	RDF075m	-0.339(-0.122)	3
3	RDF085m	0.314(0.113)	5
4	RDF110e	-0.449(-0.162)	1
5	RDF155e	-0.092(-0.033)	10
6	RDF085p	0.333(0.120)	4
7	Mor10m	0.361(0.130)	2
8	Mor12m	0.220(0.079)	8
9	G3e	-0.226(-0.081)	7
10	RTu+	0.272(0.098)	6
	Constant	7.196	
C: PLS regression statistics		Values	
n		29	
r		0.939	
s		0.391	
F		62.879	
FIT		4.964	
LOF		0.210	
AIC		0.202	
Q^2_{LOO}		0.853	
Q^2_{L5O}		0.858	
r^2_{Test}		0.763	

[a]Regression coefficient of PLS factor and its standard error. [b]Coefficients of MLR-like PLS equation in terms of descriptors for their original values; f.c. is fraction contribution of regression coefficient, computed from the normalized regression coefficients obtained from the autoscaled (zero mean and unit s.d.) data.

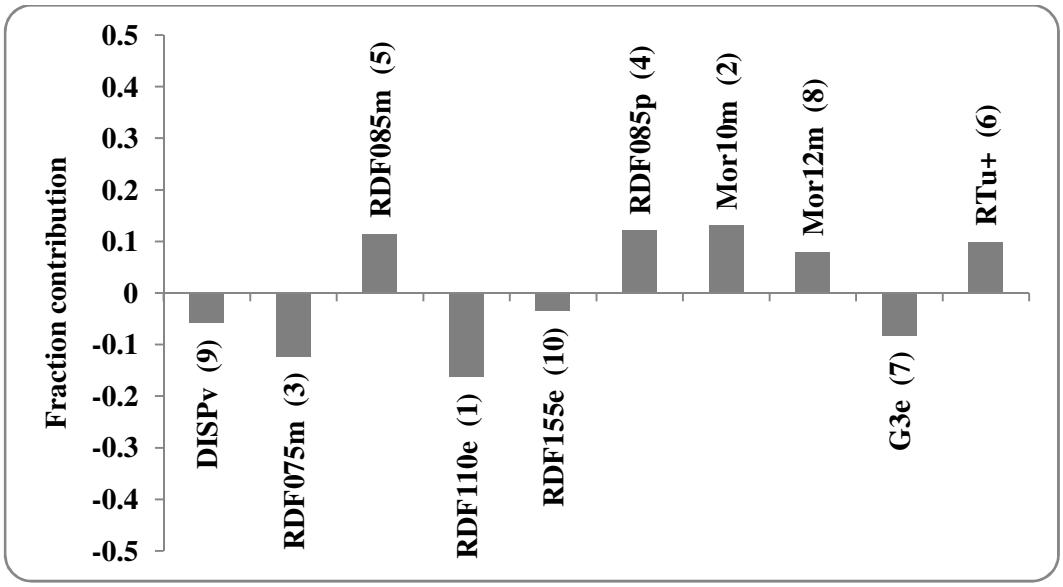


Figure 3.10: Plot of fraction contribution of MLR-like PLS coefficients (normalized) against ten CP-MLR identified descriptors (Table 3.15) associated with DPP-4 inhibitory activity of triazolopiperazines.

The PLS analysis has suggested RDF110e as the most determining descriptor for modeling the activity of the compounds (descriptor S. No. 4 in Table 4; Figure 2). The other nine significant descriptors in decreasing order of significance are Mor10m, RDF075m, RDF085p, RDF085m, RTU+, G3e, Mor12m, DISPv and RDF155e. All descriptors are part of Eqs. (3.29) to (3.38) and convey same inference in the PLS model as well. It is also observed that PLS model from the dataset devoid of 10 descriptors (Table 3.13) is inferior in explaining the activity of the analogues.

The other inhibitory activity reported for DPP-8 enzyme system has also analyzed quantitatively. A total number of 10 models in two parameters and 46 models in three parameters, having $r^2_{Test} > 0.5$, were obtained on applying CP-MLR. For the sake of brevity, highly significant four models in three parameters emerged through CP-MLR are shown below.

$pIC_{50} = 3.193 + 2.361(0.305)RDF085m – 2.142(0.718)RDF150p$

$+ 1.732(0.582)Mor10p$

n = 19, r = 0.915, s = 0.390, F = 25.793, Q^2_{LOO} = 0.741, Q^2_{L5O} = 0.738
r^2_{Test} = 0.523, FIT = 2.763, LOF = 0.257, AIC = 0.234 (**3.39**)

pIC_{50} = 3.847 + 2.302(0.325)RDF085m – 1.778(0.726)RDF150p
+ 1.237(0.479)Mor10m
n = 19, r = 0.906, s = 0.409, F = 23.001, Q^2_{LOO} = 0.747, Q^2_{L5O} = 0.724
r^2_{Test} = 0.697, FIT = 2.464, LOF = 0.283, AIC = 0.257 (**3.40**)

pIC_{50} = 7.192 + 2.588(0.594)RDF115m – 5.919(1.176)Mor23m
– 1.685(0.728)E3e
n = 19, r = 0.905, s = 0.411, F = 22.788, Q^2_{LOO} = 0.709, Q^2_{L5O} = 0.756
r^2_{Test} = 0.577, FIT = 2.441, LOF = 0.285, AIC = 0.259 (**3.41**)

pIC_{50} = 9.741 – 1.025(0.450)RDF145p – 8.388(1.092)Mor23m
– 1.702(0.442)H5m
n = 19, r = 0.900, s = 0.422, F = 21.355, Q^2_{LOO} = 0.703, Q^2_{L5O} = 0.719
r^2_{Test} = 0.574, FIT = 2.288, LOF = 0.300, AIC = 0.273 (**3.42**)

It is evident from the models that higher values of atomic mass weighted radial distribution functions 8.5 and 11.5 (descriptors RDF085m and RDF115m, respectively), and lower values of atomic polarizabilities weighted radial distribution functions 14.5 and 15.0 (descriptors RDF145p and RDF150p, respectively) would supplement the activity. Atomic mass weighted 3D-MoRSE signals 10 and 23 (descriptors Mor10m and Mor23m, respectively) in addition to atomic polarizabilities weighted 10 (descriptor Mor10p) have shown prevalence to explain the DPP-8 inhibitory activity. A higher value of descriptors Mor10p and Mor10m is conducive to activity whereas a higher value of descriptor Mor23m is unfavorable to the activity. The negative correlation of WHIM descriptor (E3e, 3rd component accessibility WHIM index/ weighted by atomic Sanderson electronegativities) and physicochemical properties weighted spatial autocorreation GETAWAY descriptor (H5m, H autocorrelation of lag 5/weighted

by atomic masses) recommended a lower value of these descriptors for elevated DPP-8 inhibitory activity. These models are able to explain nearly 84% variance in the observed DPP-8 inhibitory activities. None of the identified models has shown any chance correlation in the randomization study (100 simulations per model). The values greater than 0.5 of Q^2 index is in accordance to a reasonable internal validation and r^2_{Test} values reflect upon the good predictive power of above mentioned QSAR models. The pIC_{50} values of training and test set compounds calculated using Eqs. (3.39) to (3.42) and predicted from LOO procedure have been included in Table 3.16. The goodness of fit or agreement between observed and calculated activities for the training and test set compounds is shown in Figure 3.11.

Table 3.16: Observed and modeled DPP-8 inhibitory activities.

S. No.	$pIC_{50}(M)$[a]								
		Eq. (3.39)		Eq. (3.40)		Eq. (3.41)		Eq. (3.42)	
	Obsd[b]	Calc	Pred[c]	Calc	Pred[c]	Calc	Pred[c]	Calc	Pred[c]
1[d]	4.32	4.51	-[d]	4.35	-[d]	4.25	-[d]	4.96	-[d]
2	4.64	4.55	4.53	4.21	4.10	4.67	4.68	4.74	4.75
3	4.48	4.01	3.91	4.13	4.07	4.41	4.39	4.31	4.27
4[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
5	4.12	4.49	4.60	4.32	4.38	4.30	4.34	4.18	4.21
6[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
7	4.77	4.96	4.99	4.96	4.99	4.81	4.81	5.08	5.18
8	4.18	4.55	4.62	4.74	4.80	4.17	4.17	4.22	4.23
9	5.22	4.75	4.50	4.54	4.33	4.81	4.62	5.58	5.61
10[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
11[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
12	4.15	4.24	4.26	4.33	4.35	4.37	4.40	4.30	4.33
13	4.36	4.78	4.81	4.87	4.90	4.63	4.65	4.42	4.43
14[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
15	5.10	4.79	4.76	4.85	4.82	4.56	4.49	4.94	4.91
16[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
17	5.80	5.52	5.48	5.75	5.74	5.71	5.69	5.27	5.05
18	5.52	4.99	4.89	5.07	5.00	5.05	4.94	4.85	4.76
19	4.14	4.94	5.02	4.98	5.07	4.73	4.97	4.29	4.33
20[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
21[d]	4.52	4.38	-[d]	4.88	-[d]	4.19	-[d]	4.89	-[d]
22	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]

23	6.21	6.11	6.09	6.04	6.00	5.58	5.51	6.18	6.17
24	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
25	6.44	6.49	6.51	6.52	6.54	6.72	6.85	6.56	6.60
26[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
27[d]	5.10	5.08	-[d]	5.66	-[d]	5.89	-[d]	5.42	-[d]
28[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
29[d]	5.96	5.08	-[d]	5.67	-[d]	5.98	-[d]	6.28	-[d]
30[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
31	6.48	6.42	6.40	6.35	6.30	6.53	6.55	6.51	6.52
32[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
33	6.49	6.76	6.92	6.43	6.42	6.53	6.56	6.08	5.80
34	4.40	4.63	4.72	4.57	4.66	5.29	5.55	5.52	5.74
35	6.20	5.95	5.88	6.19	6.18	5.96	5.89	5.86	5.80
36[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
37[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
38[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
39	5.30	5.07	4.76	5.16	4.96	5.16	5.13	5.10	5.00

[a]On molar basis; [b]Taken from ref. [737]; [c]Leave-one-out (LOO) procedure; [d]Compound included in test set and [e]Compound with uncertain activity, not part of data set.

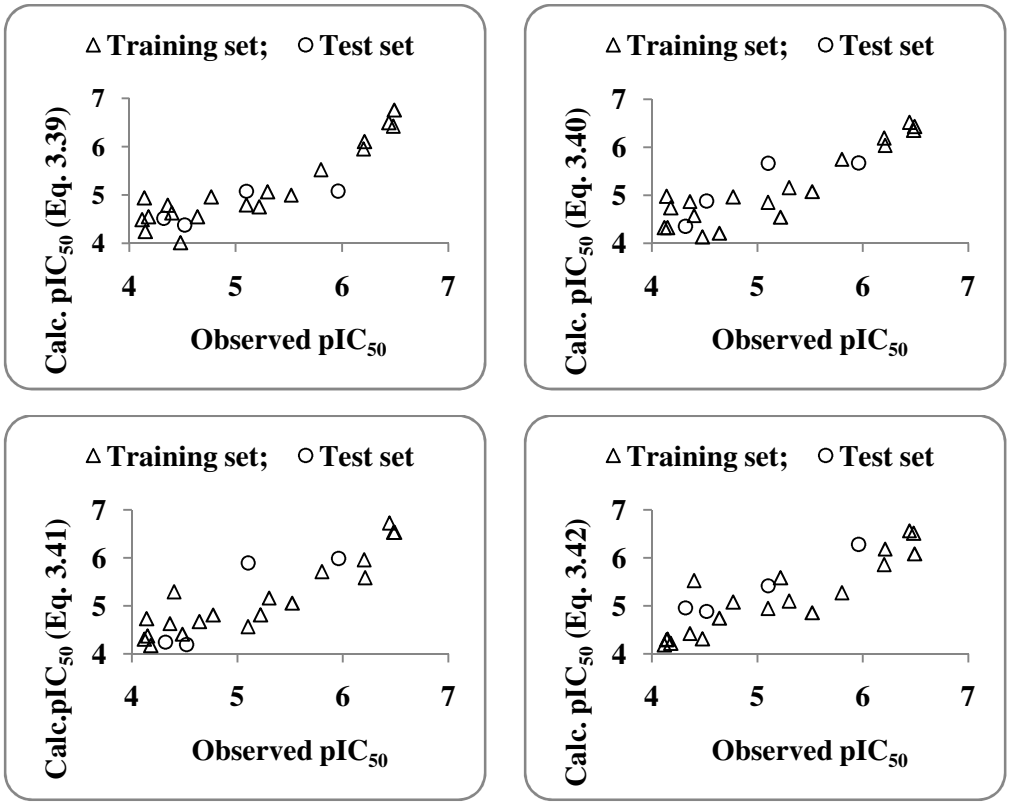


Figure 3.11: Plot of observed and calculated pIC_{50} values for DPP-8.

2.3.1.2. Applicability domain (AD)

On analyzing the model AD in the Williams plot (Figure 3.12) of the model based on the whole dataset (Table 3.17), it has appeared that none of the compounds were identified as an obvious outlier for the DPP-4 inhibitory activity if the limit of normal values for the *Y* outliers (response outliers) was set as 3 (standard deviation) units. One of the compounds (S. No. 37, Table 3.12) was found to have leverage (*h*) values greater than the threshold leverages (*h**) in a plot derived for Eqn. (3.38a) reflecting it as a chemically influential compound.

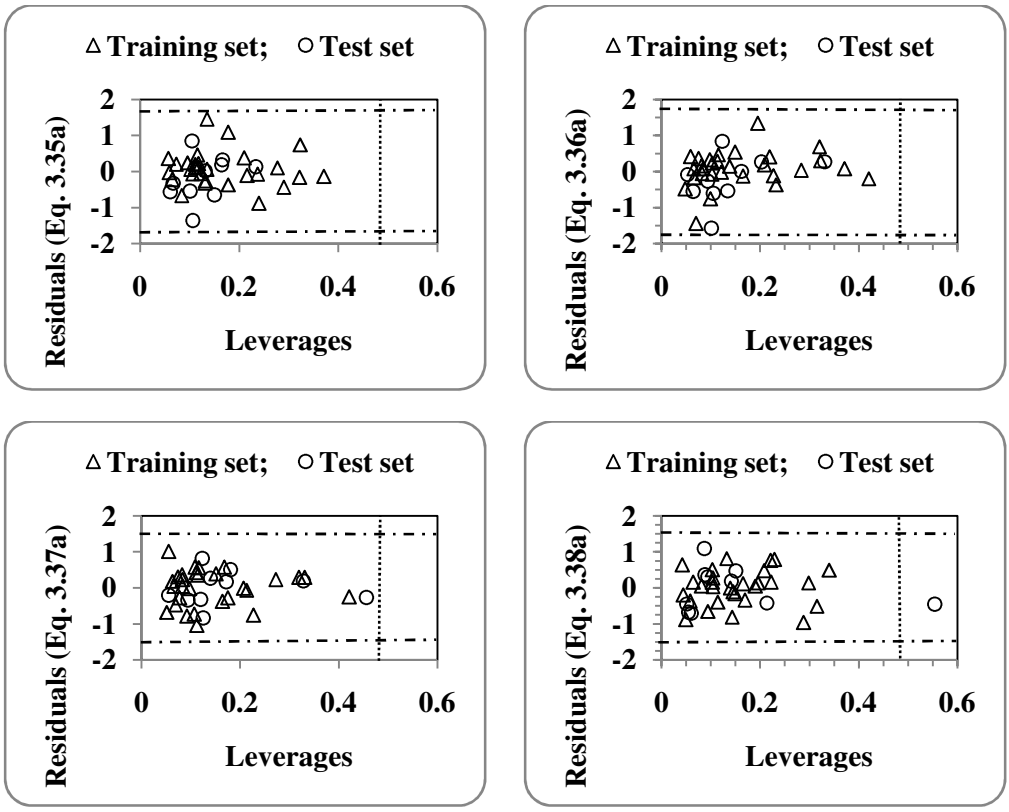


Figure 3.12: Williams plot for the training-set and test- set compounds for DPP-4 inhibitory activity. The horizontal dotted line refers to the residual limit (±3×standard deviation) and the vertical dotted line represents threshold leverage h* (= 0.46).

Table 3.17: Models derived for the whole data set (n = 39) for the DPP-4 inhibitory activity in descriptors identified through CP-MLR.

Model	r	s	F	Eq.
pIC_{50} = 7.197 – 1.490(0.627)DISPv –2.823(0.602)RDF110e +1.505(0.497)Mor10m +3.198(0.527)RDF085p + 1.662(0.587)RTu+	0.875	0.559	21.562	(3.35a)
pIC_{50} = 5.414 – 2.495(0.586)RDF110e +3.383(0.537)RDF085p +2.209(0.474)Mor10m + 0.864(0.530)Mor12m + 1.788(0.608)RTu+	0.863	0.582	19.383	(3.36a)
pIC_{50} = 6.582 – 3.073(0.427)RDF110e –2.047(0.468)RDF155e +2.889(0.448)Mor10m +2.772(0.419)RDF085p +1.321(0.459)Mor12m	0.892	0.520	25.914	(3.37a)
pIC_{50} = 7.778 – 4.456(0.682)RDF075m +5.559(0.715)RDF085m+1.223(0.463)Mor10m – 2.183(0.455)G3e + 1.251(0.485)RTu+	0.880	0.547	22.780	(3.38a)

For both the training set and test set, the suggested model matches the high-quality parameters with good fitting power and the capability of assessing external data. Furthermore, almost all of the compounds was within the AD of the proposed model and were evaluated correctly.

2.3.3. CONCLUSIONS

The DPP-4 and DPP-8 inhibitory activity of triazolopiperazines have been quantitatively analyzed in terms of 3D-Dragon descriptors. The derived QSAR models have shown that atomic properties played pivotal role in terms of weighted radial distribution functions, 3D-MoRSE signals, component symmetry directional WHIM index and moment expansions. The CP-MLR indentified RDF, 3D-MoRSE, WHIM and GETAWAY descriptors weighted or unweighted with atomic properties endow relevant molecular 3D information about molecular size, shape, symmetry, atom distribution, effective position of substituents and fragments in the molecular space hold promise for rationalizing the DPP-4 and DPP-8 inhibitory actions of triazolopiperazines. The values of statistical parameters, Q^2_{LOO} and r^2_{Test} ensure that the models have validated internally and externally, both and the predictions are reliable and acceptable. PLS analysis has further confirmed the dominance of the CP-MLR identified descriptors.

Applicability domain analysis revealed that the suggested models have acceptable predictability. All the compounds are within the applicability domain of the proposed models and were evaluated correctly.

REFERENCES

699. Wideman RD and Kieffer TJ. Trends Endocrinol Metab (2009) 20:280-286.

700. Triplitt C, Wright A and Chiquette E. Pharmacotherapy (2006) 26:360-374.

701. Gallwitz B. Treat Endocrinol (2005) 4:361-370.

702. Holst JJ. Gastroenterology (1994) 107:1048-1055.

703. Drucker DJ. Diabetes (1998) 47:159-169.

704. Vilsboll T and Holst JJ. Diabetologia (2004) 47:357-366.

705. Elahi D, McAloon-Dyke M, Fukagawa NK, Meneilly GS, Sclater AL, Minaker KL, Habener JF and Andersen DK. Regul Pept (1994) 51:63-75.

706. Nauck MA, Bartels E, Oerskov C, Ebert R and Creutzfedlt WJ. J Clin Endocrinol Metabol (1993) 76:912- 917.

707. Gentilella R, Bianchi C, Rossi A and Rotella CM. Diabetes Obes Metab (2009) 11:544-556.

708. Gonzalez C, Beruto V, Keller G, Santoro S and Di Girolamo G. Expert Opin Invest Drugs (2006) 15:887-895.

709. Holst JJ. Diabetes Metab Res and Rev (2002) 18:430-441.

710. Hansen L, Deacon CF, Orskov C and Holst JJ. Endocrinology (1999) 140: 5356-5363.

711. Nagakura T, Yasuda N, Yamazaki K, Ikuta H, Yoshikawa S, Asano O and Tanaka I. Biochem Biophys Res Commun (2001) 284:501-506.

712. Marguet D, Baggio L, Kobayashi T, Bernard A, Pierres M, Nielsen PF, Ribel U, Watanabe T, Drucker DJ and Wagtmann N. Proc Natl Acad Sci USA (2000) 97:6874-6879.

713. Durinx C, Lambeir A, Bosmans E, Falmagne J, Berghmans R, Haemers A, Scharpe S and De Meester I. Eur J Biochem (2000) 267:5608-5613.

714. Iwaki-Egawa S, Watanabe Y, Kikuya Y and Fujimoto Y. J Biochem

(1998) 124:428-433.

715. Villhauer EB, Brinkman JA, Naderi GB, Burkey BF, Dunning BE, Prasad K, Mangold BL, Russell ME and Hughes TE. J Med Chem (2003) 46:2774-2789.

716. Kim D, Wang L, Beconi M, Eiermann GJ, Fisher MH, He H, Hickey GJ, Kowalchick JE, Leiting B, Lyons K, Marsilio F, McCann ME, Patel RA, Petrov A, Scapin G, Patel SB, Roy RS, Wu JK, Wyvratt MJ, Zhang BB, Zhu L, Thornberry NA and Weber AE. J Med Chem (2005) 48:141-151.

717. Augeri DJ, Robl JA, Betebenner DA, Magnin DR, Khanna A, Robertson JG, Wang A, Simpkins LM, Taunk P, Huang Q, Han SP, Abboa-Offei B, Cap M, Xin L, Tao L, Tozzo E, Welzel GE, Egan DM, Marcinkeviciene J, Chang SY, Biller SA, Kirby MS, Parker RA and Hamann LG. J Med Chem (2005) 48:5025-5037.

718. Feng J, Zhang Z, Wallace MB, Stafford JA, Kaldor SW, Kassel DB, Navre M, Shi L, Skene RJ, Asakawa T, Takeuchi K, Xu R, Webb DR and Gwaltney SL. J Med Chem (2007) 50:2297-2300.

719. Havale SH and Pal M. Bioorg Med Chem (2009) 17:1783-1802

720. Brigance RP, Meng W, Fura A, Harrity T, Wang A, Zahler R, Kirby MS and Hamman LG. Bioorg Med Chem Lett (2010) 20:4395-4398

721. Meng W, Brigance RP, Chao HJ, Fura A, Harrity T, Marcinkkeviciene J, O'Connor SP, Tamura JK, Xie D, Zhang Y, Klei HE, Kish K, Weigett CA, Turdi H, Wang A, Zahler R, Kirby MS and Hamann LG. J Med Chem (2010) 53:5620-5628.

722. Yaron A and Naider F. Crit Rev Biochem Mol Biol (1993) 28:31-81.

723. Mentlein R, Gallwitz B and Schmidt WE. Eur J Biochem (1993) 214: 829-835.

724. Kieffer TJ, McIntosh CHS and Pederson RA. Endocrinology (1995) 136:3585-3596.

725. Deacon CF, Nauck MA, Toft-Nielsen M, Pridal L, Willms B and Holst JJ.

Diabetes (1995) 44:1126-1131.

726. Ashworth DM, Atrash B, Baker GR, Baxter AJ, Jenkins PD, Jones DM and Szelke M. Bioorg Med Chem Lett (1996) 6:2745-2748.

727. Ashworth DM, Atrash B, Baker GR, Baxter AJ, Jenkins PD, Jones DM and Szelke M. Bioorg Med Chem Lett (1996) 6:1163-1166.

728. Ahren B, Gomis R, Standl E, Mills D and Schweizer AA. 64th Scientific Sessions of the American Diabetes Association. Orlando,FL. (2004) Late-breaking 7-LB (Abstract).

729. Tsu H, Chen X, Chen C-T, Lee S-J, Chang C-N, Kao K-H, Coumar MS, Yeh Y-T, Chien C-H, Wang H-S, Lin K-T, Chang Y-Y, Wu S-H, Chen Y-S, Lu I-L, Wu S-Y, Tsai T-Y, Chen W-C, Hsieh H-P, Chao Y-S and Jiaang W-T. J Med Chem (2006) 49:373-380.

730. Kim D, Wang L, Beconi M, Eiermann GJ, Fisher MH, He H, Hickey GJ, Kowalchick JE, Leiting B, Lyons K, Marsilio F, McCann ME, Patel RA, Petrov A, Scapin G, Patel SB, Sinha Roy R, Wu JK, Wyvratt MJ, Zhang BB, Zhu L, Thornberry NA and Weber AE. J Med Chem (2005) 48:141-151.

731. Edmondson SE, Fisher MH, Kim D, Maccoss M, Parmee ER, Weber AE and Xu J. US Patent 6699871B2, March 2, 2004.

732. Villhauer EB, Brinkman JA, Naderi GB, Burkey BF, Dunning BE, Prasad K, Mangold BL, Russell ME and Hughes TE. J Med Chem (2003) 46: 2774-2789.

733. Magnin DR, Robl JA, Sulsky RB, Augeri DJ, Huang Y, Simpkins LM, Taunk PC, Betebenner DA, Robertson JG, Abboa-Offei BE, Wang A, Cap M, Xin L, Tao L, Sitkoff DF, Malley MF, Gougoutas JZ, Khanna A, Huang Q, Han S, Parker RA and Hamann LG. J Med Chem (2004) 47:2587-2598.

734. Kim D, Kowalchick JE, Edmondson SD, Mastracchio A, Xu JM, Wu JK,

Eiermann GJ, Leiting B, Pryor K, Patel RA, He H, Lyons KA, Thornberry NA and Weber AE. Bioorg Med Chem Lett (2007) 17:3373-3377.

735. Xu J, Ok HO, Gonzalez EJ, Colwell LF Jr, Habulihaz B, He H, Leiting B, Lyona KA, Marsilio F, Patel RA, Wu JK, Thornberry NA, Weber AE and Parmee ER. Bioorg Med Chem Lett (2004) 14:4759-4762.

736. Brockunier L, He J, Colwell LF Jr, Habulihaz B, He H, Leiting B, Lyons KA, Marsilio F, Patel R, Teffera Y, Wu JK, Thornberry NA, Weber AE and Parmee ER. Bioorg Med Chem Lett (2004) 14:4763-4766.

737. Kim D, Kowalchick JE, Brockunier LL, Parmee ER, Eiermann GJ, Fisher MH, He H, Leiting B, Lyons K, Scapin G, Patel SB, Petrov A, Pryor KD, Roy RS, Wu JK, Zhang X, Wyvratt MJ, Zhang BB, Zhu L, Thornberry N and Weber AE. J Med Chem (2008) 51:589-602.

738. Hemmer MC, Steinhauer V and Gasteiger J. Vib Spectrosc (1999) 19:151-164.

739. Silverman BD. J Chem Inform Comput Sci (2000) 40:1470-1476.

740. Schuur JH, Selzer P and Gasteiger J. J Am Chem Soc (1996) 36:334-344.

741. (a) Consonni V, Todeschini R and Pavan M. J Chem Inform Comput Sci (2002) 42:682-692. (b) Consonni V, Todeschini R, Pavan M and Gramatica P. J Chem Inform Comput Sci (2002) 42:693-705.

742. (a) Todeschini R, Lasagni M and Marengo E. J Chemom (1994) 8:263-273. (b) Todeschini R and Gramatica P. 3D QSAR in drug design. (Kubinyi H, Folkers G and Martin YC Eds.), Kluwer/ESCOM, Dordrecht (The Netherlands) (1998) Vol 2:355-380.

Chapter 4

QSAR Studies on PPAR Agonists

CHAPTER 4

QSAR STUDIES ON PPAR AGONISTS

1. INTRODUCTION

Elevated plasma glucose in the presence of high endogenous insulin levels is characteristic of type 2 diabetes (T2D). T2D is a complex metabolic disorder because insulin resistance and impaired insulin secretion lead to abnormal metabolism of glucose, lipids and amino acids. The quality of life of diabetic patients slowly decreases due to developed long-term micro- and macro- vascular complications like neuropathy, retinopathy, nephropathy, myocardial infarction, stroke, and lower limb amputation as the progression of the disease progresses [743, 744]. Thus, T2D come with a defect in pancreatic β-cell and is characterized by resistance of insulin in the liver and peripheral tissues [745] and due to lack of physical activity and excessive food intake, is presumed to attain epidemic proportions [746], become a chronic metabolic disorder. The prevalence of T2D in developed and developing countries is rising speedily and it is expected that number of diabetics to reach 380 million by 2025 [747, 748].

The treatment of T2D is currently aimed at to improve insulin secretion by reducing hyperglycemia or to reduce the insulin resistance of peripheral tissues. Most of such types of commonly used therapies were developed without considering therapeutic target. Therefore, attempts were made to identify more suitable therapeutic strategies with better insight of the disease's pathogenesis [749]. Peroxisomes proliferators activated receptors (PPARs), belonging to the family of nuclear receptors, are ligand-activated transcription factors [750]. Three subtypes namely PPARα, PPARγ and PPARβ/(δ) have been identified after the discovery in 1990 by Issemen and Green [751]. These receptors are extensively involved in glucose and lipid homeostasis [752-754]. A number of agonists in this class have progressed to the clinical phase and marketed as anti-diabetic drugs [755, 756]. Among the PPAR subtypes the most extensively investigated subtype is PPARγ.

2. MODELING STUDIES

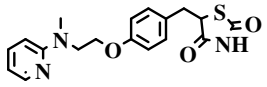
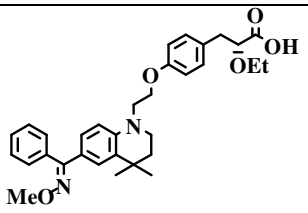
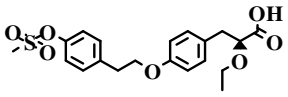
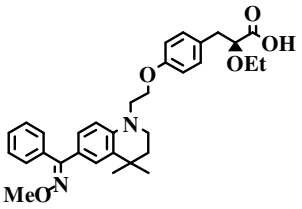
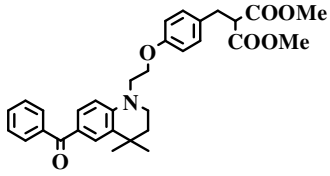
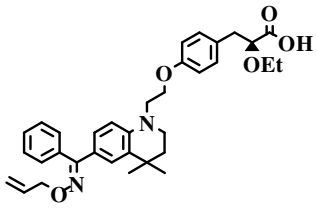
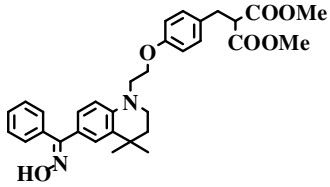
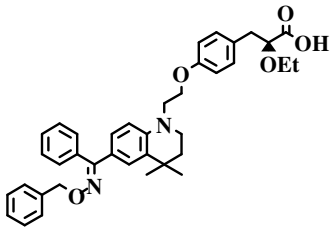
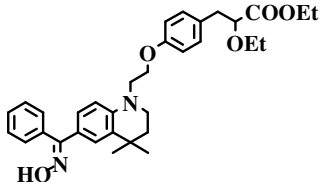
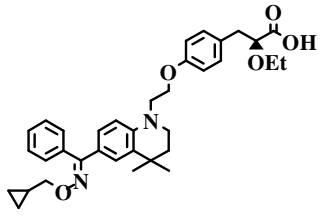
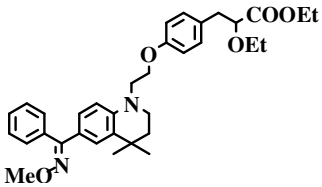
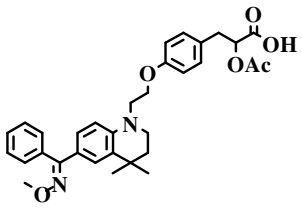
2.1. 4,4-DIMETHYL-1,2,3,4-TETRAHYDROQUINOLINES AS PARα/γ AGONIST

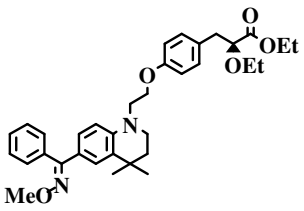
In present era of drug development, the hypolipidemic fibrates and glitazones class of insulin sensitizers, full-agonists of PPARα [750] and PPARγ [757, 758], respectively, has motivated pharmaceutical companies to focus on developing more potent and dual acting agonists belonging to these two subtypes. In the treatment of dyslipidemic T2D dual-acting PPARα/γ agonists such as Tesaglitazar and Muraglitazar have been observed as a very attractive option [752, 756, 759-764]. These compounds may also circumvent or reduce the main side effects such as weight gain or edema induced by the full PPARγ agonists like TZDs [765]. The ligand–protein interactions of a typical PPAR agonists revealed that the acidic head group of ligand, known as carboxylic acid, is involved in up to four hydrogen bonds with the receptor which is crucial part for activation of PPAR. The central aromatic moiety is located in a hydrophobic pocket while the cyclic tail tolerates more polar substituents [754]. Based on the typical topology of synthetic PPAR agonists 4,4-dimethyl-1,2,3,4-tetrahydroquinoline has been considered as novel cyclic tail to design novel PPARγ selective agonists and/or dual PPARα/γ agonists [766]. A new series of 4,4-dimethyl-1,2,3,4-tetrahydroquinoline-based compounds as effective PPARγ selective agonists and dual-acting agonists of PPARα and PPARγ has been reported [767, 768].

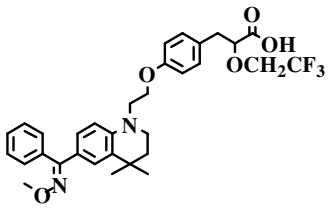
The reported eighteen tetrahydroquinoline derivatives are considered as the data set for this study [767, 768]. The structures of these analogues are given in Table 4.1. These derivatives were evaluated for binding affinity to human PPARγ using a competitive binding assay with [^{3}H]Rosiglitazone. Functional activity was determined in a transient transfection assay using pGAL4hPPARα and pGAL4hPPARγ [767, 768].

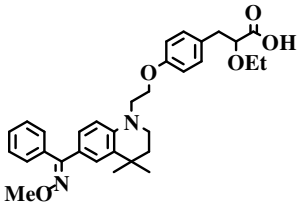
The reported binding affinity in terms of pK_i(M) and transactivation activity in terms of pEC_{50}(M) of these congeners is presented in Table 4.2.

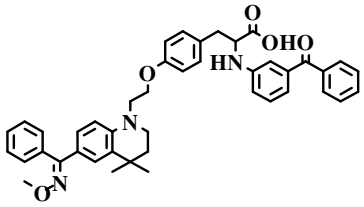
Table 4.1: Structures[a] of tetrahydroquinoline derivatives.

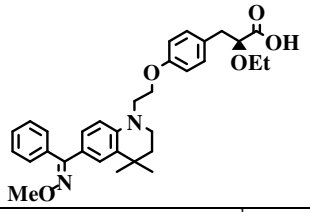
Cpd.	Structure	Cpd.	Structure
1[b]		10	
2[c]		11	
3		12	
4		13	
5		14	
6		15	

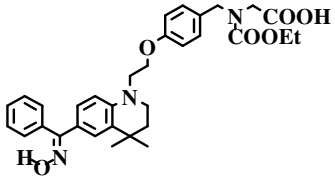
7 

16 

8 

17 

9 

18 

[a]Taken from reference [767, 768], [b]Rosiglitazone and [c]Tesaglitazar.

Table 4.2: Reported biological actions of tetrahydroquinoline derivatives.

Cpd.	Obsd.[a]			Cpd.	Obsd.[a]		
	Binding	Transactivation			Binding	Transactivation	
	PPARγ	hPPARα	hPPARγ		PPARγ	hPPARα	hPPARγ
	pK_i (M)[b]	pEC_{50}(M)[c]			pK_i (M)[b]	pEC_{50}(M)[c]	
1	8.10	5.00	8.40	10	7.05	6.71	7.19
2	7.74	6.38	7.43	11[d]	7.74	6.94	8.11
3	5.00	-[e]	6.82	12	7.27	7.14	7.89
4	5.00	-[e]	6.71	13	7.74	7.47	8.15
5	6.21	6.72	7.32	14[d]	6.60	8.05	8.12
6	5.00	7.52	7.72	15	-[e]	-[e]	-[e]
7[d]	5.00	7.54	7.85	16	7.48	6.00	6.90
8[d]	7.32	7.92	7.96	17	7.27	5.00	7.85
9	-[e]	7.54	7.85	18	-[e]	-[e]	5.00

[a]Taken from ref. [767, 768]; [b]On molar basis, K_i represents the binding affinity to human PPARγ; [c]On molar basis; [d]Compound included in test set; and [e]Inactive compound, not part of data set.

For modeling purpose the data set has been sub-divided into training set (for model development) and test set (for external prediction or validation). The selection of test set compounds was made using an in-house written randomization program. The test and training set compounds are mentioned in Table 4.2.

2.1.1. RESULTS AND DISCUSSION

2.1.1.1. QSAR RESULTS

For the compounds in Table 4.1, a total number of 479 descriptors belonging to 0D- to 2D- classes of DRAGON have been computed and were subjected to CP-MLR analysis. Descriptors which are inter-correlated beyond 0.9 (descriptor vs. descriptor, $r > 0.9$) and poorly correlated with biological actions (descriptor vs. activity, $r < 0.1$) has been excluded prior to the application of CP-MLR procedure. In this way the reduced descriptor data set contained 55, 39 and 67 as relevant descriptors for PPARγ binding, and hPPARα and hPPARγ transactivation activities, respectively. The descriptors have been scaled between the intervals 0 to 1 [769] to ensure that a descriptor will not dominate simply because it has larger or smaller pre-scaled value compared to the other descriptors and the scaled descriptors would have equal potential to influence the QSAR models.

Initially, the pEC_{50} values pertaining to hPPARα and hPPARγ transactivation actions were correlated to pK_i values corresponding to PPARγ binding activity, and pEC_{50} values pertaining to hPPARα and hPPARγ transactivations for all active congeners to confer the diversity between the binding and transactivation activities, and hPPARα and hPPARγ transactivations. The derived correlations are given below:

pK_i (PPARγ) = -0.465 pEC_{50} (hPPARα) + 10.122

n = 13, r = 0.458, s = 0.932, F = 2.927 **(4.1)**

pK_i (PPARγ)= 1.051 pEC_{50} (hPPARγ) -1.313

n = 15, r = 0.484, s = 1.052, F = 3.970 (**4.2**)

pEC_{50} (hPPARγ) = 0.059 pEC_{50} (hPPARα) +7.362

n = 14, r = 0.138, s = 0.430, F = 0.232 **(4.3)**

where n, r, s and F represent respectively the number of data points, the multiple correlation coefficient, the standard deviation and the F-ratio between the variances of calculated and observed activities. All these equations have divulged not very much significant statistical parameter. No correlation between EC_{50} values obtained from transactivation PPARγ tests and K_i values from binding tests suggested that these derivatives may have a binding site different from the Rosiglitazone binding site. This ensures us that the biological actions in terms of binding and or transactivation are independent. Therefore, we have considered all types of biological endpoints as the dependent variables in the subsequent parametric analysis.

The PPARγ binding activity of titled compounds was investigated with 55 relevant 0D-, 1D- and 2D-descriptors. A training set consisting 11 compounds was considered for the development of QSAR models and test set involving 04 (nearly one-fourth of the total) compounds for the external validation of derived significant models. CP-MLR resulted one model in one parameter and ten models in two parameters having r^2_{Test}> 0.5. These models shared 12 descriptors and are listed in Table 4.3 along with their physical meaning, average regression coefficient and total incidences. The sign of the regression coefficients indicates the direction of influence of explanatory variables in above models. The positive regression coefficient associated to a descriptor will augment the activity profile of a compound while the negative coefficient will cause detrimental effect to it.

Table 4.3: Identified descriptors[a] along with their physical meaning, average regression coefficient and incidence[b], in modeling the binding and transactivation activity.

Descriptor; average regression coefficient and (incidence) in analysis for the:		
Binding activity	Transactivation activity	
PPARγ	hPPARα	hPPARγ
Constitutional descriptors (**CONST**):		
MW; -1.458(1)	AMW; -2.340(1)	Me; -0.774 (3)
AMW; 1.565(1),	Me; -1.254(1)	RBN; -1.380(1)
Me; 1.078(1)		
Topological descriptors (**TOPO**):		

MAXDP; -1.529(1)	MAXDP; -1.943(2)	HNar; 1.543 (3)
X2A; 1.826(1)	IC1; -2.188(2)	IVDE; -2.034(1)
T(N..O); -1.760(1)	T(N..N); -2.268(6)	IC2; -1.111(1)
		SIC4; 2.000(1)
2D autocorrelations (**2D-AUTO**):		
GATS4v; 1.478(1)	MATS7m; -1.546(1)	MATS5v; 2.294(3)
GATS2e; -1.508(1)	MATS8m; -1.431(1)	MATS8v; 1.333(1)
	GATS5v; -1.998(1)	MATS5e; -1.393(4)
	GATS6e; 1.326(1)	MATS8e; 1.243(6)
		GATS2e; 0.908(1)
		GATS6e; 1.316(5)
		GATS8e; -1.102(1)
Functional groups (**FUNC**):		
nCconjR; -1.022(1)		nCs; -0.631(1)
nROR; 1.006(1)		nCt; 0.827(1)
nHDon; 2.742(5)		
Atom-centered fragments (**ACF**):		
O-060; -2.424(6)		C-006; -1.035(3)
		C-008; 1.544(8)
Empirical descriptors (**EMP**):		
		Hy; 1.529(1)

[a]The descriptors are identified from the two parameter models for PPARγ binding activity and hPPARα transactivation activity profiles , emerged from CP-MLR protocol with filter-1 as 0.3, filter-2 as 2.0, filter-3 as 0.5 and filter-4 as $0.3 \leq q^2 \leq 1.0$ with a training set of 11 and 10 compounds, respectively; and for PPARγ transactivation activity profile three parameter models with filter-1 as 0.3, filter-2 as 2.0, filter-3 as 0.878 and filter-4 as $0.3 \leq q^2 \leq 1.0$ with a training set of 13 compounds. [b]The average regression coefficient of the descriptor corresponding to all models and the total number of its incidence. The arithmetic sign of the coefficient represents the actual sign of the regression coefficient in the models. **CONST:** MW, molecular weight; AMW; average molecular weight; Me, mean atomic Sanderson electronegativity (scaled on Carbon atom); RBN, number of rotatable bonds; **TOPO:** MAXDP maximal electrotopological positive variation; X2A, average connectivity index chi-2; T(N..O), sum of topological distances between N..O; T(N..N), sum of topological distances between N..N; HNar, Narumi harmonic index; IVDE, mean information content vertex degree equality; IC1, information content index (neighborhood symmetry of 1-order); IC2, information content index (neighborhood symmetry of 2-order); SIC4, structural information content (neighborhood symmetry of 4-order); **2D-AUTO:** MATS7m, Moran autocorrelation of lag-7/ weighted by atomic masses; MATS8m, Moran autocorrelation of lag-7/ weighted by atomic masses; MATS5v, Moran autocorrelation of lag-5/ weighted by atomic van der Waals volumes; MATS8v, Moran autocorrelation of lag-8/ weighted by atomic van der Waals volumes; MATS5e, Moran autocorrelation of lag-5/ weighted by atomic Sanderson electronegativities; MATS8e, Moran autocorrelation of lag-8/ weighted by atomic Sanderson electronegativities; GATS4v, Geary autocorrelation of lag-4/ weighted by atomic van der Waals volumes; GATS5v, Geary autocorrelation of lag-5/ weighted by atomic van der Waals volumes; GATS2e, Geary autocorrelation of lag-2/ weighted by atomic Sanderson electronegativities; GATS6e, Geary autocorrelation of lag-6/ weighted by atomic Sanderson electronegativities; GATS8e, Geary autocorrelation of lag-8/ weighted by atomic Sanderson electronegativities; **FUNC:** nCconjR; number of exo-conjugated C(sp2); nROR; number of aliphatic ethers; nHDon; number of donor atoms for H-bonds (with N and O); nCs; number of total secondary C(sp3); nCt; number of total tertiary C(sp3); **ACF:** O-060, Al-O-Ar/Ar-O-Ar/R..O..R/R-O-C=X; C-006, CH2RX; C-008, CHR2X; **EMP:** Hy, hydrophilic factor.

The selected highly significant two parameter models, emerged in CP-MLR for the PPARγ binding activity are given below.

pK_i = -1.529(0.437)MAXDP + 2.379(0.394)nHDon + 6.328

n = 11, r = 0.940, s = 0.455, F = 30.721, q^2_{LOO} = 0.811,

q^2_{L3O} = 0.781, r^2_{Test}= 0.791 **(4.4)**

pK_i = 1.006(0.295)nROR + 2.952(0.396)nHDon + 4.866

n = 11, r = 0.938, s = 0.462, F = 29.711, q^2_{LOO} = 0.764,

q^2_{L3O} = 0.784, r^2_{Test}= 0.662 **(4.5)**

pK_i = 1.826(0.534)X2A – 2.258(0.378)O-060 + 7.011

n = 11, r = 0.934, s = 0.478, F = 27.460, q^2_{LOO} = 0.813,

q^2_{L3O} = 0.810, r^2_{Test}= 0.603 **(4.6)**

pK_i = 1.565(0.479)AMW – 2.295(0.387)O-060 + 6.892

n = 11, r = 0.930, s = 0.491, F = 25.847, q^2_{LOO} = 0.740,

q^2_{L3O} = 0.706, r^2_{Test}= 0.725 **(4.7)**

The data within the parentheses are the standard errors associated with regression coefficients. The descriptors, participated in above models, are from constitutional (AMW), topological (MAXDP and X2A), functional group (nHDon and nROR) and atom-centered fragment (O-060) class. Constitutional class descriptors are molecular connectivity and conformations independent 0D descriptors. The emerged constitutional class descriptor AMW (average molecular weight) has shown positive correlation to activity favoring high average molecular weight of a molecule for elevated binding activity.

Topological class descriptors are based on a graph representation of the molecule and are numerical quantifiers of molecular topology obtained by the application of algebraic operators to matrices representing molecular graphs and whose values are independent of vertex numbering or labeling. They can be sensitive to one or more structural features of the molecules such as size, shape, symmetry, branching and cyclicity and can also encode chemical information concerning atom type and bond multiplicity. The negative contribution of descriptor MAXDP (maximal electrotopological positive

variation) and positive contribution of descriptor X2A (average connectivity index, chi-2) suggested that a lower value of descriptor MAXDP and a higher value of X2A would be supportive to the activity.

Descriptors nHDon and nROR are functional group class descriptors. Descriptor nHDon represents number of donor atoms for H-bonds (with N and O) and nROR corresponds to number of aliphatic ethers. Presence and or higher number of both the types of functionality in a molecular structure would be favorable to the binding activity.

Descriptor O-060 is representative of atom centered fragments (ACF) class. ACF class descriptors are based on the counting of 120 atom centered fragments, defined by Ghose-Crippen in a molecular structure. Descriptor O-060 represents Al-O-Ar/Ar-O-Ar/R..O..R/R-O-C=X type fragments in a molecular structure. The negative sign of correlation coefficient of this descriptor recommends absence of such types of fragments for elevated PPARγ binding profile. Based on the total number of incidences, it is also clear that descriptors O-060 and nHDon appeared as most relevant descriptors to explain the binding profiles of titled compound (Table 4.3).

In above equations (4.4) to (4.7), the F-values are significant at 99% level. Value greater than 0.5 of both the indices q^2_{LOO} and q^2_{L3O} showed internal robustness of the models whereas accountability of selected test-set for external validation reflected through the r^2_{Test} values (> 0.5). These models are able to estimate up to 88.36 percent of variance in observed activity of the compounds. The derived statistical parameters of these four models in two parameters have shown the statistically significance, therefore, these models were used to calculate the PPARγ binding activity profiles of all the compounds and are included in Table 4.4 for the sake of comparison with observed ones. A close agreement between them has been observed. Additionally, the graphical display, showing the variation of observed versus calculated activities is given in Figure 4.1 to ensure the goodness of fit for each of these four models.

Table 4.4: Observed, calculated and predicted PPARγ binding activities of Tetrahydroquinolines.

Cpd.	pK_i (M)[a]										
	Obs[b]	Eq. (4.4)		Eq. (4.5)		Eq. (4.6)		Eq. (4.7)		PLS	
		Cal[c]	Pre[c]	Cal[c]	Pre[c]	Cal[c]	Pre[c]	Cal[c]	Pre[c]	Cal[c]	Pre[c]
1	8.10	8.35	8.47	7.82	7.60	7.69	7.61	8.46	8.83	8.16	8.19
2	7.74	7.52	7.46	7.35	7.28	7.71	7.45	7.02	6.66	7.55	7.18
3	5.00	4.80	4.62	4.87	4.76	4.80	4.68	5.08	5.13	4.83	4.74
4	5.00	5.67	5.85	4.87	4.76	4.75	4.58	5.09	5.15	5.28	5.37
5	6.21	5.51	5.32	5.87	5.77	6.04	6.01	5.84	5.75	5.70	5.59
6	5.00	5.50	5.63	5.87	6.14	6.09	6.24	5.75	5.95	5.57	5.71
7[d]	5.00	5.50	-[d]	5.87	-[d]	6.09	-[d]	5.75	-[d]	5.57	-[d]
8[d]	7.32	7.34	-[d]	7.35	-[d]	7.27	-[d]	7.07	-[d]	7.13	-[d]
9[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
10	7.05	7.34	7.39	7.35	7.41	7.27	7.32	7.07	7.08	7.13	7.14
11[d]	7.74	7.34	-[d]	7.35	-[d]	7.27	-[d]	7.07	-[d]	7.13	-[d]
12	7.27	7.32	7.33	7.35	7.36	7.38	7.40	7.05	6.98	7.19	7.17
13	7.74	7.29	7.22	7.35	7.28	7.32	7.24	7.16	7.02	7.21	7.11
14[d]	6.60	7.31	-[d]	7.35	-[d]	7.17	-[d]	6.97	-[d]	7.15	-[d]
15[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
16	7.48	7.28	7.25	7.35	7.32	7.74	7.80	7.86	7.96	7.92	8.04
17	7.27	7.29	7.34	7.82	8.24	7.06	7.00	7.48	7.52	7.33	7.38
18[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]

[a]On molar basis, K_i represents the binding affinity to human PPARγ using a competitive binding assay with an appropriate radioligand [^{3}H]Rosiglitazone; [b]Taken from ref. [767, 768]; [c]Leave-one-out (LOO) procedure; [d]Compound included in test set and [e]Compound with uncertain activity, not part of data set.

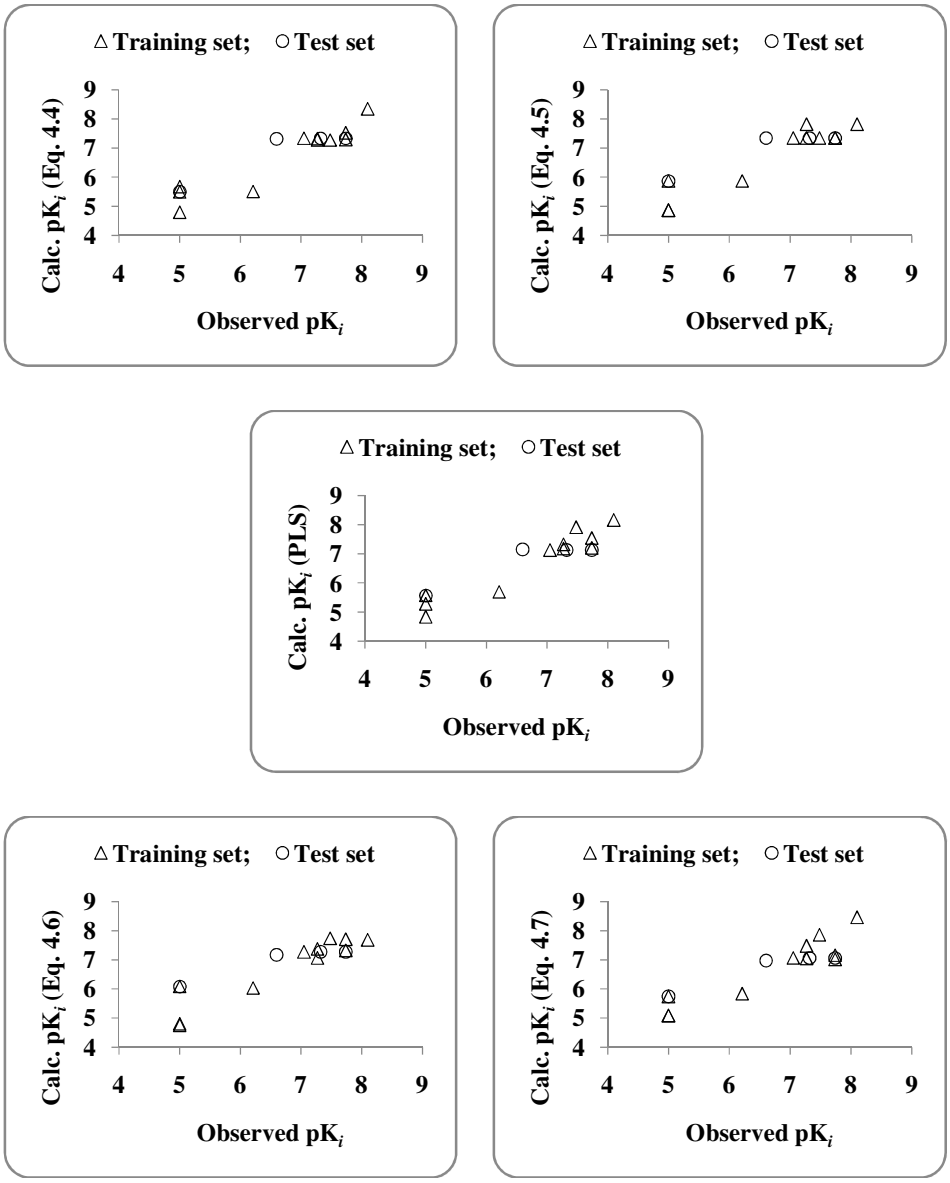


Figure 4.1: Plot of observed and calculated pK$_i$ values for training- and test-set compounds.

A PLS has also been carried out on 12 descriptors (identified through CP-MLR) to support the study. The results of PLS analysis are given in Table 4.5. For this purpose, the descriptors have been autoscaled (zero mean and unit s.d.) to give each one of them equal weight in the analysis. In the PLS cross-validation, two components have been found to be the optimum for these

12 descriptors and they explained 91.4 percent variance in the activity (r^2 = 0.914). The MLR-like PLS coefficients of these 12 descriptors are given in Table 4.5. The calculated activity values of training- and test-set compounds are in close agreement to that of the observed ones and are listed in Table 4.4. For the sake of comparison, the plot between observed and calculated activities (through PLS analysis) for the training- and test-set compounds is given in Figure 4.1.

Table 4.5: PLS and MLR-like PLS models from the descriptors of two parameter CP-MLR models for PPARγ binding affinity.

A: PLS equation			
PLS components		PLS coefficient (s.e.)[a]	
Component-1		0.483(0.057)	
Component-2		0.306(0.077)	
Constant		6.714	
B: MLR-like PLS equation			
S. No.	Descriptor	MLR-like coefficient (f.c.)[b]	Order
1	MW	0.025(0.017)	10
2	AMW	0.118(0.082)	4
3	Me	0.071(0.049)	7
4	MAXDP	-0.167(-0.116)	3
5	X2A	0.103(0.071)	5
6	T(N..O)	0.020(0.014)	11
7	GATS4v	-0.025(-0.018)	9
8	GATS2e	-0.087(-0.060)	6
9	nCconjR	-0.020(-0.014)	12
10	nROR	0.070(0.048)	8
11	nHDon	0.364(0.252)	2
12	O-060	-0.376(-0.260)	1
	Constant	6.529	
C: PLS regression statistics		Values	
n		11	
r		0.956	
s		0.392	
F		42.800	
q^2_{LOO}		0.850	
q^2_{L3O}		0.887	
r^2_{Test}		0.762	

[a]Regression coefficient of PLS factor and its standard error. [b]Coefficients of MLR-like PLS equation in terms of descriptors for their original values; f.c. is fraction contribution of regression coefficient, computed from the normalized regression coefficients obtained from the autoscaled (zero mean and unit s.d.) data.

Figure 4.2 shows a plot of the fraction contribution of normalized regression coefficients of these descriptors to the activity (Table 4.5).

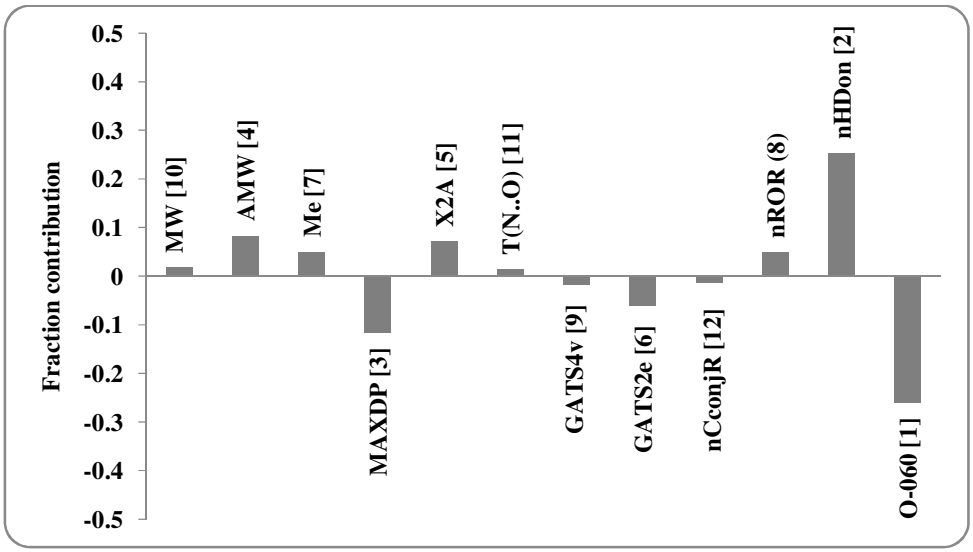


Figure 4.2: Plot of fraction contribution of MLR-like PLS coefficients (normalized) against 12 identified descriptors (Table 4.5) associated with PPARγ binding affinity of the compounds.

Descriptors in decreasing order of significance in PLS analysis are O-060, nHDon, MAXDP, AMW, X2A, GATS2e, Me, nROR, GATS4v, MW, T(N..O) and nCconjR. Among these descriptors, O-060, nHDon, MAXDP, AMW, X2A and nRORare part of Equations discussed above and convey same inferences in PLS analysis. The positive contributions of constitutional class descriptors MW (molecular weight) and Me (mean atomic Sanderson electronegativity scaled on Carbon atom); and topological class descriptor T(N..O), representing the sum of topological distances between N and O atoms advocated that higher values of these are helpful in improving the activity profile. Whereas lower values of descriptors GATS4v (Geary autocorrelation of lag-4/ weighted by atomic van der Waals volumes), GATS2e (Geary autocorrelation of lag-2/ weighted by atomic Sanderson electronegativities) and number of exo-conjugated C(sp2) (descriptornCconjR) would be supportive to enhance the activity. It is also observed that PLS

model from the dataset devoid of 12 descriptors (Table 4.5) remained inferior in explaining the activity of the analogues.

QSAR rationales, with the same test-set used earlier for the analysis of PPARγ binding activity, have also been obtained for other reported activity profile pertaining to hPPARα and hPPARγ transactivation. A descriptor pool of 39 and 67 relevant descriptors for hPPARα and hPPARγ transactivation, respectively, were subjected to CP-MLR analysis. CP-MLR resulted a total number of 08 models in two parameters sharing 9 descriptors for hPPARα activity. For the hPPARγ activity 15 three parameters models sharing 18 descriptors were obtained. The shared descriptors along with their physical meaning, average regression coefficient and total incidences for both the analysis have been given in Table 4.3. The selected models emerged through CP-MLR are mentioned below.

pEC_{50}(hPPARα)= – 1.686(0.388)IC1 – 1.933(0.396)T(N..N) + 7.965
n = 10, r = 0.945, s = 0.355, F = 29.412, q^2_{LOO} = 0.837,
q^2_{L3O} = 0.847, r^2_{Test}= 0.610 (**4.8**)

pEC_{50}(hPPARα)= – 2.013(0.469)MAXDP – 2.690(0.426)IC1 + 8.450
n = 10, r = 0.933, s = 0.391, F = 23.596, q^2_{LOO} = 0.740,
q^2_{L3O} = 0.707, r^2_{Test}= 0.749 (**4.9**)

pEC_{50}(hPPARα)= – 2.340(0.422)AMW – 1.874(0.517)MAXDP + 7.881
n = 10, r = 0.916, s = 0.435, F = 18.364, q^2_{LOO} = 0.643,
q^2_{L3O} = 0.571, r^2_{Test}= 0.765 **(4.10)**

pEC_{50}(hPPARα)= –1.976(0.518)T(N..N) –1.546(0.540)MATS7m +7.995
n = 10, r = 0.905, s = 0.462, F = 15.876, q^2_{LOO} = 0.724,
q^2_{L3O} = 0.700, r^2_{Test}= 0.617 **(4.11)**

Newly appeared descriptors IC1 and T(N..N) are topological class descriptors whereas descriptor MATS7m belong to 2D-autocorrelations (2D-AUTO) class. The 2D-AUTO descriptors, ATSke, GATSke and MATSke have their origin in autocorrelation of topological structure of Broto-Moreau,

of Moran and of Geary, respectively. The computation of these descriptors involves the summation of different autocorrelation functions corresponding to the different fragment lengths and lead to different autocorrelation vectors corresponding to the lengths of the structural fragments. Also a weighting component in terms of a physicochemical property has been embedded in these descriptors. As a result, these descriptors address the topology of the structure or parts thereof in association with a selected physicochemical property. In these descriptors' nomenclature, the penultimate character, a number, indicates the number of consecutively connected edges considered in its computation and is called as the autocorrelation vector of lag k (corresponding to the number of edges in the unit fragment). The very last character of the descriptor's nomenclature indicates the physicochemical property considered in the weighting component – m for atomic mass, e for atomic Sanderson electronegativity and p for atomic polarizability - for its computation.

All the descriptors, participated in Eqs. (4.8) to (4.11), have shown negative correlation to activity as evinced from the signs of the correlation coefficients thus lower values of information content index of 1^{st} order neighborhood symmetry (descriptor IC1), sum of topological distances between N..N (descriptor T(N..N)), maximal electrotopological positive variation (descriptor MAXDP), average molecular weight (descriptor AMW) and Moran autocorrelation of lag-7/ weighted by atomic masses (descriptor MATS7m) would be beneficiary to the hPPARα activity.

The derived statistical parameters models have revealed that these models are statistically significant. The values greater than 0.5 of indices q^2_{LOO} and q^2_{L3O} have accounted the internal robustness of models and the r^2_{Test} values greater than 0.5 are accountable for external validation. These models are able to estimate up to 89.36 percent of variance in observed activity of the compounds. These models were, therefore, used to calculate the activity profiles of all the compounds and are included in Table 4.6 for the sake of comparison with observed ones. A close agreement between them has been observed.

Considering the number of observation in the data set for the hPPARγ transactivation profile, models with up to three descriptors were explored. Following are the selected three-descriptor models, obtained from CP-MLR, for the hPPARγ transactivation.

Table 4.6: Observed and calculated transactivation activities of tetrahydroquinoline analogues.

Cpd.	Transactivation $pEC_{50}(M)$[a]									
	hPPARα					hPPARγ				
		Calculated					Calculated			
	Obs.[b]	Eq. (4.8)	Eq. (4.9)	Eq. (4.10)	Eq. (4.11)	Obs.[b]	Eq. (4.12)	Eq. (4.13)	Eq. (4.14)	Eq. (4.15)
1	5.00	5.01	5.58	5.10	5.04	8.40	8.31	8.62	8.14	8.48
2	6.38	6.71	6.44	5.98	6.91	7.43	7.34	7.57	7.71	7.23
3	-[c]	7.81	6.19	5.28	7.76	6.82	7.35	6.75	6.87	7.17
4	-[c]	6.73	6.19	6.33	7.32	6.71	6.56	6.44	6.55	6.74
5	6.72	7.19	6.71	6.74	7.16	7.32	7.03	7.51	7.70	7.02
6	7.52	7.60	7.35	6.86	7.01	7.72	7.44	7.58	7.76	7.32
7[d]	7.54	7.60	7.35	6.86	7.01	7.85	7.44	7.58	7.76	7.32
8[d]	7.92	6.91	7.11	7.39	6.70	7.96	7.45	7.49	7.69	7.54
9	7.54	6.91	7.11	7.39	6.70	7.85	7.45	7.49	7.69	7.54
10	6.71	6.91	7.11	7.39	6.70	7.19	7.45	7.49	7.69	7.54
11[d]	6.94	6.91	7.11	7.39	6.70	8.11	7.45	7.49	7.69	7.54
12	7.14	6.87	7.03	7.41	7.32	7.89	7.90	7.52	7.69	7.74
13	7.47	7.30	7.67	7.20	7.47	8.15	8.34	7.85	7.66	8.52
14[d]	8.05	7.04	7.28	7.51	7.63	8.12	7.98	7.85	7.68	7.88
15	-[c]	6.51	6.48	6.93	6.54	-[c]	7.10	7.23	7.59	7.43
16	6.00	5.92	5.44	6.14	6.08	6.90	7.23	6.87	6.69	6.71
17	5.00	5.07	5.04	5.26	5.09	7.85	7.72	8.19	7.96	7.90
18	-[c]	4.84	5.34	6.19	5.22	5.00	5.09	5.34	5.11	5.33

[a]On molar basis, determined in a transient transfection assay using pGAL4hPPARα and pGAL4hPPARγ; [b]Taken from ref. [767, 768]; [c]Inactive compound, not part of data set and [d]Compound included in test set.

pEC_{50}(hPPARγ) = 2.386(0.312)MATS5v + 1.369(0.237)MATS8e
+ 0.827(0.241)nCt + 4.807
n = 13, r = 0.950, s = 0.311, F = 28.279, q^2_{LOO} = 0.640,
q^2_{L3O} = 0.713, r^2_{Test}= 0.545 **(4.12)**

pEC_{50}(hPPARγ) = 1.089(0.346)HNar + 1.129(0.232)MATS8e
+ 1.363(0.213)C-008 + 5.156

n = 13, r = 0.950, s = 0.313, F = 27.929, q^2_{LOO} = 0.736,
q^2_{L3O} = 0.760, r^2_{Test}= 0.613 **(4.13)**

pEC_{50}(hPPARγ) = – 0.710(0.242)Me + 1.090(0.240)MATS8e
+ 1.595(0.215)C-008 + 5.808
n = 13, r = 0.946, s = 0.325, F = 25.748, q^2_{LOO} = 0.785,
q^2_{L3O} = 0.701, r^2_{Test}= 0.766 **(4.14)**

pEC_{50}(hPPARγ)= 2.482(0.329)MATS5v + 1.390(0.308)GATS6e
– 0.983(0.231)C-006 + 5.707
n = 13, r = 0.946, s = 0.325, F = 25.723, q^2_{LOO} = 0.709,
q^2_{L3O} = 0.808, r^2_{Test}= 0.560 **(4.15)**

In all above equations (4.12) to (4.15) the F-values remained significant at 99% level. The values, greater than 0.5, obtained for the indices q^2_{LOO}, q^2_{L3O}, and r^2_{Test} ascertained the internal robustness and external validation of the models. These models are capable to explain up to 90.40 percent of variance in observed activity of the compounds. The derived statistical parameters are in tune to statistical significance. The activity profile of all the compounds calculated using these equations is in the close agreement to the observed ones and the same are included in Table 4.6.

2D-autocorrelations class descriptors MATS5v (Moran autocorrelation of lag-5/ weighted by atomic van der Waals volumes), MATS8e (Moran autocorrelation of lag-8/ weighted by atomic Sanderson electronegativities) and GATS6e (Geary autocorrelation of lag-6/ weighted by atomic Sanderson electronegativities) added positively to the inhibitory activity suggesting that a higher values of descriptors MATS5v, MATS8e and GATS6e would be helpful to augment the activity. Constitutional class descriptors Me (mean atomic Sanderson electronegativity scaled on Carbon atom) favors low value of mean atomic Sanderson electronegativity for elevated activity.

Descriptor HNar, corresponds to Narumi harmonic index, is a topological class descriptor. The positive contribution of descriptor HNar suggested that a higher value of it would be supportive to the activity. The

other participated descriptors are nCt (from the functional group class), and C-006 and C-008 (from the atom-centered fragments). Number of total tertiary C(sp3) (descriptor nCt) and CHR2X type atom centered fragment (descriptor C-008) correlated positively to the activity suggested that a higher value of these will augment the activity. On the other hand negative correlation of descriptor C-006 advocated that CH2RX type structural fragments would be detrimental to the activity.

2.1.1.2. APPLICABILITY DOMAIN (AD)

To analyze the applicability domain (AD) a Williams plot of the model based on the whole data set (Table 4.7) has been constructed that is shown in Figure 4.3.

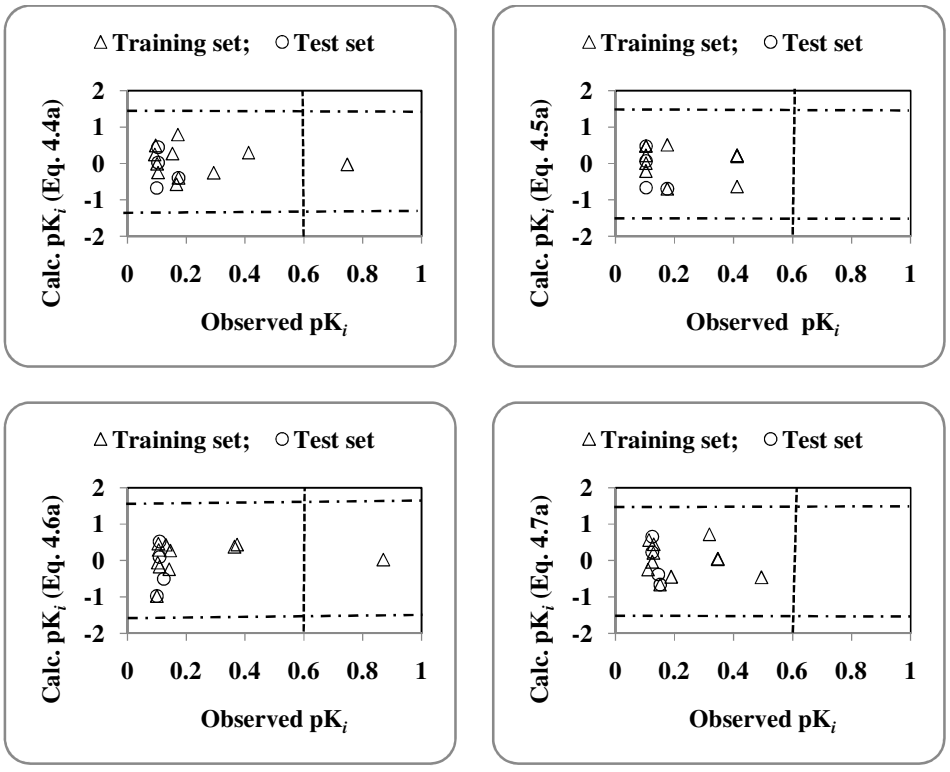


Figure 4.3: Williams plot for the training-set and test- set for binding affinity of PPARγ for the compounds in Table 4.1. The horizontal dotted line refers to the residual limit (±3×standard deviation) and the vertical dotted line represents threshold leverage h* (= 0.6).

Table 4.7: Models derived for the whole data set (n = 15) for the PPARγ binding affinity in descriptors identified through CP-MLR.

Model	r	s	F	q^2_{LOO}	Eq.
pK_i= -1.523(0.412)MAXDP +2.483(0.378)nHDon + 6.229	0.930	0.457	38.897	0.808	(4.4a)
pK_i= 0.915(0.296)nROR + 3.123(0.401)nHDon + 4.780	0.916	0.499	31.601	0.739	(4.5a)
pK_i= 1.962(0.579)X2A – 2.374(0.386)O-060 + 6.938	0.904	0.534	26.829	0.745	(4.6a)
pK_i= 1.668(0.441)AMW – 2.457(0.363)O-060 + 6.894	0.914	0.505	30.802	0.727	(4.7a)

The analysis revealed that none of the compound has been identified as an obvious 'outlier' for the PPARγ binding activity if the limit of normal values for the Y outliers (response outliers) was set as 3×(standard deviation) units and compounds **2** and **17** appeared as chemically influential compounds. Furthermore, the suggested model matches the high quality parameters with good fitting power and the capability of assessing external data and all of the compounds were within the applicability domain of the proposed model and were evaluated correctly.

2.1.2. CONCLUSIONS

This study has provided a rational approach for the development of tetrahydroquinoline derivatives as PPARα/γ agonists. The descriptors identified in CP-MLR analysis for the PPARγ binding activity have highlighted the role of average molecular weight (AMW), maximal electrotopological positive variation (MAXDP), average connectivity index i.e. chi-2 (X2A) to explain the binding actions in addition to presence of donor atoms for H-bonds with N and O (nHDon), aliphatic ethers (nROR) and absence of Al-O-Ar/Ar-O-Ar/R..O..R/R-O-C=X type fragments in a molecular structure (O-060) have also shown prevalence to optimize the PPARγ binding activity of titled compounds. PLS analysis has further confirmed the dominance of the CP-MLR identified descriptors and applicability domain analysis revealed that the suggested model for PPARγ binding activity matches the high quality parameters with good fitting power and the capability

of assessing external data and all of the compounds was within the applicability domain of the proposed model and were evaluated correctly.

Derived statistical significant models for hPPARα transactivation activity revealed that lower values of information content index of 1st order neighborhood symmetry (descriptor IC1), sum of topological distances between N..N (descriptor T(N..N)), maximal electrotopological positive variation (descriptor MAXDP), average molecular weight (descriptor AMW) and Moran autocorrelation of lag-7/ weighted by atomic masses (descriptor MATS7m) would be beneficiary to the hPPARα activity. Role of atomic van der Waals volumes and electronegativities to explain the hPPARγ transactivation activity is evinced through participation of descriptors MATS5v, MATS8e, GATS6e and Me. Additionally a higher value of Narumi harmonic index (HNar), number of total tertiary C(sp3) (descriptor nCt), presence of CHR2X type atom centered fragment (descriptor C-008) and absence of CH2RX type structural fragments (descriptor C-006) will augment the hPPARγ transactivation activity.

2.2. BENZYLPYRAZOLE ACYLSULFONAMIDES AS PPARγ AGONISTS

In present scenario development of new and safer antidiabetic agents which may lower hemoglobin A_{1c} (HbA_{1c}) levels and improve the lipid profile of patients simultaneously is ardently needed [770-773].PPARγ is expressed predominantly in adipose tissue, in a lesser extent in the intestine, mammary gland, endothelium, liver, skeletal muscle and in other tissues throughout the body. PPARγ plays a pivotal role in many physiological processes such as adipogenesis, glucose and lipid homeostasis, insulin sensitivity, inhibition of inflammatory responses, cell proliferation and promotion of terminal differentiation [752, 774, 775]. Introduction of troglitazone, pioglitazone hydrochloride and rosiglitazone maleate (the representatives of thiazolidinediones (TZDs)) as insulin sensitizers and the fact that TZDs are high-affinity PPARγ ligands [757] has opened new avenues for extensive

research in the area of antidiabetic drug discovery and development [754, 755, 776].

Efforts were made to indentify novel classes of PPAR ligands, based on several approaches such as PPARα/γ dual agonists, PPARγ/δ dual agonists and PPARα/γ/δ pan agonists, as second-generation insulin sensitizers [776]. Numerous reported non-TZD PPARγ ligands belonging to different chemical classes are mostly carboxylic acids. A novel class of benzylpyrazole acylsulfonamides as non-thiazolidinedione (TZD), non-carboxylic-acid-based selective PPARγ agonists has been reported by Rikimaru et al. [777].

The reported twentyeight benzylpyrazole acylsulfonamides are considered as the data set for this study [777]. The general structure of these derivatives is given in Figure 4.4.

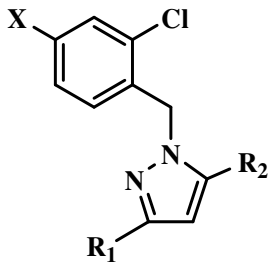
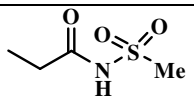
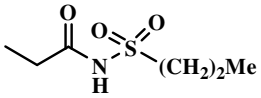
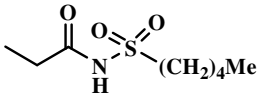
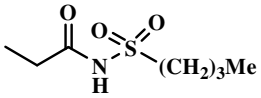
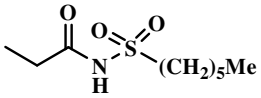
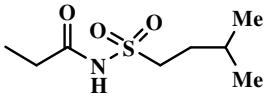
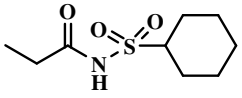
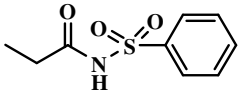
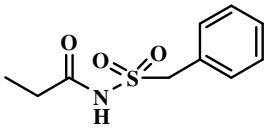
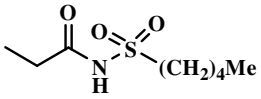
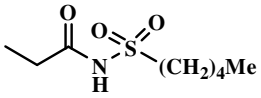
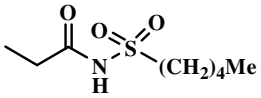
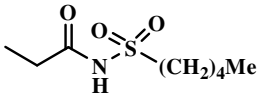
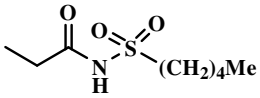
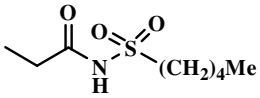
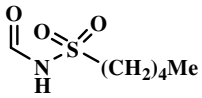
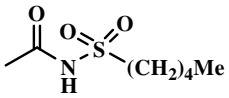
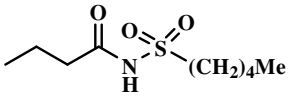
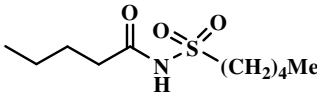
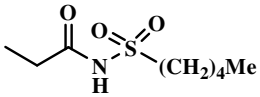
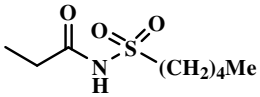
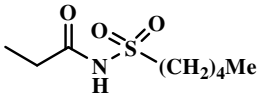
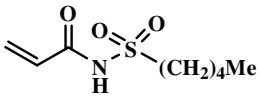
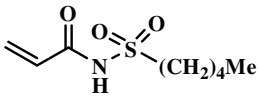
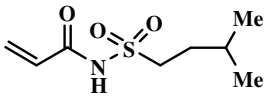


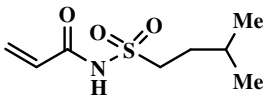
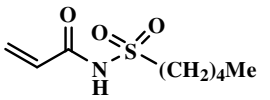
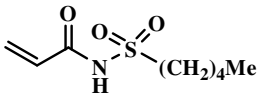
Figure 4.4: General structure of the benzylpyrazole acylsulfonamides.

The structural variations of these analogues are mentioned in Table 4.8. These derivatives were evaluated for their transactivation activity against human PPARγ stably expressed in Chinese hamster ovary (CHO) cells. Transactivation activities were assessed by a luciferase reporter gene assay using (*R*)-5-(3-{4-[(2-Furan-2-yl-5-methyl-1,3-oxazol-4-yl)methoxy]-3-methoxyphenyl}propyl)-1,3-oxazolidine-2,4-dione as the reference PPARγ agonist [778] and were reported as EC_{50} and the same are also presented in Table 4.8 as pEC_{50} on molar basis. For modeling purpose the data set has been sub-divided into training set (for model development) and test set (for external prediction or validation). The selection of test set compounds was made using an in-house written randomization program. The test and training set compounds are also mentioned in Table 4.8.

Table 4.8: Structural variations and reported PPARγ transactivation activities of benzylpyrazole acylsulfonamides.

Cpd.	X	R_1	R_2	$pEC_{50}(M)$[a]
1	Cl	Isopropoxy		6.35
2[b]	Cl	Isopropoxy		6.80
3	Cl	Isopropoxy		7.80
4	Cl	Isopropoxy		7.40
5[b]	Cl	Isopropoxy		7.68
6	Cl	Isopropoxy		7.77
7	Cl	Isopropoxy		6.82
8	Cl	Isopropoxy		6.96
9	Cl	Isopropoxy		6.59
10	Cl	Butoxy		8.12
11[b]	Cl	$MeO(CH_2)_2O$		7.51
12[b]	Cl	Benzyloxy		7.85

13	Cl	2-Pyridylmethoxy		6.92
14	Cl	Isopropyl		7.89
15[b]	Cl	phenyl		7.96
16	Cl	Butoxy		-[c]
17	Cl	Butoxy		-[d]
18[b]	Cl	Butoxy		7.17
19	Cl	Butoxy		7.54
20	CF_3	Isopropoxy		8.00
21[b]	CF_3	Isopropyl		8.08
22	CF_3	Cyclopropyl		7.89
23	Cl	Isopropoxy		7.70
24	CF_3	Isopropoxy		8.03
25	CF_3	Isopropoxy		8.08

26	CF_3	$MeO(CH_2)_2O$		8.05
27	Cl	Isopropyl		8.09
28	CF_3	Cyclopropyl		7.92

[a]EC_{50} (the effective concentration for 50% response of a given compound's intrinsic maximum response) on molar basis, taken from reference [777]; [b]Compound included in test set; [c]Inactive compound, not part of data set; [d]Compound with uncertain activity, not part of data set.

2.2.1. RESULTS AND DISCUSSION

2.2.1.1. QSAR RESULTS

A total number of 484 descriptors, belonging to 0D- to 2D- modules of DRAGON software, have been computed to obtain most appropriate models describing the biological activity. Prior to model development procedure, all those descriptors that are inter-correlated beyond 0.90 and showing a correlation of less than 0.1 with the biological endpoints (descriptor versus activity, $r < 0.1$) were excluded. This procedure has reduced the total descriptors from 484 to 107 as relevant ones to explain the biological actions of titled compounds. For the purpose of modeling study, 7 compounds have been included in the test set for the validation of the models derived from 19 training set compounds. All the 107 significant descriptors have been subjected to CP-MLR analysis with default "filters" set in it. Statistical models in two descriptors have been derived to achieve the best relationship correlating PPARγ transactivation activity. A total number of seven models in two descriptors, having $r^2_{Test} > 0.5$, were obtained through CP-MLR. The selected models in two descriptors are given below.

$pEC_{50} = 6.337 + 1.221(0.229)BELm5 + 1.017(0.225)JGI4$

$n = 19, r = 0.871, s = 0.291, F = 25.371, Q^2_{LOO} = 0.658, Q^2_{L5O} = 0.663$

$r^2_{Test} = 0.549$, FIT = 2.206, LOF = 0.114, AIC = 0.116 **(4.16)**

$pEC_{50} = 6.319 + 1.342(0.254)BELm5 + 0.810(0.218)JGI2$

$n = 19, r = 0.841, s = 0.321, F = 19.337, Q^2_{LOO} = 0.594, Q^2_{L5O} = 0.594$

$r^2_{Test} = 0.503$, FIT = 1.681, LOF = 0.139, AIC = 0.142 **(4.17)**

$pEC_{50} = 6.387 + 1.080(0.265)BELm5 + 1.055(0.300)GGI4$

$n = 19, r = 0.831, s = 0.330, F = 17.991, Q^2_{LOO} = 0.597, Q^2_{L5O} = 0.583$

$r^2_{Test} = 0.680$, FIT = 1.564, LOF = 0.147, AIC = 0.149 **(4.18)**

$pEC_{50} = 6.405 + 1.189(0.276)BELm5 + 0.836(0.274)GGI2$

$n = 19, r = 0.809, s = 0.349, F = 15.186, Q^2_{LOO} = 0.538, Q^2_{L5O} = 0.501$

$r^2_{Test} = 0.597$, FIT = 1.320, LOF = 0.165, AIC = 0.167 **(4.19)**

Most of the descriptors GGI2, GGI4, JGI2 and JGI4 participated in above models are from the GALVEZ class and the remained one BELm5 is the modified Burden eigenvalue (BCUT class descriptor). All the descriptors have shown positive influence on the activity as evident from the signs of regression coefficients. Thus a higher value of Galvez descriptors GGI2 (2^{nd} order topological charge index), GGI4 (4^{th} order topological charge index), JGI2 (2^{nd} order mean topological charge index) and JGI4 (4^{th} order mean topological charge index) in addition to a higher value of the lowest eigenvalue n.5 of Burden matrix/weighted by atomic masses (descriptor BELm5) would be beneficiary to the activity.

The two descriptor models could estimate nearly 76% in observed activity of the compounds. Considering the number of observation in the dataset, models with up to three descriptors were explored. It has resulted in 21 three-parameter models with test set $r^2 > 0.50$. These models (with 107 descriptors) were identified in CP-MLR by successively incrementing the filter-3 with increasing number of descriptors (per equation). For this, the optimum *r*-bar value of the preceding level model (= 0.854) has been used as the new threshold of filter-3 for the next generation.

These models have shared 26 descriptors among them. All these 26 descriptors along with their brief meaning, average regression coefficients, and total incidence are listed in Table 4.9, which will serve as a measure of their estimate across these models.

Table 4.9: Identified descriptors[a] along with their class, average regression coefficient and incidence[b], in modeling the PPARγ transactivation activities of benzylpyrazole acylsulfonamides.

Descriptor class, average regression coefficient and (**incidence)**	
Topological descriptors (**TOPO**)	HNar, -0.692(**1**); MAXDP, 0.941(**2**); BAC, 1.180(**2**); Lop, 1.714(**10**); Uindex, 1.527(**1**); BIC3, 0.678(**1**); T(N..O), -0.778(**3**)
Modified Burden Eigen values (**BCUT**)	BELm5, 1.069(**6**); BEHv2, 0.986(**6**); BELv8, 1.368(**1**); BEHm3, 0.659(**1**)
Galvez Topological charge indices (**GLVZ**)	GGI2, 0.621(**1**); GGI4, 0.885(**4**); GGI7, 0.659(**1**); JGI2, 0.517(**1**); JGI4, 0.825(**4**); JGT, 0.583(**1**)
2D autocorrelations (**2D-AUTO**):	MATS8m, -0.765(**2**); MATS4v, 0.890(**6**); MATS3e, 1.538(**2**); MATS3p, -0.690(**1**); MATS5p, -1.079(**1**); GATS5p, 0.603(**2**)
Empirical descriptors (**EMP**)	Hy, -7.354(**1**)
Functional groups (**FUNC**)	nCrH2, -0.856(1)
Properties (**PROP)**	MLOGP, 0.523(**1**)

[a]The descriptors are identified from the three parameter models for PPARγ binding activity transactivation activity emerged from CP-MLR protocol with filter-1 as 0.3, filter-2 as 2.0, filter-3 as 0.854 and filter-4 as $0.3 \leq q^2 \leq 1.0$ with a training set of 19 compounds. [b]The average regression coefficient of the descriptor corresponding to all models and the total number of its incidence. The arithmetic sign of the coefficient represents the actual sign of the regression coefficient in the models. **TOPO:** HNar, Narumi harmonic index; MAXDP, maximal electrotopological positive variation; BAC, Balaban centric index; Lop, Lopping centric index; Uindex, Balaban U index; BIC3, bond information content of 3rd order neighborhood symmetry; T(N..O), sum of topological distances between N..O; **BCUT:** BEHm3, highest eigenvalue n.3 of Burden matrix/weighted by atomic masses; BELm5, lowest eigenvalue n.5 of Burden matrix/weighted by atomic masses; BEHv2, highest eigenvalue n.2 of Burden matrix/weighted by van der Waals volumes; BELv8, lowest eigenvalue n.8 of Burden matrix/weighted by van der Waals volumes; **GLVZ:** GGI2, topological charge index of order 2; GGI4, topological charge index of order 4; GGI7, topological charge index of order 7; JGI2, mean topological charge index of order 2; JGI4, mean topological charge index of order 4; JGT, global topological charge index; **2D-AUTO:** MATS8m, Moran autocorrelation of lag-8/ weighted by atomic masses; MATS4v, Moran autocorrelation of lag-4/ weighted by atomic van der Waals volumes; MATS3e, Moran autocorrelation of lag-3/ weighted by atomic Sanderson electronegativities; MATS3p, Moran autocorrelation of lag-3/weighted by atomic polarizabilities; MATS5p, Moran autocorrelation of lag-5/ weighted by atomic polarizabilities; GATS5p, Geary autocorrelation of lag-5/

weighted by atomic polarizabilities; **FUNC:** nCrH2, number of ring secondary C(sp3); **EMP:** Hy, hydrophilic factor; **PROP**: MLOGP, Moriguchi octanol-water partition coefficient (logP).

Following are the selected three-descriptor models for the PPARγ transactivation activities of benzylpyrazole acylsulfonamides emerged through CP-MLR.

pEC_{50} = 5.456 + 1.120(0.297)MAXDP + 1.679(0.208)Lop
+ 0.807(0.169)JGI4
n = 19, r = 0.936, s = 0.215, F = 35.559, Q^2_{LOO} = 0.718, Q^2_{L5O} = 0.606
r^2_{Test} = 0.523, FIT = 3.809, LOF = 0.078, AIC = 0.071 **(4.20)**

pEC_{50} = 6.380 + 1.267(0.185)Lop – 0.583(0.222)T(N..O)
+ 1.022(0.208)GGI4
n = 19, r = 0.927, s = 0.229, F = 30.866, Q^2_{LOO} = 0.760, Q^2_{L5O} = 0.733
r^2_{Test} = 0.512, FIT = 3.307, LOF = 0.088, AIC = 0.080 **(4.21)**

pEC_{50} = 5.986 + 1.243(0.193)Lop + 0.894(0.228)GGI4
+ 0.523(0.235)MLOGP
n = 19, r = 0.920, s = 0.240, F = 27.678, Q^2_{LOO} = 0.704, Q^2_{L5O} = 0.631
r^2_{Test} = 0.532, FIT = 2.965, LOF = 0.097, AIC = 0.088 **(4.22)**

pEC_{50} = 5.683 + 0.761(0.349)MAXDP + 1.560(0.241)Lop
+ 0.892(0.230)GGI4
n = 19, r = 0.919, s = 0.241, F = 27.382, Q^2_{LOO} = 0.685, Q^2_{L5O} = 0.773
r^2_{Test} = 0.556, FIT = 2.933, LOF = 0.098, AIC = 0.089 **(4.23)**

The newly appeared descriptors in above models, MAXDP, Lop and T(N..O), are topological class descriptors whereas MLOGP belongs to properties class. Descriptors MAXDP, Lop and MLOGP have shown positive and descriptor T(N..O), showed negative correlation to the activity. The signs of regression coefficients advocated that higher values of maximal electrotopological positive variation (descriptor MAXDP), Lopping centric index (descriptor Lop) and Moriguchi octanol-water partition coefficient i.e. logP (descriptor MLOGP) would be incremental to the activity. On the other

hand a higher value of sum of topological distances between N..O would be deleterious to the activity.

These models have accounted for nearly 88% variance in the observed activities. In the randomization study (100 simulations per model), none of the identified models has shown any chance correlation. The values greater than 0.5 of Q^2 index is in accordance to a reasonable robust QSAR model. The pEC_{50} values of training set compounds calculated using Eqs. (4.20) to (4.23) and predicted from LOO procedure have been included in Table 4.10. The models (4.20) to (4.23) are validated with an external test set of 7 compounds listed in Table 4.8. The predictions of the test set compounds based on external validation are found to be satisfactory as reflected in the test set r^2 (r^2_{Test}) values and the same is reported in Table 4.10. The goodness of fit between observed and calculated activities is shown in Figure 4.5.

Table 4.10: Observed and modeled PPARγ transactivation activity of benzylpyrazole acylsulfonamides.

S. No.	pEC_{50}(M)[a]										
		Eq. (4.20)		Eq. (4.21)		Eq. (4.22)		Eq. (4.23)		PLS	
	Obsd[b]	Calc	Pred[c]	Calc	Pred[c]	Calc	Pred[c]	Calc	Pred[c]	Calc	Pred[c]
1	6.35	6.48	7.08	6.89	7.09	6.61	6.86	6.46	7.05	6.57	6.64
2[d]	6.80	6.91	-[d]	7.09	-[d]	6.87	-[d]	6.84	-[d]	6.85	-[d]
3	7.80	7.70	7.69	7.57	7.53	7.54	7.50	7.65	7.62	7.58	7.55
4	7.40	7.57	7.58	7.48	7.49	7.38	7.37	7.46	7.47	7.24	7.22
5[d]	7.68	7.81	-[d]	7.65	-[d]	7.69	-[d]	7.81	-[d]	7.84	-[d]
6	7.77	7.93	7.95	7.75	7.75	7.69	7.68	7.78	7.78	7.62	7.61
7	6.82	6.86	6.89	6.67	6.60	6.72	6.67	6.81	6.80	6.87	6.88
8	6.96	6.63	6.49	6.67	6.54	6.73	6.63	6.65	6.52	6.69	6.60
9	6.59	6.64	6.67	6.73	6.80	6.77	6.85	6.72	6.78	6.66	6.68
10	8.12	7.79	7.68	7.66	7.58	7.70	7.61	7.80	7.69	7.73	7.62
11[d]	7.51	7.76	-[d]	7.40	-[d]	7.38	-[d]	7.78	-[d]	7.46	-[d]
12[d]	7.85	7.43	-[d]	7.46	-[d]	7.61	-[d]	7.59	-[d]	7.57	-[d]
13	6.92	7.42	7.53	7.00	7.11	7.29	7.35	7.58	7.68	7.33	7.38
14	7.89	8.12	8.17	8.00	8.01	7.92	7.92	7.89	7.89	7.95	7.96
15[d]	7.96	7.62	-[d]	7.64	-[d]	7.71	-[d]	7.65	-[d]	7.68	-[d]
16[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
17[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
18[d]	7.17	7.54	-[d]	7.56	-[d]	7.76	-[d]	7.69	-[d]	7.74	-[d]
19	7.54	7.57	7.59	7.64	7.68	7.99	8.25	7.80	7.86	7.83	7.97
20	8.00	7.95	7.94	7.98	7.98	8.01	8.01	8.01	8.01	8.03	8.04

21[d]	8.08	8.37	-[d]	8.42	-[d]	8.39	-[d]	8.25	-[d]	8.38	-[d]
22	7.89	7.93	7.94	8.04	8.07	8.05	8.09	7.91	7.91	7.87	7.87
23	7.70	7.57	7.55	7.57	7.55	7.50	7.47	7.56	7.53	7.63	7.63
24	8.03	7.81	7.80	7.98	7.98	7.97	7.96	7.92	7.90	8.09	8.09
25	8.08	8.00	7.99	8.17	8.20	8.12	8.13	8.05	8.04	8.17	8.21
26	8.05	8.08	8.08	8.01	7.99	7.97	7.87	8.20	8.24	8.05	8.05
27	8.02	7.99	7.98	8.00	8.00	7.88	7.87	7.80	7.78	8.02	8.02
28	7.92	7.80	7.78	8.04	8.06	8.02	8.03	7.81	7.80	7.92	7.92

[a]On molar basis; [b]Taken from ref. [777]; [c]Leave-one-out (LOO) procedure; [d]Compound included in test set; [e]Compound with uncertain activity or inactive, not part of data set.

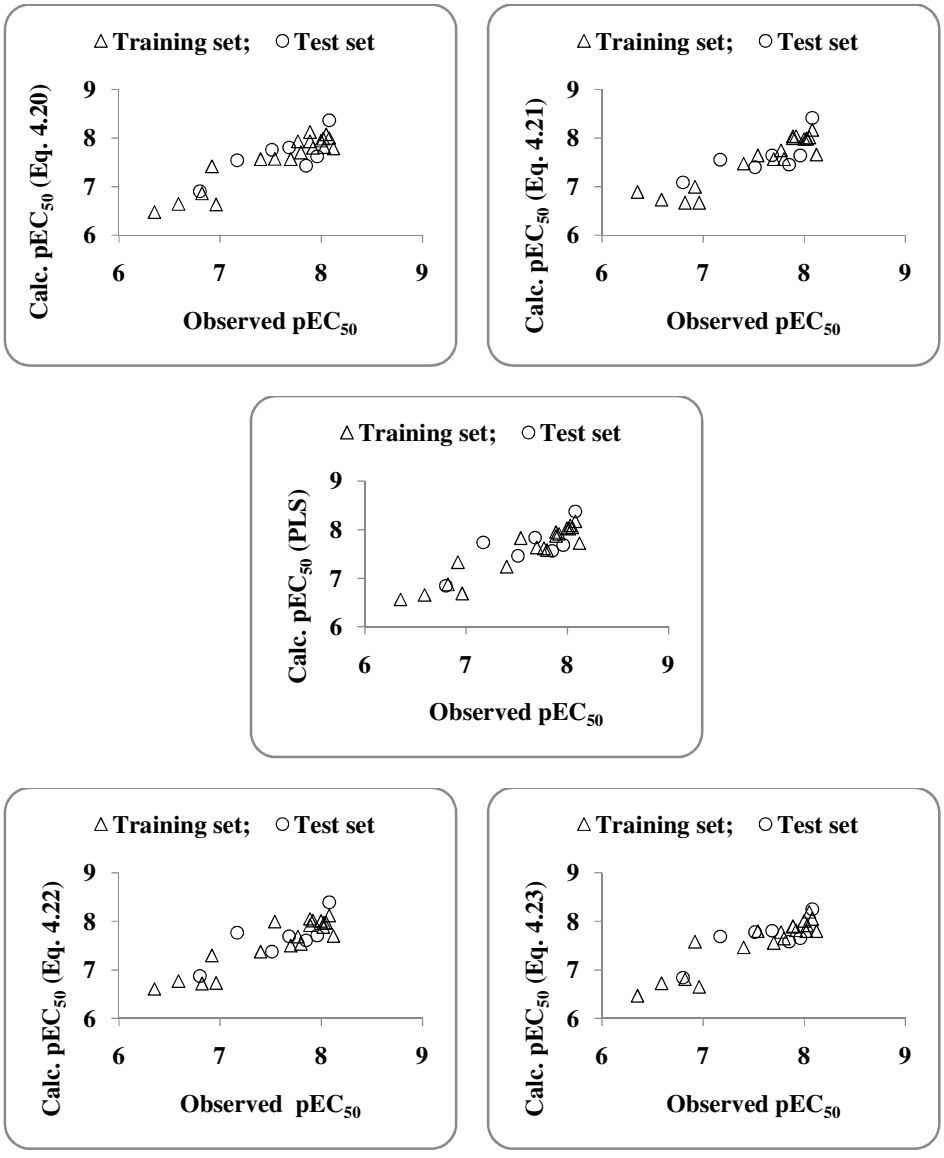


Figure 4.5: plot of observed and calculated pEC_{50} values of training- and test-set compounds for PPARγ transactivation.

A partial least square (PLS) analysis has been carried out on these 26 CP-MLR identified descriptors (Table 4.8) to facilitate the development of a "single window" structure–activity model. For the purpose of PLS, the descriptors have been auto-scaled (zero mean and unit SD) to give each one of them equal weight in the analysis. In the PLS cross-validation, two components are found to be the optimum for these 10 descriptors and they explained 88.36% variance in the activity (r^2 = 0.940, Q^2_{LOO} = 0.819, s = 0.202, F = 60.955, r^2_{Test} = 0.517). The MLR-like PLS coefficients of these 26 descriptors are given in Table 4.11.

Table 4.11: PLS and MLR-like PLS models from the descriptors of three parameter CP-MLR models for PPARγ transactivation activities.

A: PLS equation									
PLS components					PLS coefficient (s.e.)[a]				
Component-1					-0.171(0.016)				
Component-2					-0.078(0.021)				
Constant					7.571				
B: MLR-like PLS equation									
S. No.	Descriptor	MLR-like coef.[b]	(f.c.)[c]	Order	S. No.	Descriptor	MLR-like coef.[b]	(f.c.)[c]	Order
1	HNar	-0.149	-0.043	10	14	GGI7	0.116	0.037	12
2	MAXDP	-0.026	-0.006	24	15	JGI2	0.053	0.020	20
3	BAC	0.156	0.048	9	16	JGI4	0.100	0.033	14
4	Lop	0.287	0.092	1	17	JGT	-0.018	-0.005	25
5	Uindex	0.194	0.064	5	18	MATS8m	-0.204	-0.063	6
6	BIC3	0.111	0.026	18	19	MATS4v	0.216	0.065	4
7	T(N..O)	-0.151	-0.040	11	20	MATS3e	0.310	0.067	3
8	BEHm3	0.061	0.019	21	21	MATS3p	-0.022	-0.006	23
9	BELm5	0.272	0.088	2	22	MATS5p	-0.006	-0.002	26
10	BEHv2	-0.111	-0.033	13	23	GATS5p	0.094	0.029	17
11	BELv8	0.210	0.061	7	24	nCrH2	-0.117	-0.032	15
12	GGI2	0.068	0.022	19	25	Hy	-0.442	-0.016	22
13	GGI4	0.110	0.031	16	26	MLOGP	0.194	0.053	8
				Constant		6.558			
C: PLS regression statistics					Values				
n					19				
r					0.940				
s					0.202				
F					60.955				
FIT					5.300				
LOF					0.055				
AIC					0.056				
Q^2_{LOO}					0.819				

Q^2_{L5O}	0.797
r^2_{Test}	0.517

[a]Regression coefficient of PLS factor and its standard error. [b]Coefficients of MLR-like PLS equation in terms of descriptors for their original values;[c]f.c. is fraction contribution of regression coefficient, computed from the normalized regression coefficients obtained from the autoscaled (zero mean and unit s.d.) data.

For the sake of comparison, the plot showing goodness of fit between observed and calculated activities (through PLS analysis) for the training and test set compounds is also given in Figure 4.5. Figure 4.6 shows a plot of the fraction contribution of normalized regression coefficients of these descriptors to the activity.

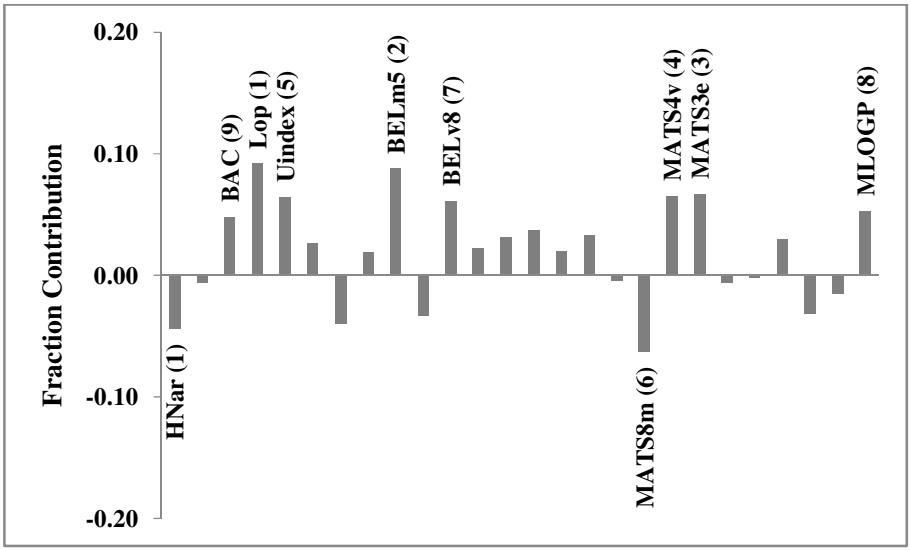


Figure 4.6: Plot of fraction contribution of MLR-like PLS coefficients (normalized) against 26 CP-MLR identified descriptors (Table 4.11) associated with PPARγ transactivation activity of benzylpyrazole acylsulfonamides.

The PLS analysis has suggested Lop as the most determining descriptor for modeling the activity of the compounds (descriptor S. No. 4 in Table 4.11; Figure 4.6). The other nine significant descriptors in decreasing order of significance are BELm5, MATS3e, MATS4v, Uindex, MATS8m, BELv8, MLOGP, BAC and HNar. Descriptors Lop, BELm5 and MLOGP are part of Eqs. (4.16) to (4.23) and convey same inference in the PLS model as well.

It is inferred from the PLS analysis that a higher values of 2D autocorrelation descriptors MATS3e (Moran autocorrelation of lag-3/ weighted by atomic Sanderson electronegativities) and MATS4v (Moran autocorrelation of lag-4/ weighted by atomic van der Waals volumes), topological descriptors Uindex (Balaban U index) and BAC (Balaban centric index); and modified Burden eigenvalue class descriptor BELv8 (lowest eigenvalue n.8 of Burden matrix/weighted by van der Waals volumes) would be advantageous to the activity. Based on the similar grounds a lower value of Moran autocorrelation of lag-8/ weighted by atomic masses (descriptor MATS8m) and Narumi harmonic index (descriptor HNar) will be supportive to the activity. It is also observed that PLS model from the dataset devoid of CP-MLR identified 26 descriptors (Table 4.11) is inferior in explaining the activity of the analogues.

2.2.1.2. APPLICABILITY DOMAIN (AD)

To analyze the applicability domain (AD) a Williams plot of the model based on the whole data set (Table 4.12) has been constructed that is shown in Figure 4.7.

From the analysis it has appeared that none of the compounds were identified as an obvious outlier for the PPARγ transactivation activities if the limit of normal values for the *Y* outliers (response outliers) was set as 3 (standard deviation) units. One compound listed in Table 4.8 at S. No. 1 found to have leverage (h) values greater than the threshold leverage (h*) suggesting it as chemically influential compound.

For both the training-set and test-set, the suggested model matches the high quality parameters with good fitting power and the capability of assessing external data. Furthermore, all of the compounds were within the applicability domain of the proposed model and were evaluated correctly.

Table 4.12: Models derived for the whole data set (n = 26) for the PPARγ transactivation activity in descriptors identified through CP-MLR.

Model	r	s	F	Eq.
pEC_{50} = 5.430 +1.259(0.267)MAXDP	0.905	0.237	33.591	(4.20a)

+1.583(0.197)Lop+0.795(0.140)JGI4				
pEC_{50} = 6.419 + 1.191(0.189)Lop –0.572(0.219)T(N..O)+1.029(0.185)GGI4	0.893	0.251	29.140	(4.21a)
pEC_{50} = 6.063 +1.135(0.190)Lop +0.922(0.199)GGI4+0.471(0.189)MLOGP	0.891	0.254	28.302	(4.22a)
pEC_{50} = 5.699 + 0.796(0.290)MAXDP +1.409(0.210)Lop+0.977(0.186)GGI4	0.896	0.248	30.001	(4.23a)

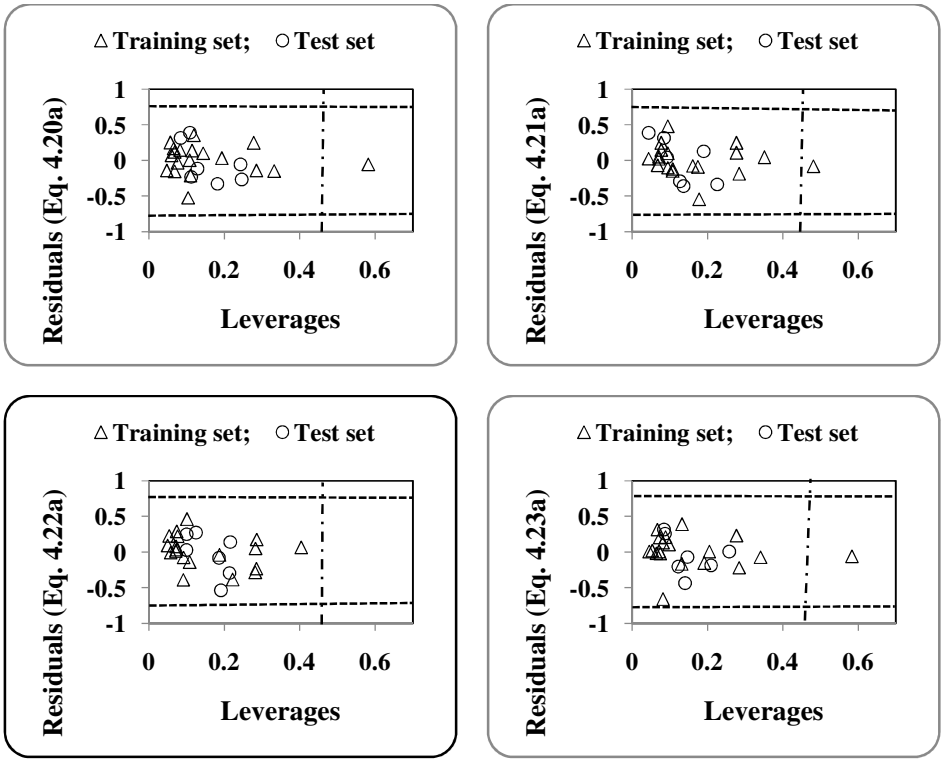


Figure 4.7: Williams plot for the training-set and test- set compounds for PPARγ transactivation activity. The horizontal dotted line refers to the residual limit (±3×standard deviation) and the vertical dotted line represents threshold leverage h* (= 0.46).

2.2.2. CONCLUSIONS

The PPARγ transactivation activity of benzylpyrazole acylsulfonamide derivatives have been quantitatively analyzed in terms of 0D- to 2D-Dragon descriptors. This study has provided a rational approach for the development of titled derivatives as PPARγ agonists. The descriptors identified in CP-MLR

analysis for the PPARγ transactivation activity have highlighted the role of atomic properties (mass, electronegativity, van der Waals volumes and polarizability) in terms of weighted 2D autocorrelations and BCUT descriptors and electronic content in terms of Galvez charge indices and maximal electrotopological positive variation (MAXDP). Additionally, Balaban's U and centric indices (Uindex and BAC, respectively), Lopping centric index (Lop), topological distance between N..O and hydrophobicity accounting parameter MLOGP have also shown prevalence to optimize the PPARγ transactivation of titled compounds. PLS analysis has further confirmed the dominance of the CP-MLR identified descriptors and applicability domain analysis revealed that the suggested model matches the high quality parameters with good fitting power and the capability of assessing external data and all of the compounds was within the applicability domain of the proposed model and were evaluated correctly.

2.3. PYRIDYLOXYBENZENE-ACYLSULFONAMIDES AS PPARγ AGONISTS

In the field of antidiabetic drug discovery and development, the findings that TZDs are high affinity ligands for peroxisome proliferator-activated receptor γ (PPARγ), opened channels for the extensive research [752, 755, 775, 776, 779]. The binding of TZD activates PPARγ which functions as an essential transcriptional regulator of glucose and lipid homeostasis. PPARγ is the most broadly studied subtype among the three PPAR subtypes (namely designated as PPARα, PPARγ, and PPARδ). PPARγ, expressed predominantly in adipose tissue, regulate the expression of a constellation of genes which is closely related to adipocyte differentiation, glucose and lipid metabolism, insulin sensitivity, inflammatory responses and cell proliferation [757, 780]. The majority of reported PPARγ ligands like TZD, oxazolidinone and tetrazole possess a carboxylic acid or its heterocyclic bioisostere [780-786]. There is also an example of non-TZD and non-carboxylic acid PPARγ agonists [777]. A novel class of pyridyloxybenzene-acylsulfonamides

as non-thiazolidinedione (TZD), non-carboxylic-acid-based selective PPARγ agonists has been reported by Rikimaru et al. [787].

The reported thirty four pyridyloxybenzene-acylsulfonamides are considered as the data set for this study [787]. These derivatives were evaluated for their transactivation activity against human PPARγ stably expressed in Chinese hamster ovary (CHO)–K1 cells. Transactivation activities were assessed by a luciferase reporter gene assay using (*R*)-5-(3-{4-[(2-Furan-2-yl-5-methyl-1,3-oxazol-4-yl)methoxy]-3-methoxyphenyl}propyl)-1,3oxazolidine -2,4-dione [778] as the reference PPARγ agonist and were reported as EC_{50}. The general structure of these analogues is represented in Figure 4.8 and the structural variations of these analogues along with their reported pEC_{50}, on molar basis, are mentioned in Table 4.13.

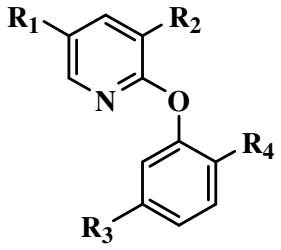
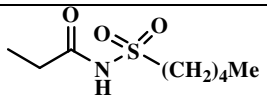
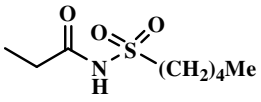
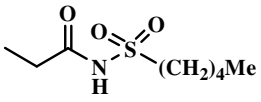
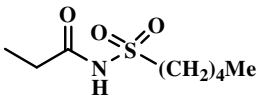
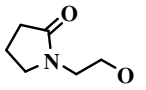
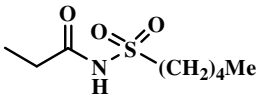
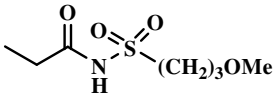
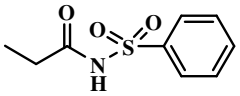
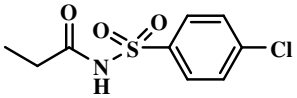
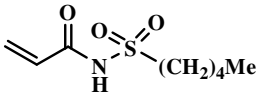
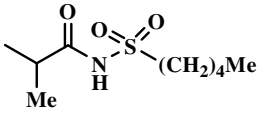
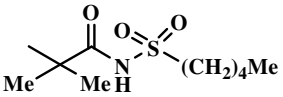
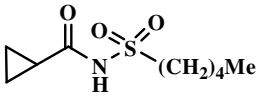
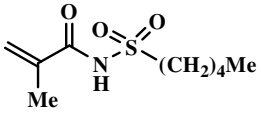
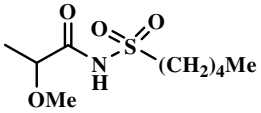
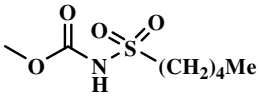
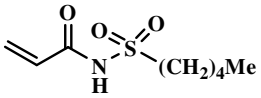
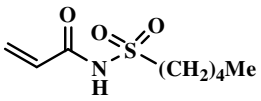
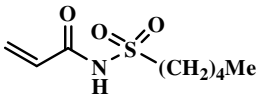
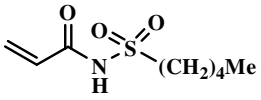
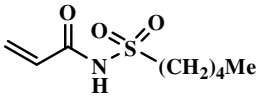
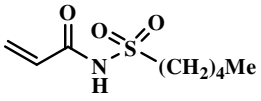
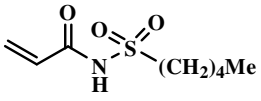
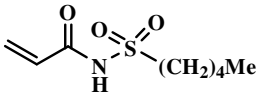
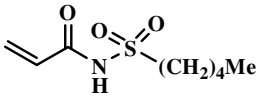
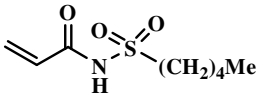
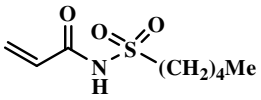
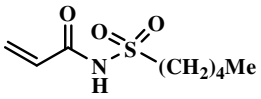
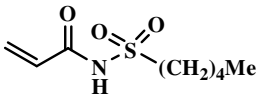
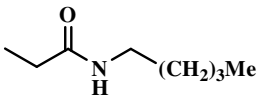


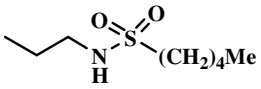
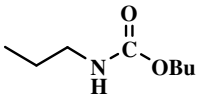
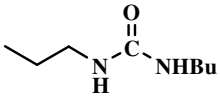
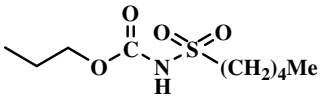
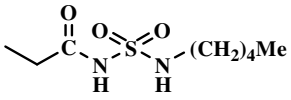
Figure 4.8: General structure of pyridyloxybenzene-acylsulfonamides.

Table 4.13: Structural variations and reported PPARγ transactivation activities of pyridyloxybenzene-acylsulfonamides.

S. No.	R_1	R_2	R_3	R_4	pEC_{50}[a]
1	CF_3	Cl	H		7.21
2	CF_3	Cl	*i*-PrO		8.59
3[b]	CF_3	Cl	$MeO(CH_2)_2O$		8.82

4	CF_3	Cl	$Et_2NC(=O)CH_2O$		8.16
5	CF_3	Cl			8.47
6	CF_3	Cl	$MeO(CH_2)_2O$		7.96
7[b]	CF_3	Cl	$MeO(CH_2)_2O$		7.49
8[b]	CF_3	Cl	$MeO(CH_2)_2O$		8.07
9[b]	CF_3	Cl	$MeO(CH_2)_2O$		7.82
10	CF_3	Cl	$MeO(CH_2)_2O$		8.64
11[b]	CF_3	Cl	$MeO(CH_2)_2O$		8.51
12	CF_3	Cl	$MeO(CH_2)_2O$		8.26
13[b]	CF_3	Cl	$MeO(CH_2)_2O$		8.28
14[b]	CF_3	Cl	$MeO(CH_2)_2O$		8.17
15	CF_3	Cl	$MeO(CH_2)_2O$		8.00
16*	CF_3	Cl	$MeO(CH_2)_2O$		7.03

17[c]	CF_3	H	$MeO(CH_2)_2O$		-[c]
18[c]	H	Cl	$MeO(CH_2)_2O$		-[c]
19	Cl	Cl	$MeO(CH_2)_2O$		6.72
20[c]	NH_2	Me	$MeO(CH_2)_2O$		-[c]
21[c]	MeCONH	Me	$MeO(CH_2)_2O$		-[c]
22	CF_3	Cl	*i*-PrO-		7.59
23	CF_3	Cl	*c*-$PrO(CH_2)_2O$		8.21
24	CF_3	Cl	$MeO(CH_2)_3O$		8.13
25[b]	CF_3	Cl	$HO(CH_2)_2O$		7.10
26	CF_3	Cl	$MeC(=O)(CH_2)_3O$		7.23
27	CF_3	Cl	$MeSO_2(CH_2)_3O$		6.49
28	CF_3	Cl	$NC(CH_2)_3O$		7.82
29[b]	CF_3	Cl	*i*-PrO		6.54

30	CF_3	Cl	$MeO(CH_2)_2O$		7.49
31	CF_3	Cl	$MeO(CH_2)_2O$		7.42
32[b]	CF_3	Cl	$MeO(CH_2)_2O$		7.32
33	CF_3	Cl	$MeO(CH_2)_2O$		8.35
34	CF_3	Cl	$MeO(CH_2)_2O$		8.96

[a]EC_{50} (the effective concentration for 50% response of a given compound's intrinsic maximum response) on molar basis, taken from reference [787]; [b]Compound included in test set; [c]Compound with uncertain activity, not part of data set; *Benzene ring instead of pyridine ring.

2.3.1. RESULTS AND DISCUSSION

2.3.1.1. QSAR RESULTS

For the purpose of modelling study, one third of total active compounds (10) have been included in the test set for the validation of the models derived from remaining 20 training set compounds. Compounds at S. No. 17, 18, 20 and 21 (Table 4.13) having uncertain activities are not part of data set. Dragon software computed a total number of 496 descriptors, belonging to 0D- to 2D-modules but after the reduction of descriptor data set only 120 relevant descriptors were obtained. These 120 significant descriptors have been subjected to CP-MLR analysis with default "filters" set in it. Statistical models in two, three and four descriptors have been explored to achieve the best relationship correlating PPARγ transactivation activity. The obtained two and three descriptor models are given below.

$pEC_{50} = 7.153 + 2.563(0.541)Qindex – 2.460(0.554)BEHm4$

$n = 20, r = 0.769, s = 0.452, F = 12.352, Q^2_{LOO} = 0.490, Q^2_{L5O} = 0.225$

$r^2_{Test} = 0.297$, FIT = 1.027, LOF = 0.271, AIC = 0.276 **(4.24)**

pEC_{50} = 7.420 + 2.279(0.575)Wap – 1.875(0.536)BEHm4

n = 20, r = 0.712, s = 0.496, F = 8.768, Q^2_{LOO} = 0.303, Q^2_{L5O} = 0.185

r^2_{Test} = 0.141, FIT = 0.730, LOF = 0.327, AIC = 0.333 **(4.25)**

pEC_{50} = 6.582 + 1.811(0.554)MW – 3.195(0.680)T(O..S)

+ 1.699(0.527)GATS1p

n = 20, r = 0.823, s = 0.413, F = 11.268, Q^2_{LOO} = 0.540, Q^2_{L5O} = 0.573

r^2_{Test} = 0.505, FIT = 1.165, LOF = 0.279, AIC = 0.256 **(4.26)**

The descriptors Qindex, Wap and T(O..S) participated in above models are topological descriptors. Descriptors BEHm4, MW and GATS1p are from the constitutional (CONST), modified Burden eigenvalue (BCUT) and 2D-autocorrealation (2D-AUTO) classes, respectively. Except BEHm4 and T(O..S), all the descriptors have shown positive influence on the activity as evident from the signs of regression coefficients. Thus a higher value of descriptors Qindex (Quadratic index), Wap (all-path Wiener index), MW (molecular weight) and GATS1p (Geary autocorrelation of lag-1/weighted by atomic polarizabilities) in addition to a lower value of the highest eigenvalue n.4 of Burden matrix/weighted by atomic masses (descriptor BEHm4) and sum of topological distances between O and S atoms (descriptor T(O..S)) would be beneficiary to the activity. The three descriptor model could estimate nearly 68% variance in observed activity of the compounds.

Considering the number of observation in the dataset, models with up to four descriptors were explored. It has resulted in 37 models with test set $r^2 >$ 0.50. These models (with 120 descriptors) were identified in CP-MLR by successively incrementing the filter-3 with increasing number of descriptors (per equation). For this, the optimum *r*-bar value of the preceding level model (=0.786) has been used as the new threshold of filter-3 for the next generation. These models have shared 43 descriptors among them. All these shared descriptors along with their brief meaning, average regression coefficients, and total incidence are listed in Table 4.14, which will serve as a measure of their estimate across these models.

Table 4.14: Identified descriptors[a] along with their class, average regression coefficient and incidence[b], in modeling the PPARγ transactivation activities of pyridyloxybenzene-acylsulfonamides.

Descriptor class, average regression coefficient and (incidence)	
Constitutional descriptors (**CONST**)	MW, 2.286 (17); nBM, -2.033 (1); nCIC, 0.826 (2); ARR,-1.692 (2); RBN, 1.203 (2); RBF, 0.966 (3); nDB, -1.756 (2); nN, 1.083 (1)
Topological descriptors (**TOPO**)	AAC, 1.986 (3); Qindex, 2.617 (6); GNar, 1.813 (1); JhetZ, -1.360 (5); MAXDP, 1.319 (1); X1A, -1.363 (7); X2A, -1.112 (1); X1Av, -1.388 (1); S2K, -1.479 (2); Lop, -1.586 (1); IDDE, 1.574 (6); SIC2, -1.944 (1); VEA1, 1.165 (1); T(N..Cl), 0.990 (6); T(O..S), -2.922 (16)
Modified Burden Eigen values (**BCUT**)	BEHm4, 1.720 (1) and -2.330 (15); BEHm7, 1.462 (4); BELm7, -1.572 (2); BELm8, -1.897 (1); BEHv1, -0.976 (3); BELv4, -2.323 (4); BELv8, 2.764 (1); BELp3, 1.379 (2)
Galvez Topological charge indices (**GALVEZ**)	GGI4, 1.946 (2); JGI3, 1.927 (2); JGI4, 1.330 (1); JGI5, 0.796 (1); JGT, 1.588 (2)
2D autocorrelations (**2D-AUTO**)	MATS1v, -0.963 (1); MATS2e, 1.600 (1); MATS3e, 1.048 (1); GATS1v, 1.554 (2); GATS1p, 2.198 (11)
Atom centered fragments (**ACF**)	H-046, 2.005 (2); H-047, -1.255 (1)

[a]The descriptors are identified from the three parameter models for PPARγ binding activity transactivation activity emerged from CP-MLR protocol with filter-1 as 0.3, filter-2 as 2.0, filter-3 as 0.786 and filter-4 as $0.3 \leq q^2 \leq 1.0$ with a training set of 20 compounds. [b]The average regression coefficient of the descriptor corresponding to all models and the total number of its incidence. The arithmetic sign of the coefficient represents the actual sign of the regression coefficient in the models. **CONST:** MW, molecular weight; nBM, number of multiple bonds; nCIC, number of rings; ARR, aromatic ratio ; RBN, number of rotatable bonds; RBF, rotatable bond fraction; nDB, number of double bonds; nN, number of Nitrogen atoms; **TOPO:** AAC,mean information index on atomic composition; Qindex, Quadratic index; GNar, Narumi geometric topological index; JhetZ, Balaban-type index from Z weighted distance matrix (Barysz matrix); MAXDP, maximal electrotopological positive variation; X1A, average connectivity index chi-1; X2A, average connectivity index chi-2; X1Av, average valence connectivity index chi-1; S2K, 2-path Kier alpha-modified shape index; Lop, Lopping centric index; IDDE, mean information content on the

distance degree equality; SIC2, structural information content (neighborhood symmetry of 2-order); VEA1, eigenvector coefficient sum from adjacency matrix; T(N..Cl), sum of topological distances between N..Cl; T(O..S),- sum of topological distances between O..S; **BCUT:** BEHm4, highest eigenvalue n.4 of Burden matrix/weighted by atomic masses; BEHm7, highest eigenvalue n.7 of Burden matrix/weighted by atomic masses; BELm7, lowest eigenvalue n.7 of Burden matrix/weighted by atomic masses; BELm8, lowest eigenvalue n.8 of Burden matrix/weighted by atomic masses; BEHv1, highest eigenvalue n.1 of Burden matrix/weighted by van der Waals volumes; BELv4, lowest eigenvalue n.4 of Burden matrix/weighted by van der Waals volumes; BELv8, lowest eigenvalue n.8 of Burden matrix/weighted by van der Waals volumes; BELp3, lowest eigenvalue n.3 of Burden matrix/weighted by atomic polarizabilities; **GALVEZ:** GGI4, topological charge index of order 4; JGI3,mean topological charge index of order 3; JGI4, mean topological charge index of order 4; JGI5, mean topological charge index of order 5; JGT, global topological charge index; **2D-AUTO:** MATS1v, Moran autocorrelation of lag-1/ weighted by atomic van der Waals volumes; MATS2e, Moran autocorrelation of lag-2/ weighted by atomic Sanderson electronegativities; MATS3e, Moran autocorrelation of lag-3/ weighted by atomic Sanderson electronegativities; GATS1v, Geary autocorrelation of lag-1/ weighted by atomic van der Waals volumes; GATS1p, Geary autocorrelation of lag-1/weighted by atomic polarizabilities; **ACF:** H-046, H attached to C0(sp3) no X attached to next C atom; H-047, H attached to C1(sp3) / C0(sp2).

The selected four-descriptor models for the PPARγ transactivation activities of pyridyloxybenzene-acylsulfonamides emerged through CP-MLR are presented through Eqs. (4.27) to (4.30).

$$pEC_{50} = 6.025 + 2.640(0.407)Qindex + 1.131(0.345)T(N..Cl)$$
$$- 2.345(0.437)BEHm4 + 1.131(0.392)GATS1v$$
$$n = 20, r = 0.899, s = 0.329, F = 15.918, Q^2_{LOO} = 0.716, Q^2_{L5O} = 0.661$$
$$r^2_{Test} = 0.602, FIT = 1.768, LOF = 0.225, AIC = 0.180 \quad \textbf{(4.27)}$$

$$pEC_{50} = 6.554 + 0.828(0.308)RBN + 2.519(0.427)Qindex$$
$$+ 0.840(0.357)T(N..Cl) - 2.369(0.447)BEHm4$$
$$n = 20, r = 0.894, s = 0.337, F = 14.988, Q^2_{LOO} = 0.691, Q^2_{L5O} = 0.636$$
$$r^2_{Test} = 0.548, FIT = 1.665, LOF = 0.236, AIC = 0.189 \quad \textbf{(4.28)}$$

$$pEC_{50} = 6.422 + 0.645(0.257)RBF + 2.773(0.422)Qindex$$
$$+ 1.013(0.358)T(N..Cl) - 2.246(0.477)BEHm4$$
$$n = 20, r = 0.889, s = 0.344, F = 14.205, Q^2_{LOO} = 0.665, Q^2_{L5O} = 0.655$$

r^2_{Test} = 0.512, FIT = 1.578, LOF = 0.247, AIC = 0.197 **(4.29)**

pEC_{50} = 7.084 + 1.332(0.501)MW – 0.924(0.339)X1A
– 3.321(0.577)T(O..S) + 2.193(0.481)GATS1p
n = 20, r = 0.885, s = 0.349, F = 13.683, Q^2_{LOO} = 0.648, Q^2_{L5O} = 0.628
r^2_{Test} = 0.573, FIT = 1.520, LOF = 0.254, AIC = 0.203 **(4.30)**

The newly appeared descriptors in above models are, T(N..Cl) and X1A (topological descriptors); RBN and RBF (constitutional descriptors); and GATS1v (a 2D-AUTO class descriptor). Descriptors T(N..Cl), RBN, RBF and GATS1v have correlated positively to the PPARγ transactivation whereas descriptor X1A influenced it negatively.

Thus from the signs of regression coefficients of these descriptors it is evident that higher values of the sum of topological distances between N and Cl atoms (descriptor T(N..Cl)), presence of more number of rotatable bonds (descriptor RBN), higher value of rotatable bond fraction (descriptor RBF) in a molecular structure and a higher value of Geary autocorrelation of lag-1/weighted by atomic polarizabilities (GATS1p) would be beneficial to the activity, whereas a lower value of descriptor X1A (average connectivity index chi-1) would be advantageous to the activity.

These models have accounted for nearly 81% variance in the observed activities. In the randomization study (100 simulations per model), none of the identified models has shown any chance correlation. The values greater than 0.5 of Q^2 index is in accordance to a reasonable robust QSAR model. The pEC_{50} values of training set compounds calculated using Eqs. (4.27) to (4.30) and predicted from LOO procedure have been included in Table 4.15.

The models (4.27) to (4.30) are validated with an external test set of 10 compounds mentioned in Table 4.13. The predictions of the test set compounds based on external validation are found to be satisfactory as reflected in the test set r^2 (r^2_{Test}) values and the same is reported in Table 4.15. The plot showing goodness of fit between observed and calculated activities for the training and test set compounds is given in Figure 4.9.

Table 4.15: Observed and modeled PPARγ transactivation activity of pyridyloxybenzene-acylsulfonamides.

S. No.	pEC50(M)[a]								
	Obsd.[b]	Eq. (4.27)		Eq. (4.28)		Eq. (4.29)		Eq. (4.30)	
		Calc.	Pred.[c]	Calc.	Pred.[c]	Calc.	Pred.[c]	Calc.	Pred.[c]
1	7.21	7.23	7.24	7.29	7.33	7.43	7.48	7.29	7.35
2	8.59	8.24	8.18	8.32	8.28	8.39	8.36	8.14	8.06
3[d]	8.82	8.08	-[d]	8.12	-[d]	8.14	-[d]	7.96	-[d]
4	8.16	8.43	8.52	8.56	8.79	8.45	8.57	8.55	8.72
5	8.47	8.62	8.70	8.52	8.55	8.49	8.50	8.53	8.56
6	7.96	8.21	8.25	8.12	8.14	8.24	8.29	8.16	8.20
7[d]	7.49	7.58	-[d]	7.85	-[d]	8.01	-[d]	7.40	-[d]
8[d]	8.07	7.65	-[d]	7.84	-[d]	8.24	-[d]	7.19	-[d]
9[d]	7.82	7.61	-[d]	7.51	-[d]	7.49	-[d]	7.67	-[d]
10	8.64	8.20	8.16	8.28	8.23	8.24	8.19	8.19	8.15
11[d]	8.51	8.52	-[d]	8.62	-[d]	8.56	-[d]	8.58	-[d]
12	8.26	8.50	8.60	8.46	8.55	8.53	8.64	8.26	8.26
13[d]	8.28	7.67	-[d]	7.61	-[d]	7.53	-[d]	7.91	-[d]
14[d]	8.17	8.33	-[d]	8.35	-[d]	8.30	-[d]	8.42	-[d]
15	8.00	8.21	8.24	8.12	8.14	8.24	8.28	8.19	8.24
16	7.03	7.36	7.43	7.40	7.47	7.26	7.33	7.39	7.42
17[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
18[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
19	6.72	6.81	6.91	6.76	6.82	6.77	6.85	6.39	6.14
20[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
21[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
22	7.59	7.73	7.75	7.68	7.70	7.70	7.73	7.85	7.89
23	8.21	7.69	7.55	7.67	7.52	7.65	7.48	8.06	8.02
24	8.13	7.63	7.59	7.58	7.53	7.45	7.37	7.77	7.73
25[d]	7.10	7.31	-[d]	7.42	-[d]	7.41	-[d]	7.06	-[d]
26	7.23	7.54	7.58	7.68	7.71	7.60	7.63	7.60	7.67
27	6.49	6.49	6.49	6.50	6.51	6.57	6.88	6.55	6.84
28	7.82	7.77	7.74	7.85	7.86	7.92	7.96	7.53	7.42

29[d]	6.54	7.20	-[d]	7.28	-[d]	7.26	-[d]	6.89	-[d]
30	7.49	7.86	7.92	7.81	7.85	7.70	7.73	7.94	8.03
31	7.42	7.22	7.13	7.18	7.07	7.20	7.09	7.65	7.75
32[d]	7.32	7.63	-[d]	7.58	-[d]	7.59	-[d]	7.37	-[d]
33	8.35	8.37	8.38	8.39	8.39	8.31	8.30	8.20	8.16
34	8.96	8.62	8.49	8.56	8.43	8.57	8.43	8.49	8.37

[a]On molar basis; [b]Taken from ref. [787]; [c]Leave-one-out (LOO) procedure; [d]Compound included in test set; [e]Compound with uncertain activity, not part of data set.

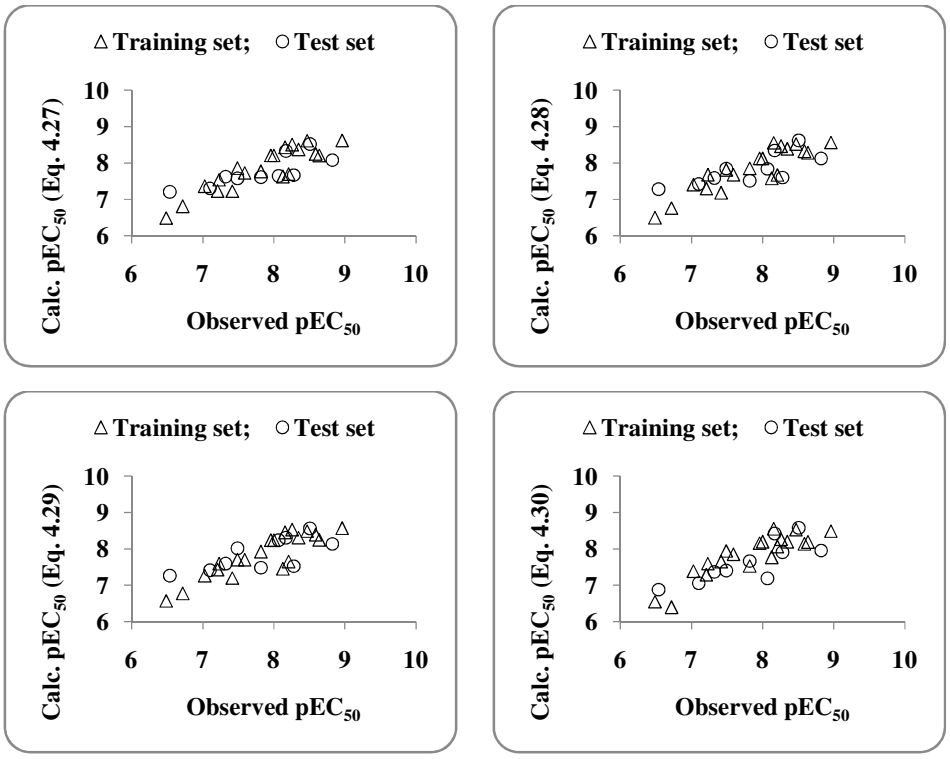


Figure 4.9: Plot of observed and calculated pEC_{50} values of training- and test-set compounds for PPARγ transactivation.

2.3.1.2. APPLICABILITY DOMAIN (AD)

To analyze the applicability domain (AD) a Williams plot of the model based on the whole data set (Table 4.16) has been constructed that is shown in Figure 4.10. On analyzing the model AD in the Williams plot it has appeared that none of the compounds were identified as an obvious outlier for the PPARγ transactivation activities if the limit of normal values for the *Y* outliers

(response outliers) was set as 3 (standard deviation) units. Two compounds listed in Table 4.12 at S. No. **8** and **27** found to have leverage (h) values greater than the threshold leverage (h*) suggesting them as chemically influential compounds. For both the training-set and test-set, the suggested model matches the high quality parameters with good fitting power and the capability of assessing external data.

Furthermore, all of the compounds were within the applicability domain of the proposed model and were evaluated correctly.

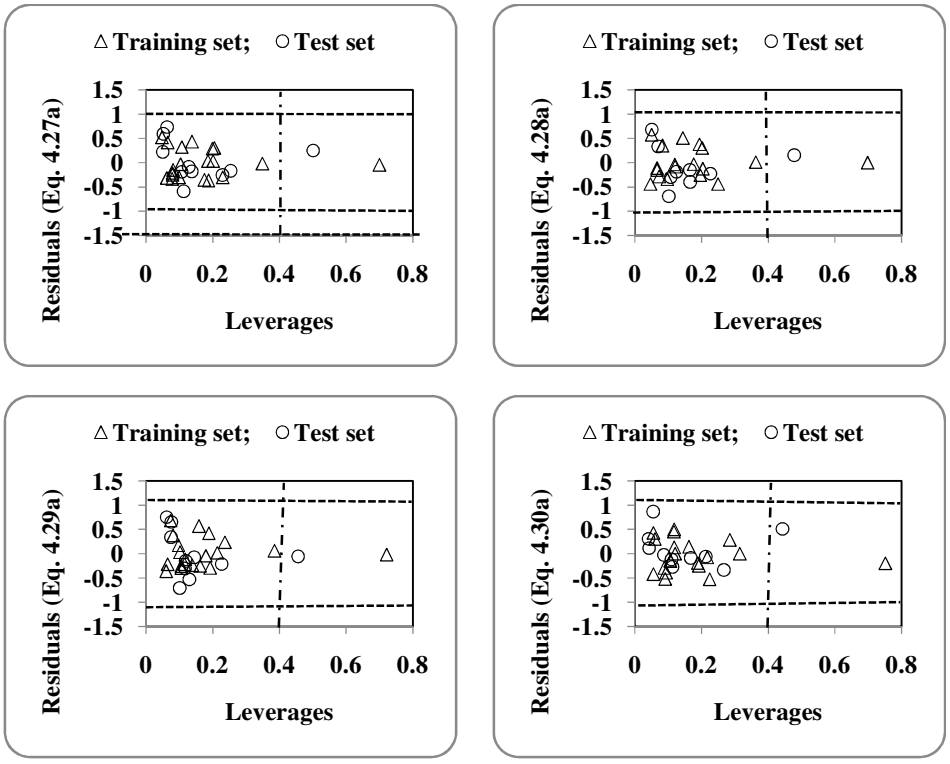


Figure 4.10: Williams plot for PPARγ transactivation activity. The horizontal dotted line refers to the residual limit (±3×standard deviation) and the vertical dotted line represents threshold leverage h* (= 0.40).

Table 4.16: Models derived for the whole data set (n = 30) for the PPARγ transactivation activity in descriptors identified through CP-MLR.

Model	r	s	F	Eq.
pEC_{50} = 5.889 + 2.887(0.392)Qindex +1.221(0.303)T(N..Cl)–2.381(0.461)BEHm4	0.865	0.359	18.664	(4.27a)

+ 1.164(0.317)GATS1v				
pEC_{50} = 6.468 +0.850(0.281)RBN +2.669(0.429)Qindex + 0.890(0.319)T(N..Cl) – 2.416(0.494)BEHm4	0.846	0.382	15.847	(4.28a)
pEC_{50} = 6.399 + 0.663(0.248)RBF +2.872(0.430)Qindex + 0.919(0.328)T(N..Cl) – 2.353(0.526)BEHm4	0.836	0.393	14.569	(4.29a)
pEC_{50} = 7.116 + 1.730(0.450)MW – 0.903(0.321)X1A – 3.351(0.577)T(O..S) + 1.853(0.359)GATS1p	0.860	0.366	17.816	(4.30a)

2.3.2. CONCLUSIONS

QSAR rationales have been obtained for the PPARγ transactivation activity of pyridyloxybenzene-acylsulfonamides in terms of 0D- to 2D-Dragon descriptors. The descriptors identified in CP-MLR analysis have highlighted the role of atomic mass, van der Waals volumes and polarizability through weighted 2D autocorrelations (GATS1v and GATS1p), modified Burden eigenvalue (BEHm4) and molecular weight (MW). Sum of topological distances between O and S (descriptor T(O..S)), and N and Cl (descriptor T(N..Cl)), average connectivity index chi-1(X1A) and Quadratic index (Qindex) have also shown dominance to optimize the PPARγ transactivation. Descriptors RBN and RBF suggested presence of rotatable bonds in a molecular structure for better PPARγ activity. Applicability domain analysis revealed that the suggested model matches the high quality parameters with good fitting power and the capability of assessing external data and all of the compounds was within the applicability domain of the proposed model and were evaluated correctly.

REFERENCES

743. American Diabetes Association. Diabetes Care (2004) 27:S5.

744. Porte D Jr and Schwartz MW. Science (1996) 272:699-700.

745. Saltiel AR. Cell (2001)104:517-529.

746. Zimmet P, Alberti KG and Shaw J. Nature (2001) 414:782-787.

747. Kopelman PG and Hitman GA. Lancet (1998) 352:SIV5

748. King H, Aubert RE and Herman WH. Diabetes Care (1998) 21:1414-1431.

749. Moller DE. Nature (2001) 414:821-827.

750. Mangelsdorf DJ and Evans RM. Cell (1995) 83:841-850.

751. Isseman I and Green S. Nature (1990) 347:645-650.

752. Berger J and Moller DE. Annu Rev of Med (2002) 53:409-435.

753. Staels B and Fruchart J-C. Diabetes (2005) 54:2460-2470.

754. Berger JP, Akiyama TE and Meinke PT. Trends Pharmacol Sci (2005) 26:244-251.

755. Willson TM, Brown PJ, Sternbach DD and Henke BR. J Med Chem (2000) 43:527-550.

756. Henke BR, J Med Chem (2004) 47:4118-4127.

757. Lehmann JM, Moore MB, Smith-Oliver TA, Wilkinson WO, Willson TM and Kliewer SA. J Biol Chem (1995) 270:12953-2956.

758. Willson TM, Cobb JE, Cowan DJ, Wiethe RW, Correa ID, Prakash SR, Beck KD, Moore LB, Kliewer SA and Lehmann JM. J Med Chem (1996) 39:665-668.

759. Lohray BB, Lorhay VB, Bajji AC, Kalchar S, Poondra RR, Padakanti S, Chakrabarti R, Vikramadithyan RK, Misra P, Juluri S, Mamidi RNVS and Rajagopalan R. J Med Chem (2001) 44:2675-2678.

760. Sauerberg P, Petterson I, Jeppesen L, Bury PS, Mogensen JP, Wassermann K, Brand CL, Sturis J, Wöldike HF, Fleckner J, Anderson AST, Mortensen SB, Anders-Svensson L, Rasmussen HB, Lehmann SV, Polivka Z, Sindelar K, Panajotova V, Ynddal L and

Wulff EM. J Med Chem (2002) 45:789-804.

761. Ebdrup S, Petterson I, Rasmussen HB, Deussen HJ, Frost-Jensen A, Mortensen SB, Fleckner J, Pridal L, Nygaard L and Sauerberg P. J Med Chem (2003) 46:1306-1317.

762. Devasthale PV, Chen S, Jeon Y, Qu F, Shao C, Wang W, Zhang H, Farreley D, Golla R, Grover G, Harrity T, Ma Z, Moore L, Seethala R, Ren J, Cheng L, Sleph P, Sun W, Tieman A, Wetterau JR, Chang SY, Doweyko A, Chandrasena G, Humphreys JR, Sasseville VG, Ryono DE, Biller SA, Selan F, Harihanan N and Cheng PTW. J Med Chem (2005) 48:2248-2250.

763. Koyama H, Miller DJ, Boueres JK, Desai RC, Jones AB, Berger JP, MacNaul KL, Kelly LJ, Doebber TW, Wu MS, Zhou G, Wang PR, Ippolito MC, Chao YS, Agrawal AK, Franklin R, Heck JV, Wright SD, Moller DE and Sahoo SP. J Med Chem (2004) 47:3255-3263.

764. Shi GQ, Dropinski JF, McKeever BM, Xu S, Becker JW, Berger JP, MacNaul KL, Elbrecht A, Zhou G, Doebber TW, Wang P, Chao YS, Forrest M, Heck JV, Moller DE and Jones AB. J Med Chem (2005) 48:4457-4468.

765. Gervois P, Fruchart J-C and Staels B. Int J Clinic Pract (2004) 58:22-29.

766. Parmenon C, Viaud-Massuard M-C, Guillard J, Dacquet C, Ktorza A and Caignard D-H. (2006). WO2006/079719.

767. Parmenon C, Guillard J, Caignard D-H, Hennuyer N, Staels B, Audinot-Bouchez V, Boutin J-A, Ktorza A, Dacquet C and Viaud-Massuard M-C. Bioorg Med Chem Lett (2008) 18:1617-1622.

768. Parmenon C, Guillard J, Caignard D-H, Hennuyer N, Staels B, Audinot-Bouchez V, Boutin J-A, Dacquet C, Ktorza A, Viaud-Massuard M-C. Bioorg Med Chem Lett (2009) 19:2683-2687.

769. Golbraikh A and Tropsha A. J Mol Graph Model (2002) 20:269-276.

770. Ohkubo Y. Kishikawa H, Araki E, Miyata T, Lsami S, Motoyoshi S, Kojima Y, Furuyoshi N and Shichiri M. Diabs Res Clinic Pract (1995) 28:103-117.

771. Haffner SM, Alexander CM, Cook TJ, Boccuzzi SJ, Musliner TA, Pedersen TR, Kjekshus J and Pyörälä K. Arch Int Med, (1999) 159:2661-2667.

772. Pyörälä K, Pedersen TR, KJekshus J, Faergeman O, Olsson AG and Thorgeirsson G. Diabetes Care (1997) 20:614-620.

773. Sacks FM, Pfeffer MA, Moye LA, Rouleau JL, Rutherford ID, Cole TG, Brown L, Warnica JW, Arnold JMO, Wun CC, Davis BR and Braunwald E. New Eng J Med (1996) 335:1001-1009.

774. Vamecq J and Latruffe N. Curr Opin Endocrinol Diabetes (2000) 7:8-18.

775. Evans RM, Barish GD and Wang Y-X. Nat. Med. (2004) 10:355-361.

776. Cho N and Momose Y. Curr Topic Med Chem (2008) 8:1483-1507.

777. Rikimaru K, Wakabayashi T, Abe H, Imoto H, Maekawa T, Ujikawa O, Murase K, Matsuo T, Matsumoto M, Nomura C, Tsuge H, Arimura N; Kawakami K, Sakamoto J, Funami M, Mol CD, Snell GP, Bragstad KA, Sang B-C; Dougan DR, Tanaka T, Katayama N, Horiguchi Y, Momose Y. Bioorg Med Chem (2012) 20:714-733.

778. Momose Y, Maekawa T, Yamano T, Kawada M, Odaka H, Ikeda H and Sohda T. J Med Chem (2002) 45:1518-1534.

779. Henke BR. Prog Med Chem (2004) 42:1-53.

780. Loviscach M, Rehman N, Carter L, Mudaliar S, Mohadeen P, Ciaraldi TP, Veerkamp JH and Henry RR. Diabetlogia (2000) 43: 304-311.

781. Minoura H, Takeshita S, Ita M, Hirosumi J, Mabuchi M, Kawamura I, Nakajima S, Nakayama O, Kayakiri H, Oku T, Okubo-Suzuki A, Fukagawa M, Kojo H, Hanioka K, Yamasaki N, Imoto T, Kobayashi Y and Mutoh S. Eur J Pharmacol (2004) 494:273-281.

782. Minoura H, Takeshita S, Yamamoto T, Mabuchi M, Hirosumi J, Takakura S, Kawamura I, Seki J, Manda T, Ita M and Mutoh S. Eur J Pharmacol (2005) 519:182-190.

783. Kyowa Hakko Kirin Company Limited. Yanagisawa A, Otsubo W, Ueno K, Suzuki M, Matsubara M, Saeki S, Yao K, Hamaguchi A and Aratake S. (2008). PCT Int. Appl. WO2008/096820.

784. Kowa Company Limited. Miura T, Onogi K, Gomi N, Araki T, Tagashira J, Sekimoto R, Ishida R, Aoki H and Ohgiya T. (2011). PCT Int. Appl. WO2011/024468.

785. Hopkins CR, O'Neil SV, Laufersweiler MC, Wang Y, Pokross M, Mekel M, Evdokimov A, Walter R, Kontoyianni M, Petrey ME, Sabatakos G, Roesgen JT, Richardson E and Demuth TP. Bioorg Med Chem Lett (2006) 16:5659-5663.

786. Etgen GJ and Mantlo N. Curr Topic Med Chem (2003) 3:1649.

787. Rikimaru K, Wakabayashi T, Abe H, Tawaraishi T, Imoto H, Yonemori J, Hirose H, Murase K, Matsuo T, Matsumoto M, Nomura C, Tsuge H, Arimura N, Kawakami K, Sakamoto J, Funami M, Mol CD, Snell GP, Bragstad KA, Sang B-C, Dougan DR, Tanaka T, Katayama N, Horiguchi Y and Momose Y. Bioorg Med Chem (2012) 20:3332-3358.

Chapter 5

QSAR Studies on GPR119 Agonists

CHAPTER 5

QSAR STUDIES ON GPR119 AGONISTS

1. INTRODUCTION

Diabetes mellitus, characterized by high blood glucose levels, is a metabolic disorder and nearly 90% of all cases of diabetes belong to type 2 diabetes mellitus (T2DM). Type 2 diabetes mellitus (T2DM), highly associated with obesity, is due to insulin resistance and impaired pancreatic β-cell function. It is estimated in a study that nearly 350 million people suffering worldwide from diabetes [788] and it is supposed that this total will reach to 642 million by 2040 [789]. Impaired insulin secretion and insulin resistance causes hyperglycemia which in long-term increases risk of micro- and macro-vascular complications that may cause blindness, renal failure, diabetic foot disorders, heart attacks and strokes [790].

Multiple oral antidiabetic agents like sulfonylureas, meglitinides, biguanides, thiazolidinediones, α-glucosidase inhibitors and dipeptidyl-peptidase-4 (DPP-4) inhibitors have been used to cure T2DM but many patients failed to achieve glycemic control at desired level [791-794]. A large number of T2DM patients fail to reach desired HbA1c levels due to insufficient glycemic control [795]. The glucose-lowering effect of sodium-dependent glucose co-transporter 2 (SGLT2) inhibitor is devoid of hypo-glycemia or weight gain. Thus there is a need to develop a novel glucose-lowering drug to attain better glycemic control which protect pancreatic β-cells or exerts anti-obesity effects and devoid of causing hypoglycemia and cardiovascular side effects. In this direction, GPR119 [796-801] are the potential target for anti-diabetic therapy.

GPR119, a G-protein coupled receptor (GPCR), is expressed predominantly in the pancreatic β-cells and gastrointestinal L-cells. The identified endogenous agonists for the GPR119 receptor are oleoyl-lysophosphatidylcholine and oleoylethanolamide (OEA) [802, 803]. Glucose-dependent insulin secretion from pancreatic β-cells increases due to increased cellular cAMP levels on activation of the GPR119 receptor [804]. Release of

incretins like glucagon-like peptide 1 (GLP-1) and glucose-dependent insulinotropic polypeptide (GIP), from enteroendocrine cells are the results of the activation of the GPR119 receptor in the gut [805-807]. The stimulation of insulin secretion from β-cells in a glucose-dependent manner by GLP-1 and GIP protects β-cells against apoptosis [808, 809]. The activation of GPR119 is beneficial therapeutically for obesity [802, 810-812]. The GPR119 agonists demonstrated safety and tolerability in humans [813-816]. The investigations of several research groups [817, 818] on multiple small-molecule GPR119 agonists led to the development of clinical compounds which include APD668 [819], GSK1292263 [820] and MBX-2982 [821].

2. MODELING STUDIES

2.1. TRIAZOLOPYRIDINES AS hGPR119 AGONIST

The glucose-dependent dual mechanism of action of GPR119 agonists may improve glycemic control without inducing hypoglycemia. But, poor aqueous solubility of agonists causes low bioavailability, produces erratic assay results in *in vitro* studies and carries a high risk of not advancing due to potential toxicity which may not be recognized during preclinical studies [822, 823]. Therefore, an attempt to improve aqueous solubility of GPR119 agonist a novel series of triazolopyridine derivatives have been reported by Matsuda *et al*. [824]. These derivatives are based on 3*H*-[1,2,3]triazolo[4,5-*c*]pyridine scaffold and having variations at central spacer, left-hand aryl group and right-hand piperidine N-capping group.

The reported derivatives of triazolopyridine, having general structure shown in Figure 5.1, are the data set for present study.

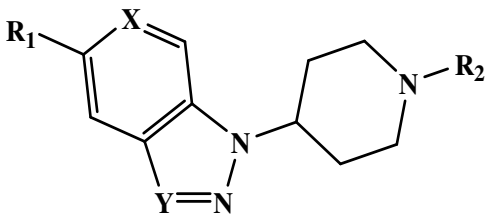
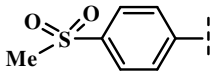
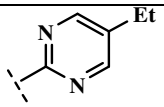
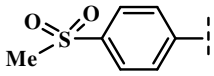
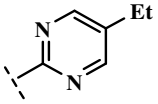
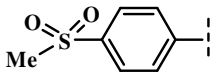
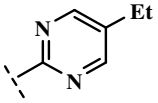
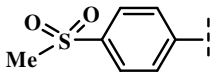
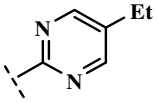
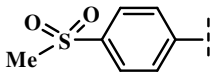
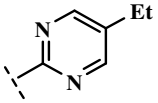
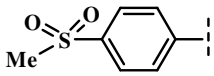
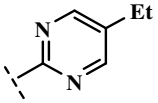
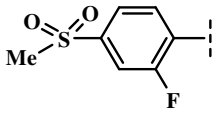
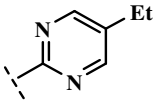
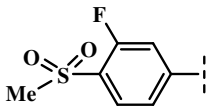
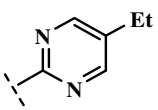
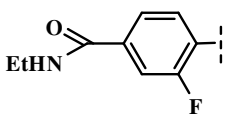
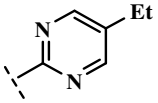
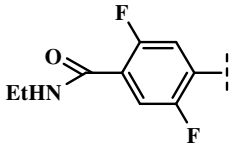
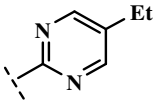
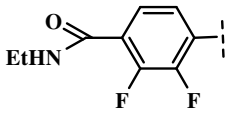
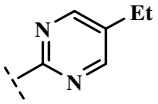
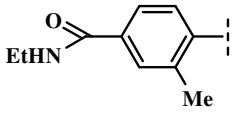
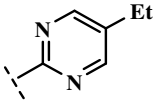
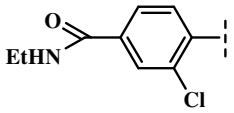
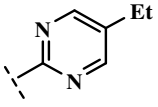
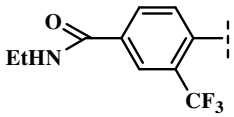
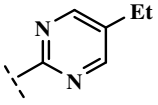
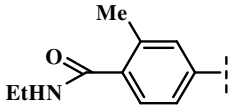
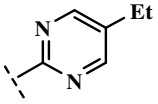
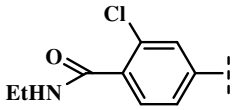
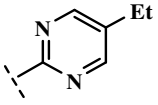
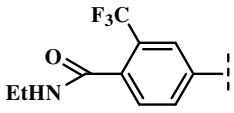
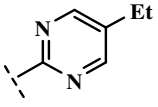
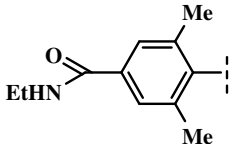
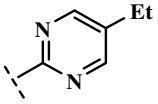
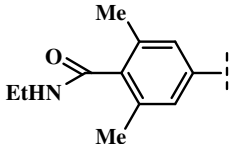
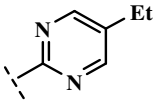
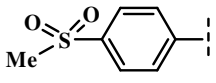
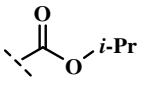


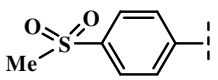
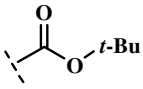
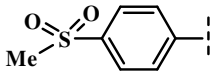
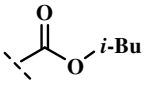
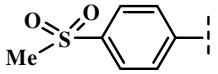
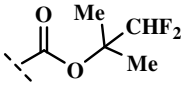
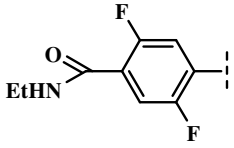
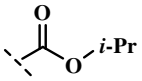
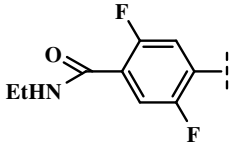
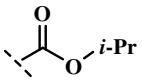
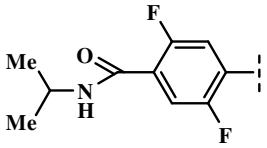
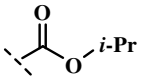
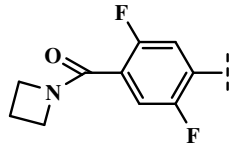
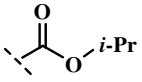
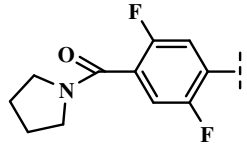
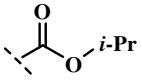
Figure 5.1: General structure of triazolopyridine derivatives.

These derivatives were evaluated for their agonistic activity against human GPR119 over-expressed in Flp-In-T-Rex-HEK293 cells by measuring changes in the cellular cAMP levels and were reported as EC_{50}. The reported activity on molar basis (as pEC_{50}) along with the structural variations of these analogues is shown in Table 5.1.

Table 5.1: Structural variations and reported hGPR119 agonistic activities of triazolopyridine derivatives.

Cpd.	R_1	X	Y	R_2	$pEC_{50}(M)$[a]
1		N	CH		7.89
2		CH	CH		7.85
3[b]		CMe	CH		7.15
4*		CH	CH		5.47
5[b]		CH	N		7.51
6		N	N		7.68
7[b]		N	N		8.15
8		N	N		8.05
9[b]		N	N		7.82

10		N	N		8.70
11		N	N		8.00
12		N	N		7.72
13		N	N		7.89
14		N	N		7.22
15		N	N		7.74
16		N	N		7.74
17		N	N		7.41
18		N	N		7.08
19		N	N		7.57
20[b]		N	CH		7.38

21[b]		N	CH		7.70
22[b]		N	CH		7.96
23		N	CH		7.64
24		N	CH		8.00
25		N	N		7.43
26		N	N		7.48
27		N	N		7.21
28		N	N		7.19

[a]EC_{50} (the the concentration of the test compound required to achieve 50% of the maximal response) on molar basis, taken from reference [824]; [b]Compound included in test set; *4-Methyl substituted indazole.

The data set was sub-divided into training set to develop models and test set to validate the models externally. The test set compounds which were selected using an in-house written randomization program, are also mentioned in Table 5.1.

2.1.1. RESULTS AND DISCUSSION

2.1.1.1. QSAR RESULTS

A total number of 492 descriptors, belonging to 0D- to 2D- modules, computed by Dragon software have been utilized to obtain most appropriate models describing the biological activity. For the purpose of modeling study, 07 (one fourth of total active) compounds have been included in the test set for the validation of the models derived from remaining 21 training set compounds. A total number of 99 relevant descriptors from 0D- to 2D- classes, which were obtained after the reduction of descriptor data set, have been subjected to CP-MLR analysis with default "filters" set in it. It has resulted in 04 models with test set $r^2 > 0.50$. These models have shared 10 descriptors among them. All these shared descriptors along with their brief meaning, average regression coefficients, and total incidence are listed in Table 5.2, which will serve as a measure of their estimate across these models.

Table 5.2: Identified descriptors[a] along with their class, average regression coefficient and incidence[b], in modeling the hGPR119 agonistic activities of triazolopyridines.

Descriptor class, average regression coefficient and (incidence)	
Topological descriptors (**TOPO**)	PW5, 2.374(3); IVDE, 1.128(2); LP1, -2.367(4)
Modified Burden Eigen values (**BCUT**)	BELm7, 0.980(1); BEHv8, -0.539(1)
2D autocorrelations (**2D-AUTO**)	MATS4m, 1.430(1); MATS2e, -0.722(1); MATS4e, 0.983 (1); MATS5e,1.223(1)
Functional group counts (**FUNC**)	nCp, -0.412(1)

[a]The descriptors are identified from the four parameter models for PPARγ binding activity transactivation activity emerged from CP-MLR protocol with filter-1 as 0.79, filter-2 as 2.0, filter-3 as 0.814 and filter-4 as $0.3 \leq q^2 \leq 1.0$ with a training set of 20 compounds. [b]The average regression coefficient of the descriptor corresponding to all models and the total number of its incidence. The arithmetic sign of the coefficient represents the actual sign of the regression coefficient in the models. **TOPO:** PW5, path/walk 5-Randic shape index; IVDE, mean information vertex degree equality; LP1; Lovasz-Pelikan index (leading eigenvalue); **BCUT:** BELm7, lowest eigenvalue n.7 of Burden matrix/weighted by atomic masses; BELm8, lowest eigenvalue n.8 of Burden matrix/weighted by atomic masses; BEHv8, highest eigenvalue n.8 of Burden matrix/weighted by van der Waals volumes; **2D-AUTO:** MATS4m, Moran autocorrelation of lag-4/ weighted by atomic masses; MATS2e, Moran autocorrelation of lag-2/ weighted by atomic Sanderson electronegativities; MATS4e, Moran autocorrelation of lag-4/ weighted by atomic Sanderson electronegativities; MATS5e, Moran autocorrelation of lag-5/ weighted by atomic Sanderson electronegativities; **FUNC:** nCp, number of total primary C(sp3).

The models in four descriptors, for the hGPR119 agonistic activities of triazolopyridines, emerged through CP-MLR are mentioned below.

pEC_{50} = 6.436 + 2.274(0.444)PW5 – 1.967(0.305)LP1
+ 1.430(0.464)MATS4m – 0.721(0.266)MATS2e
n = 21, r = 0.899, s = 0.297, F = 16.833, Q^2_{LOO} = 0.513, Q^2_{L5O} = 0.591,
r^2_{Test} = 0.532, FIT = 1.819, LOF = 0.175, AIC = 0.143 **(5.1)**

pEC_{50} = 6.587 + 2.349(0.536)PW5 + 0.916(0.454)IVDE
– 2.424(0.424)LP1 – 0.539(0.238)BEHv8
n = 21, r = 0.885, s = 0.315, F = 14.465, Q^2_{LOO} = 0.545, Q^2_{L5O} = 0.557,
r^2_{Test} = 0.661, FIT = 1.563, LOF = 0.198, AIC = 0.161 **(5.2)**

pEC_{50} = 6.865 –2.308(0.323)LP1 +0.980(0.365)BELm7 + 0.983(0.248)
MATS4e + 1.223(0.390) MATS5e
n = 21, r = 0.883, s = 0.317, F = 14.207, Q^2_{LOO} = 0.534, Q^2_{L5O} = 0.528,
r^2_{Test} = 0.579, FIT = 1.535, LOF = 0.200, AIC = 0.164 **(5.3)**

pEC_{50} = 6.474 + 2.500(0.556)PW5 + 1.340(0.474)IVDE
– 2.770(0.409)LP1 – 0.412(0.203)nCp
n = 21, r = 0.879, s = 0.323, F = 13.596, Q^2_{LOO} = 0.562, Q^2_{L5O} = 0.598,
r^2_{Test} = 0.522, FIT = 1.469, LOF = 0.207, AIC = 0.169 **(5.4)**

The participated descriptors, PW5, IVDE and LP1, in above models belong to topological class. It is apparent from the above mentioned equations that a higher value of path/walk 5-Randic shape index (PW5), and mean information vertex degree equality (IVDE) and a lower value of Lovasz-Pelikan index (LP1) would be helpful to elevate the agonistic activity. Modified Burden eigenvalue (BCUT) class descriptors BELm7 (lowest eigenvalue n.7 of Burden matrix/weighted by atomic masses) and BEHv8 (highest eigenvalue n.8 of Burden matrix/weighted by van der Waals volumes) have shown positive and negative contribution, respectively, to the activity suggesting a higher value of BELm7 and a lower value of BEHv8 beneficiary to the activity. Except MATS2e, all the participated 2D-autocoorelation

descriptors namely MATS4m, MATS4e and MATS5e contributed positively to the activity. Thus it may be inferred that a lower value of MATS2e (Moran autocorrelation of lag-2/weighted by atomic Sanderson electronegativities) and higher values of MATS4m (Moran autocorrelation of lag-4/weighted by atomic masses), MATS4e (Moran autocorrelation of lag-4/weighted by atomic Sanderson electronegativities) and MATS5e (Moran autocorrelation of lag-5/weighted by atomic Sanderson electronegativities) would be helpful for better activity.

Additionally, presence of higher number of total sp3 hybridized carbon atoms in a molecular structure (nCp, functional group class descriptor) would be detrimental to the activity. Nearly 81% variance in the observed activity has been accounted by these models. None of the CP-MLR identified model has shown any chance correlation in the randomization study (100 simulations per model). The values of Q^2 index, greater than a specified cutoff (0.5), hint that derived models are reasonable robust QSAR models. The pEC_{50} values of training set compounds calculated using Eqs. (5.1) to (5.4) and predicted from LOO procedure have been included in Table 5.3.

Table 5.3: Observed and modeled hGPR119 activity of triazolopyridines.

S. No.	pEC50(M)[a]								
	Obsd[b].	Eq. (5.1)		Eq. (5.2)		Eq. (5.3)		Eq. (5.4)	
		Calc.	Pred[c].	Calc.	Pred[c].	Calc.	Pred[c].	Calc.	Pred[c].
1	7.89	7.95	7.96	7.86	7.86	7.89	7.89	7.90	7.90
2	7.85	7.82	7.82	7.86	7.87	7.92	7.93	7.90	7.91
3[d]	7.15	6.72	-[d]	7.02	-[d]	6.79	-[d]	6.81	-[d]
4	5.47	5.67	6.18	5.71	6.22	5.91	6.44	5.70	6.19
5[d]	7.51	7.82	-[d]	7.86	-[d]	7.73	-[d]	7.90	-[d]
6	7.68	7.96	8.00	7.86	7.90	7.80	7.82	7.90	7.94
7[d]	8.15	8.16	-[d]	8.02	-[d]	8.10	-[d]	8.14	-[d]
8	8.05	7.80	7.77	7.80	7.76	8.30	8.38	7.92	7.90

9[d]	7.82	8.04	-[d]	7.88	-[d]	7.91	-[d]	7.81	-[d]
10	8.70	8.23	8.03	8.09	7.91	8.33	8.04	8.07	7.89
11	8.00	7.68	7.61	7.49	7.40	7.80	7.76	7.42	7.35
12	7.72	7.64	7.62	7.63	7.61	7.51	7.48	7.60	7.57
13	7.89	7.97	7.98	7.88	7.88	7.55	7.51	7.81	7.79
14	7.22	6.90	6.59	7.00	6.95	6.59	6.39	7.19	7.19
15	7.74	7.83	7.85	7.87	7.90	7.52	7.46	7.83	7.85
16	7.74	8.11	8.16	8.07	8.15	7.95	7.98	8.03	8.09
17	7.41	7.56	7.72	7.85	7.97	7.73	7.90	8.09	8.35
18	7.08	7.09	7.10	7.02	6.99	7.18	7.22	7.07	7.07
19	7.57	7.37	7.31	7.40	7.28	7.53	7.51	7.53	7.51
20[d]	7.38	7.19	-[d]	7.70	-[d]	7.39	-[d]	7.69	-[d]
21[d]	7.70	7.68	-[d]	7.73	-[d]	7.55	-[d]	7.62	-[d]
22[d]	7.96	7.93	-[d]	7.97	-[d]	7.62	-[d]	7.97	-[d]
23	7.64	7.26	6.69	7.43	6.98	7.85	7.96	7.52	7.30
24	8.00	7.69	7.66	7.97	7.97	7.70	7.65	7.85	7.82
25	7.43	7.71	7.74	7.97	8.12	7.45	7.46	7.85	7.96
26	7.48	7.80	7.88	7.68	7.72	7.57	7.61	7.36	7.31
27	7.21	7.36	7.37	7.26	7.27	7.19	7.19	7.22	7.22
28	7.19	7.57	7.65	7.24	7.25	7.70	7.77	7.20	7.20

[a]On molar basis; [b]Taken from ref. [824]; [c]Leave-one-out (LOO) procedure; [d]Compound included in test set.

The models (5.1) to (5.4) are validated with an external test set of 7 compounds mentioned in Table 5.1. The test set r^2 (r^2_{Test}) values greater than 0.5 of these models reflect that these models have satisfactory external validation capability. The predicted activity values of test set compounds are in tune to the observed ones and the same is mentioned in Table 5.3. The plot showing goodness of fit between observed and calculated activities for the training and test set compounds is given in Figure 5.2.

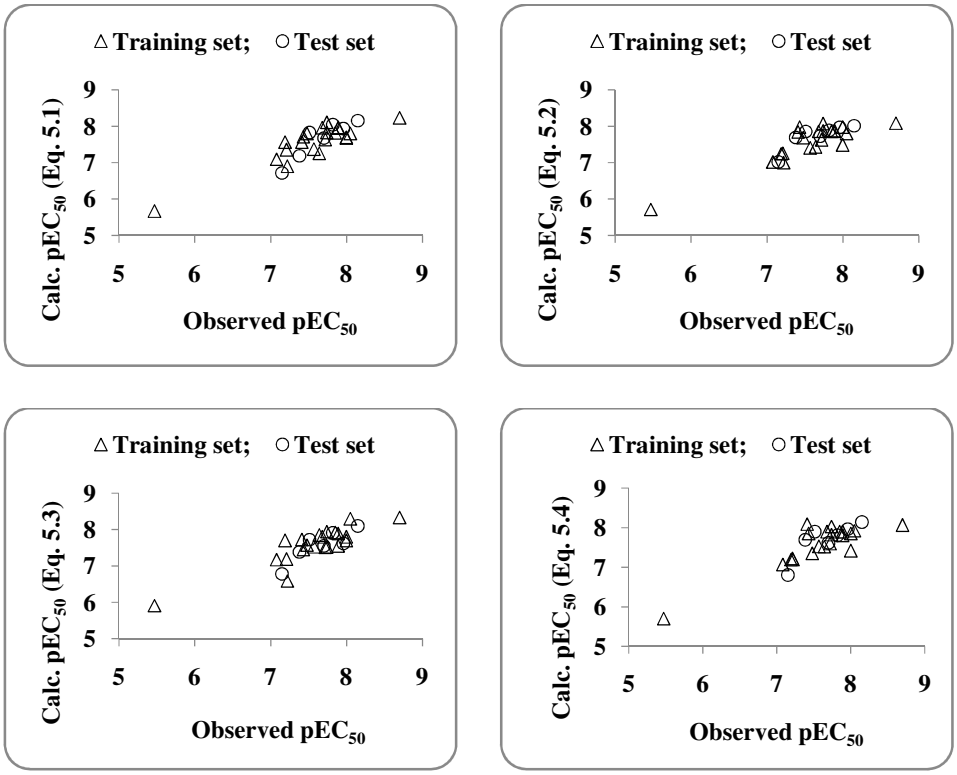


Figure 5.2: Plot of observed and calculated pEC_{50} values of training- and test-set compounds for hGPR119 agonistic activity of triazolopyridines.

2.1.1.2. Applicability domain (AD)

On analyzing the model AD in the Williams plot, shown in Figure 5.3, of the model based on the whole dataset (Table 5.4), it has appeared that none of the compounds were identified as an obvious outlier for the hGPR119 activity of triazolopyridines if the limit of normal values for the *Y* outliers (response outliers) was set as 3 (standard deviation) units. One compound listed in Table 5.1 at S. No. **4** found to have leverage (h) values greater than the threshold leverage (h*) suggesting this training set compound as chemically influential compound. For both the training-set and test-set, the suggested model matches the high quality parameters with good fitting power and the capability of assessing external data. Furthermore, all of the compounds were within the applicability domain of the proposed model and were evaluated correctly.

Table 5.4: Models derived for the whole data set (n = 28) for the hGPR119 agonistic activity in descriptors identified through CP-MLR.

Model	r	s	F	Eq.
pEC_{50} = 6.525 +2.074(0.353)PW5 –1.783(0.254)LP1+1.362(0.349)MATS4m –0.685(0.223)MATS2e	0.887	0.273	21.332	(5.1a)
pEC_{50} = 6.391 +2.536(0.444)PW5 +1.032(0.337)IVDE –2.370(0.329)LP1 –0.559(0.199)BEHv8	0.882	0.279	20.182	(5.2a)
pEC_{50} = 6.956 –2.210(0.270)LP1 +0.895(0.262)BELm7+0.924(0.204)MATS4e +1.140(0.263)MATS5e	0.874	0.288	18.632	(5.3a)
pEC_{50} = 6.326 +2.595(0.468)PW5 +1.418(0.371)IVDE –2.650(0.332)LP1 –0.380(0.169)nCp	0.869	0.293	17.803	(5.4a)

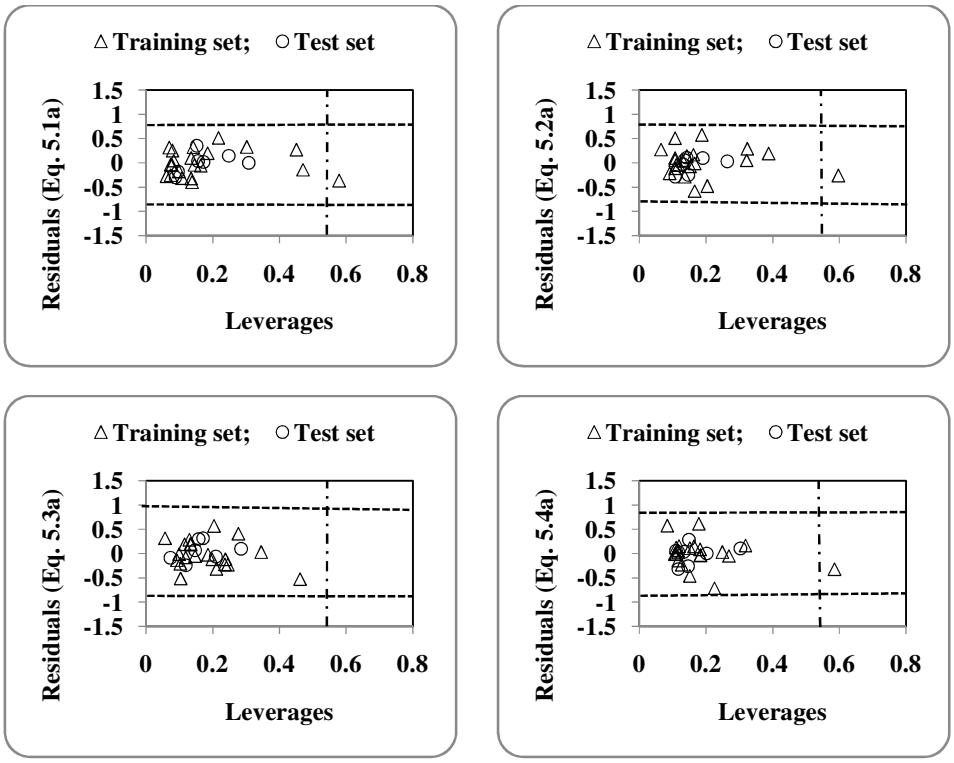


Figure 5.3: Williams plot for the training-set and test- set compounds for hGPR119 agonistic activity. The horizontal dotted line refers to the residual limit (±3×standard deviation) and the vertical dotted line represents threshold leverage h* (= 0.540).

2.1.2. CONCLUSIONS

QSAR study has been carried out on the hGPR119 agonistic activity of triazolopyridines in 0D- to 2D-Dragon descriptors. The descriptors identified in CP-MLR analysis have highlighted the role of molecular topology accounting features path/walk 5-Randic shape index (PW5), mean information vertex degree equality (IVDE), Lovasz-Pelikan index (LP1) in addition to atomic properties such as mass, van der Waals volume, and Sanderson electronegativity through weighted 2D autocorrelations (MATS4m, MATS2e, MATS4e and MATS5e) and modified Burden eigenvalues (BELm7 and BEHv8). Counts of total primary sp3 hybridized carbon atoms in a molecular structure (descriptor nCp) have also shown significance to optimize the hGPR119 agonistic activity. Applicability domain analysis revealed that the suggested model matches the high quality parameters with good fitting power and the capability of assessing external data and all of the compounds was within the applicability domain of the proposed model and were evaluated correctly.

2.2. INDOLE-BASED DERIVATIVES AS GPR119 AGONIST

Poor aqueous solubility of present GPR119 agonist causes low bioavailability has made a scope for further development of novel agonist. As an attempt to develop a novel GPR119 agonist for the treatment of T2DM, a series of indoline-based compounds has been reported by Sato *et al.* [825].The reported twenty six indole-based derivatives is considered as the data set for present study [825]. The general structure of these analogous is represented in Figure 5.4.

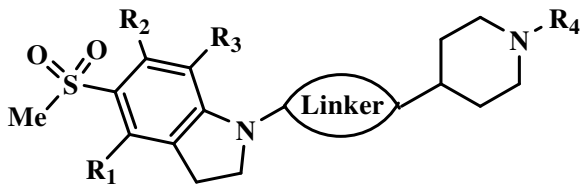
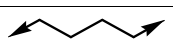
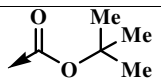
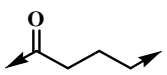
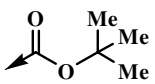
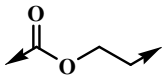
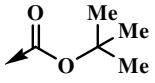
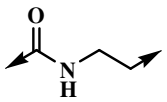
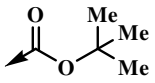
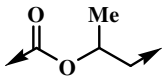
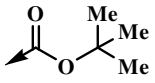
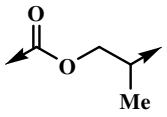
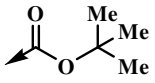
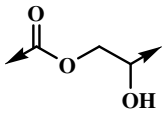
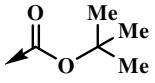
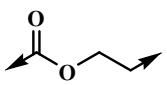
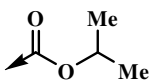
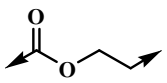
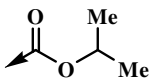
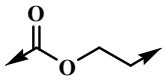
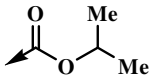
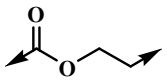
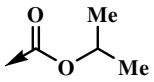
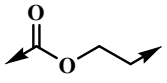
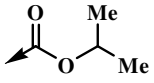
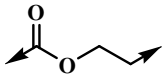
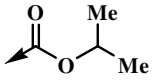
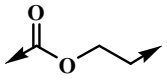
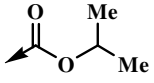
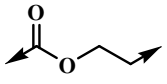
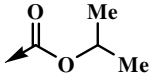
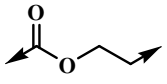
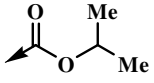
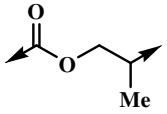
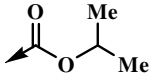
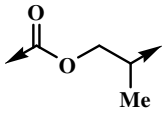
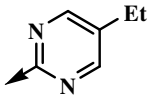
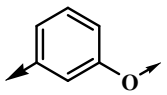
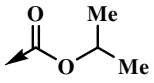
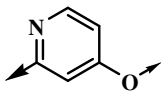
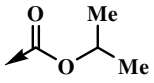
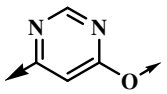
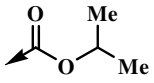
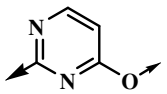
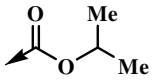
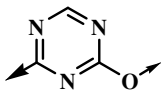
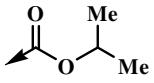
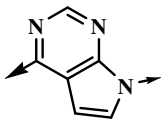
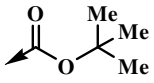


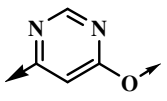
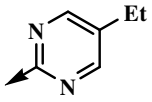
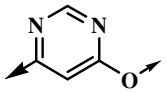
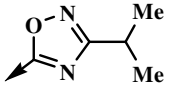
Figure 5.4: General structure of indole-based GPR119 agonists

These derivatives were evaluated for their GPR agonist activities in the reporter gene assay using CHO cells stably co-expressing cyclic AMP response element (CRE)–luciferase reporter gene (Promega) and GPR119 and were reported as EC_{50}. The reported activity on molar basis (as pEC_{50}) along with the structural variations of these analogues is shown in Table 5.5. The data set was sub-divided into training set to develop models and test set to validate the models externally. The test set compounds which were selected using an in-house written randomization program, are also mentioned in Table 5.5.

Table 5.5: Structural variations and reported GPR119 agonistic activities of indole-based derivatives.

Cpd.	R_1	R_2	R_3	Linker	R_4	$pEC_{50}(M)$[a]
1	H	H	H			5.85
2	H	H	H			6.92
3	H	H	H			7.17
4	H	H	H			5.60
5[b]	H	H	H			5.59
6	H	H	H			8.08
7	H	H	H			5.62
8	F	H	H			7.60
9[b]	Cl	H	H			6.28
10	Me	H	H			6.11

11	H	F	H			6.80
12	H	OMe	H			5.32
13	H	H	F			7.60
14	H	H	Cl			7.68
15	H	H	Me			7.36
16	H	H	OMe			5.77
17[b]	H	H	F			8.17
18	H	H	F			8.41
19	H	H	H			6.31
20[b]	H	H	H			7.06
21	H	H	H			7.77
22[c]	H	H	H			5.06
23	H	H	H			6.28
24[b]	H	H	H			7.80

25	H	H	H			7.96
26	H	H	H			8.11

[a]EC_{50} (the the concentration of the test compound required to achieve 50% of the maximal response) on molar basis, taken from reference [825]; [b]Compound included in test set; [c]Outlier compound.

2.2.1. RESULTS AND DISCUSSION

2.2.1.1. QSAR RESULTS

Primary observation of the data set revealed that one compound (S. No. 22, Table 5.5) does not fit in the trend of data set. Thus this compound has been excluded in deriving QSAR models. There are many reasons for their occurrence in QSAR studies; for example, chemicals might be acting by a mechanism different from that of the majority of the data points. It is also likely that outlier might be a result of a random experimental error that could be significant when analyzing a large data set. For the purpose of modeling study, 05 (one fifth of total active) compounds have been included in the test set for the validation of the models derived from remaining 20 training set compounds.

A total number of 485 descriptors, belonging to 0D- to 2D- modules, computed by Dragon software have been utilized to obtain most appropriate models describing the biological activity.122 relevant descriptors from 0D- to 2D- classes, which were obtained after the reduction of descriptor data set, have been subjected to CP-MLR analysis with default "filters" set in it. Statistical models in two and three descriptors have been explored to achieve the best relationship correlating GPR119 agonistic activity. All the models obtained in two descriptors were having the r^2_{Test} value less than 0.5. The obtained, all the three models, in three descriptors are given below through Eqs. (5.5) to (5.7). These models (with 122 descriptors) were identified in CP-MLR by successively incrementing the filter-3 with increasing number of descriptors (per equation). For this, the optimum *r*-bar value of the preceding

level model (=0.699, r-bar value of the two parameter model having highest r^2_{Test}) has been used as the new threshold of filter-3 for the next generation.

pEC_{50} = 7.980 + 3.264(0.936)GGI8 – 2.038(0.500)ATS7e
– 1.702(0.485)GATS1e
n = 20, r = 0.853, s = 0.559, F = 14.290, Q^2_{LOO} = 0.549, Q^2_{L5O} = 0.689
r^2_{Test} = 0.788, FIT = 1.478, LOF = 0.511, AIC = 0.470 **(5.5)**

pEC_{50} = 7.458 + 3.461(1.038)GGI8 – 2.003(0.536)ATS7e
– 3.475(1.173)Hy
n = 20, r = 0.830, s = 0.598, F = 11.846, Q^2_{LOO} = 0.531, Q^2_{L5O} = 0.510
r^2_{Test} = 0.700, FIT = 1.225, LOF = 0.584, AIC = 0.537 **(5.6)**

pEC_{50} = 8.151 + 2.585(0.891)LP1 – 1.993(0.444)BELp8
– 2.535(0.562)MATS7m
n = 20, r = 0.826, s = 0.604, F = 11.477, Q^2_{LOO} = 0.575, Q^2_{L5O} = 0.612
r^2_{Test} = 0.538, FIT = 1.187, LOF = 0.597, AIC = 0.548 **(5.7)**

The participated descriptors, ATS7e, MATS7m and GATS1e, in above models belong to 2D-AUTO class. It is apparent from the above mentioned equations that a lower values of Broto-Moreau autocorrelation of a topological structure of lag-7 weighted by atomic Sanderson electronegativities (ATS7e), Moran autocorrelation of lag-6 weighted by atomic masses and Geary autocorrelation of lag-1 weighted by atomic Sanderson electronegativities would be helpful to elevate the agonistic activity. The topological class descriptor LP1 (Lovasz-Pelikan index) and Galvez class descriptor (8^{th} order Galvez topological charge index, GGI8) shown positive correlation to the activity suggesting higher values of these as beneficial to the activity. The negative sign of correlation coefficient of modified Burden eigenvalue (BCUT) class descriptor BELp8 (lowest eigenvalue n.8 of Burden matrix/weighted by atomic polarizabilities and PROP class descriptor Hy (hydrophilic factor) advocated that a lower value of descriptor BELp8 and less hydrophilic factor or nature of molecule would be advantageous to the activity.

The three descriptor models could estimate nearly 73% variance in observed activity of the compounds. Considering the number of observation in the dataset, models with up to four descriptors were explored through CP-MLR and the result was 27 four-parameter models with test set $r^2 > 0.50$ sharing 40 descriptors among them. The shared descriptors along with their brief physical meaning, average regression coefficients, and total incidence are listed in Table 5.6, which will serve as a measure of their estimate across these models.

Table 5.6: Identified descriptors[a] along with their class, average regression coefficient and incidence[b], in modeling the GPR119 agonistic activities of indole-based compounds.

Descriptor class, average regression coefficient and (incidence)	
Constitutional descriptors (**CONST**)	AMW, 2.560(1); Mv, 2.958(1), -1.937(1); nAT, 3.015(1); nDB,0.634(1); nO,-1.343(1)
Topological descriptors (**TOPO**)	ZM2V, 2.406(1); MSD, -3.249(2); MAXDP, 1.889(1); X1A, -1.964(1); PW4, -1.570(1); AECC, -1.989(1); IC1, 3.103(1); LP1, 2.517(1)
Molecular walk counts (**MWC**)	MWC10, -0.860(1)
Modified Burden Eigen values (**BCUT**)	BEHm2, 2.038(1); BELm5, 2.459(7); BEHv5, 2.113(1); BELv1, -1.229(1); BELv2,-1.528(1); BELe8, -1.596(5); BELp8,-1.616(2)
Galvez topological charge indices (**GVZ**)	GGI5,-1.342(1);GGI8,3.396(18)
2D autocorrelations (**2D-AUTO**)	ATS6v, -3.354(1); ATS6e, -2.152(1); ATS7e, -2.466(9); ATS7p,-3.955(5); MATS6m,1.759(2); MATS7m, -2.651(1); MATS5v, -2.126(2); MATS6v, 1.977(2); MATS1e, 1.529(1); MATS3p, -2.721(4); GATS1e, -2.382(3); GATS4e, 2.088(1)
Atom centered fragments (**ACF**)	C-008, -1.072(1); C-029,1.136(2); H-046,2.491(4); H-050, -1.647(10)
Molecular properties (**PROP**)	Hy, -3.595(6)

[a]The descriptors are identified from the four parameter models for PPARγ binding activity transactivation activity emerged from CP-MLR protocol with filter-1 as 0.79, filter-2 as 2.0, filter-3 as 0.822 and filter-4 as $0.3 \leq q^2 \leq 1.0$ with a training set of 20 compounds. [b]The average regression coefficient of the descriptor corresponding to all models and the total number of its incidence. The arithmetic sign of the coefficient represents the actual sign of the regression coefficient in the models. **CONST**: AMW,average molecular weight; Mv,mean atomic van der Waals volume (scaled on Carbon atom); nAT,number of atoms; nDB,number of double bonds; nO,number of Oxygen atoms; **TOPO:** ZM2V, second Zagreb index by valence vertex degrees; MSD,mean square distance index (Balaban); MAXDP, maximal electrotopological positive

variation; X1A, average connectivity index chi-1; PW4, path/walk 4-Randic shape index; AECC, average eccentricity; IC1, information content index (neighborhood symmetry of 1-order) ; LP1, Lovasz-Pelikan index (leading eigenvalue); **MWC**: MWC10, molecular walk count of order 10; **BCUT:**BEHm2, highest eigenvalue n.2 of Burden matrix/weighted by atomic masses; BELm5, lowest eigenvalue n.5 of Burden matrix/weighted by atomic masses; BEHv5, highest eigenvalue n.5 of Burden matrix/weighted by van der Waals volumes; BELv1, lowest eigenvalue n.1 of Burden matrix/weighted by van der Waals volumes BELv2, lowest eigenvalue n.2 of Burden matrix/weighted by van der Waals volumes; BELe8, lowest eigenvalue n.8 of Burden matrix/weighted by atomic Sanderson electronegativities, BELp8, lowest eigenvalue n.8 of Burden matrix/weighted by atomic polarizabilities; **2D-AUTO:**ATS6v,Broto-Moreau autocorrelation of a topological structure - lag 6 / weighted by atomic van der Waals volumes; ATS6e,Broto-Moreau autocorrelation of a topological structure -lag 6/weighted by atomic Sanderson electronegativities; ATS7e,Broto-Moreau autocorrelation of a topological structure -lag 7/weighted by atomic Sanderson electronegativities; ATS7p,Broto-Moreau autocorrelation of a topological structure - lag 7 / weighted by atomic polarizabilities; MATS6m,Moran autocorrelation - lag 6 / weighted by atomic masses; MATS7m, Moran autocorrelation - lag 7 / weighted by atomic masses; MATS5v,Moran autocorrelation - lag 5 /weighted by atomic van der Waals volumes; MATS6v, Moran autocorrelation - lag 6 /weighted by atomic van der Waals volumes; MATS1e, Moran autocorrelation of lag-1/weighted by atomic Sanderson electronegativities; MATS3p; Moran autocorrelation of lag-3/weighted by atomic polarizabilities; GATS1e, Geary autocorrelation of lag-1/weighted by atomic Sanderson electronegativities; GATS4e, Geary autocorrelation of lag-4/weighted by atomic Sanderson electronegativities; GALVEZ: GGI5,topological charge index of order 5; GGI8, topological charge index of order 8; **ACF**:C-008,CHR2X; C-029, R--CX—X; H-046,H attached to C0(sp3) no X attached to next C; H-050, H attached to heteroatom; **PROP:** Hy, hydrophilic factor.

The selected models, in four parameters are given below.

$pEC_{50} = 6.142 + 2.359(0.639)BELm5 + 4.496(0.825)GGI8$

$– 3.661(0.828)ATS7p – 1.185(0.410)H\text{-}050$

$n = 20, r = 0.888, s = 0.509, F = 14.010, Q^2_{LOO} = 0.651, Q^2_{L5O} = 0.607$

$r^2_{Test} = 0.550$, FIT = 1.556, LOF = 0.541, AIC = 0.432 **(5.8)**

$pEC_{50} = 7.812 + 2.891(0.833)GGI8 – 2.094(0.502)ATS7e$

$– 1.072(0.372)C\text{-}008 – 1.453(0.409)H\text{-}050$

$n = 20, r = 0.888, s = 0.509, F = 13.994, Q^2_{LOO} = 0.624, Q^2_{L5O} = 0.703$

$r^2_{Test} = 0.745$, FIT = 1.554, LOF = 0.541, AIC = 0.433 **(5.9)**

$pEC_{50} = 6.316 + 2.163(0.922)GGI8 – 2.820(0.490)GATS1e$

$+ 2.088(0.589)GATS4e + 1.307(0.355)C\text{-}029$

$n = 20, r = 0.887, s = 0.511, F = 13.904, Q^2_{LOO} = 0.604, Q^2_{L5O} = 0.514$

$r^2_{Test} = 0.697$, FIT = 1.544, LOF = 0.544, AIC = 0.435 **(5.10)**

pEC_{50} = 8.553 + 2.516(0.760)LP1 – 2.066(0.380)BELp8
– 2.651(0.481)MATS7m – 2.459(0.927)Hy
n = 20, r = 0.885, s = 0.515, F = 13.610, Q^2_{LOO} = 0.699, Q^2_{L5O} = 0.637
r^2_{Test} = 0.540, FIT = 1.512, LOF = 0.553, AIC = 0.442 **(5.11)**

The newly appeared descriptors in above models C-008, C-029 and H-050 are from the atom centered fragment (ACF) class of descriptors. Descriptor BELm5 belong to BCUT class and the remaining two ATS7p and GATS4e are 2D-autocorrelations. The signs of regression coefficients of ACF descriptors suggested absence of CHR2X type fragment (descriptor C-008) and H attached to heteroatom (descriptor H-050) and presence of R--CX--X type structural fragment (descriptor C-029) beneficial to the activity.

Additionally, higher values of descriptors BELm5 (lowest eigenvalue n.5 of Burden matrix/weighted by atomic masses) and GATS4e (Geary autocorrelation of lag-4/weighted by atomic Sanderson electronegativities), and a lower value of descriptor ATS7p (Broto-Moreau autocorrelation of a topological structure - lag 7/weighted by atomic polarizabilities) would be advantageous to the agonistic activity.

Nearly 79% variance in the observed activity has been accounted by these models. None of the CP-MLR identified model has shown any chance correlation in the randomization study (100 simulations per model). The values of Q^2 index, greater than a specified cutoff (0.5), hint that derived models are reasonable robust QSAR models. The pEC_{50} values of training set compounds calculated using Eqs. (5.8) to (5.11) and predicted from LOO procedure have been included in Table 5.7.

The models (5.8) to (5.11) are validated externally with test set of 5 compounds mentioned in Table 5.5. The test set r^2 (r^2_{Test}) values greater than 0.5 of these models reflect that these models have satisfactory external validation capability. The predicted activity values of test set compounds are in tune to the observed ones and the same is mentioned in Table 5.7. The plot showing goodness of fit between observed and calculated activities for the training and test set compounds is given in Figure 5.4.

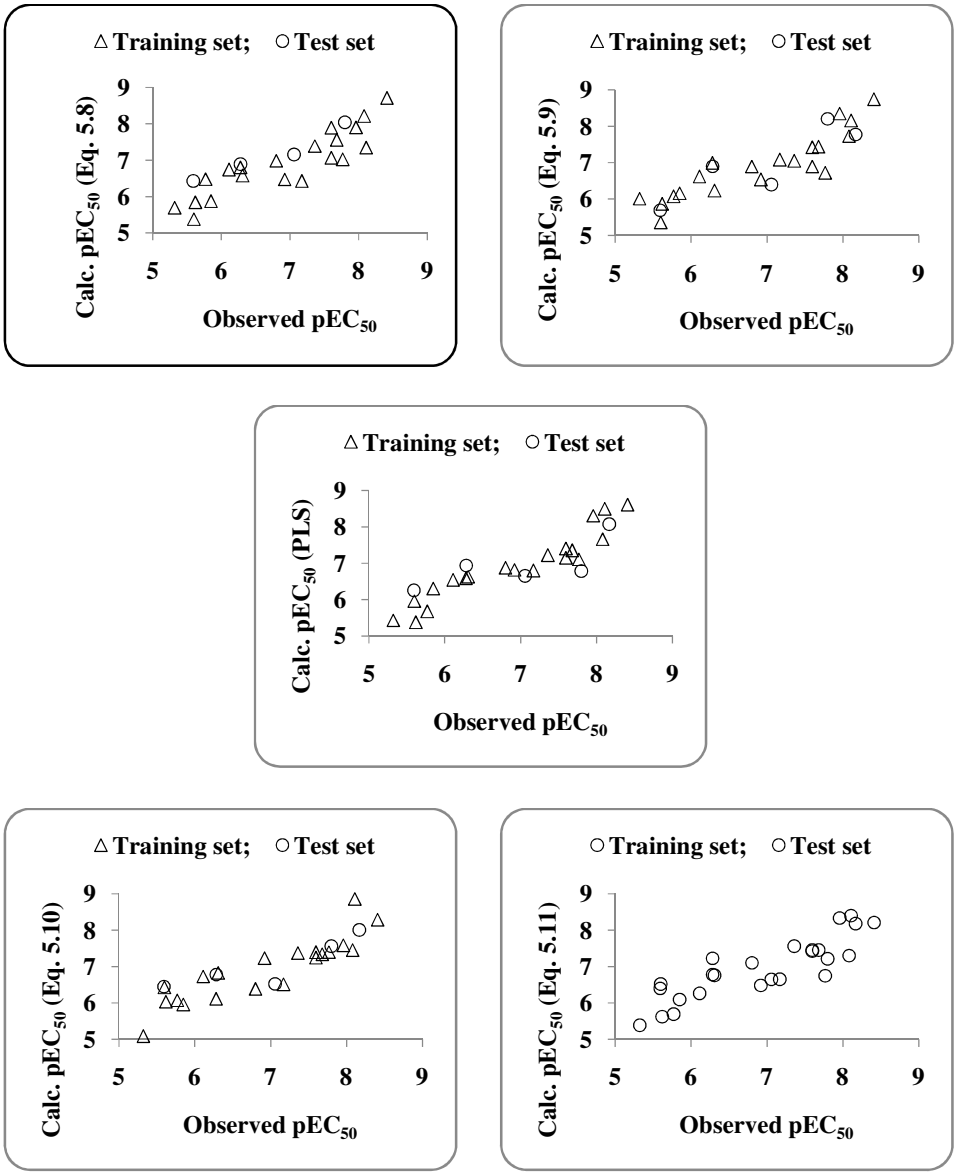


Figure 5.5: Plot of observed and calculated pEC_{50} values of training- and test-set compounds for indole-based GPR119 agonists.

Table 5.7: Observed and modeled GPR119 activity of indole-based agonists.

Cpd.	pEC50(M)[a]										
		Eq. (5.8)		Eq. (5.9)		Eq. (5.10)		Eq. (5.11)		PLS	
	Obsd[b]	Calc	Pre[c]	Calc	Pre[c]	Calc	Pre[c]	Calc	Pre[c]	Calc	Pre[c]
1	5.85	5.87	5.88	6.16	6.28	5.95	6.06	6.10	6.26	6.31	6.52
2	6.92	6.47	6.37	6.54	6.43	7.23	7.34	6.49	6.41	6.82	6.79
3	7.17	6.43	6.35	7.08	7.07	6.51	6.46	6.65	6.58	6.80	6.77

4	5.60	5.38	4.97	5.36	4.97	6.43	6.52	6.52	6.66	5.96	6.04
5[d]	5.59	6.42	-[d]	5.70	-[d]	6.46	-[d]	6.41	-[d]	6.26	-[d]
6	8.08	8.21	8.31	7.73	7.47	7.45	7.18	7.31	7.17	7.66	7.55
7	5.62	5.84	6.25	5.86	6.25	6.04	6.39	5.63	5.87	5.38	5.31
8	7.60	7.06	6.91	6.89	6.72	7.24	7.11	7.43	7.38	7.15	7.12
9[d]	6.28	6.89	-[d]	6.90	-[d]	6.79	-[d]	7.23	-[d]	6.94	-[d]
10	6.11	6.74	6.85	6.62	6.71	6.73	6.89	6.26	6.33	6.54	6.61
11	6.80	6.98	7.03	6.89	6.92	6.39	6.32	7.11	7.15	6.88	6.88
12	5.32	5.69	5.80	6.01	6.15	5.09	4.98	5.38	5.42	5.43	5.46
13	7.60	7.89	7.96	7.42	7.40	7.40	7.37	7.46	7.43	7.40	7.39
14	7.68	7.56	7.54	7.44	7.42	7.34	7.29	7.46	7.42	7.35	7.33
15	7.36	7.39	7.39	7.05	7.03	7.37	7.37	7.56	7.60	7.22	7.21
16	5.77	6.48	6.59	6.07	6.20	6.08	6.13	5.69	5.66	5.67	5.65
17[d]	8.17	9.12	-[d]	7.78	-[d]	8.01	-[d]	8.18	-[d]	8.08	-[d]
18	8.41	8.71	8.93	8.74	8.93	8.29	8.22	8.22	8.00	8.61	8.67
19	6.31	6.57	6.63	6.23	6.19	6.82	6.86	6.76	6.82	6.63	6.65
20[d]	7.06	7.15	-[d]	6.40	-[d]	6.53	-[d]	6.66	-[d]	6.65	0.00
21	7.77	7.02	6.95	6.72	6.44	7.40	7.17	6.75	6.66	7.11	7.05
22[e]	5.06	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
23	6.28	6.80	6.98	6.99	7.18	6.12	6.07	6.79	6.83	6.59	6.64
24[d]	7.80	8.04	-[d]	8.21	-[d]	7.56	-[d]	7.21	-[d]	6.78	-[d]
25	7.96	7.90	7.88	8.35	8.46	7.58	7.33	8.34	8.46	8.31	8.39
26	8.11	7.35	7.14	8.16	8.17	8.86	9.39	8.41	8.51	8.49	8.60

[a]On molar basis; [b]Taken from ref. [825]; [c]Leave-one-out (LOO) procedure; [d]Compound included in test set; [e]Oulier compound.

A partial least square (PLS) analysis has been carried out on these 13 descriptors, emerged in above mentioned models (5.8) to (5.11), to facilitate the development of a "single window" structure–activity model. For the purpose of PLS, the descriptors have been autoscaled (zero mean and unit SD) to give each one of them equal weight in the analysis. In the PLS cross-validation, two components are found to be the optimum for these 13 descriptors and they explained 87.79% variance in the activity. The MLR-like PLS coefficients of these 13 descriptors are given in Table 5.8.

The PLS analysis has suggested ATS7e as the most determining descriptor for modeling the agonistic activity of the compounds (descriptor S. No. 5 in Table 5.8; Figure 5.5). The other descriptors in decreasing order of significance are GGI8, BELp8, H-050, MATS7m, GATS1e, C-029, BELm5, GATS4e, Hy, C-008, ATS7p and LP1 and convey same inference in the PLS model as well. It is also observed that PLS model from the dataset devoid of these 13 descriptors is inferior in explaining the activity of the analogues.

Table 5.8: PLS and MLR-like PLS models from the 13 descriptors of four parameter CP-MLR models for GPR119 agonistic activities.

A: PLS equation									
PLS components					PLS coefficient (s.e.)[a]				
Component-1					-0.561(0.052)				
Component-2					0.141(0.045)				
Constant					6.916				
B: MLR-like PLS equation									
S. No.	Descriptor	MLR-like coef.[b]	(f.c.)[c]	Order	S. No.	Descriptor	MLR-like coef.[b]	(f.c.)[c]	Order
1	LP1	0.030	0.013	13	8	GATS1e	-0.217	-0.097	6
2	BELm5	0.163	0.073	8	9	GATS4e	0.133	0.059	9
3	BELp8	-0.230	-0.103	3	10	C-008	-0.100	-0.045	11
4	GGI8	0.235	0.105	2	11	C-029	0.199	0.088	7
5	ATS7e	-0.272	-0.121	1	12	H-050	-0.228	-0.102	4
6	ATS7p	-0.089	-0.040	12	13	Hy	-0.126	-0.056	10
7	MATS7m	-0.222	-0.099	5					
				Constant		6.558			
C: PLS regression statistics					Values				
n					20				
r					0.937				
s					0.363				
F					61.307				
FIT					5.108				
LOF					0.175				
AIC					0.178				
Q^2_{LOO}					0.828				
Q^2_{L5O}					0.808				
r^2_{Test}					0.538				

[a]Regression coefficient of PLS factor and its standard error. [b]Coefficients of MLR-like PLS equation in terms of descriptors for their original values;[c]f.c. is fraction contribution of regression coefficient, computed from the normalized regression coefficients obtained from the autoscaled (zero mean and unit s.d.) data.

For the sake of comparison, the plot showing goodness of fit between observed and calculated activities (through PLS analysis) for the training and test set compounds is also given in Figure 5.5. The fraction contribution of normalized regression coefficients of these descriptors to the activity is shown in Figure 5.6.

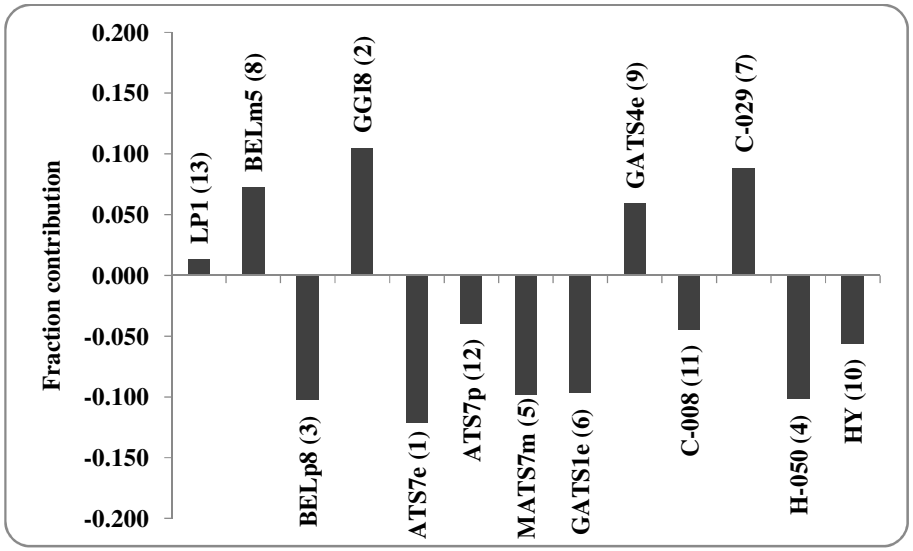


Figure 5.6: Plot of fraction contribution of MLR-like PLS coefficients (normalized) against 13 CP-MLR identified descriptors (Table 5.8) associated with GPR119 agonistic activity of indole-based derivatives.

2.2.1.2. APPLICABILITY DOMAIN (AD)

On analyzing the model AD in the Williams plot, shown in Figure 5.7, of the model based on the whole dataset (Table 5.9), it has appeared that one compound (S. No. **22**, Table 5.5) was identified as an obvious outlier for the GPR119 agonistic activity if the limit of normal values for the *Y* outliers (response outliers) was set as 2 (standard deviation) units. An outlier to a QSAR is identified normally by having a large standard residual activity and can indicate the limits of applicability of QSAR models. Two compounds, listed in Table 5.5 at S. No. **7** and **24** found to have leverage (h) values greater than the threshold leverage (h*) suggesting these compounds as chemically influential compound.

For both the training-set and test-set, the suggested model matches the high quality parameters with good fitting power and the capability of assessing external data. Furthermore, all of the compounds were within the applicability domain of the proposed model and were evaluated correctly.

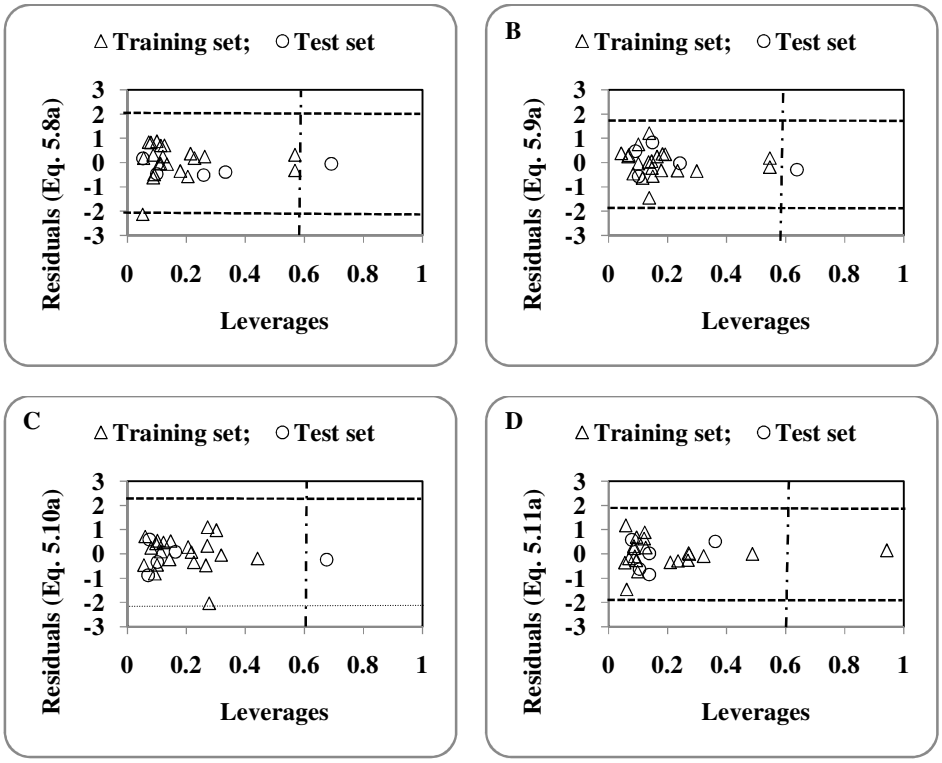


Figure 5.7: Williams plot for the training-set and test- set compounds for GPR119 agonistic activity. The horizontal dotted line refers to the residual limit (±2×standard deviation) and the vertical dotted line represents threshold leverage h* (= 0.580).

Table 5.9: Models derived for the whole data set (n = 26) for the GPR119 agonistic activity in descriptors identified through CP-MLR.

Model	r	s	F	Eq.
pEC_{50} = 6.237 + 1.680(0.679)BELm5 + 4.123(0.813)GGI8 – 3.263(0.841)ATS7p – 1.088(0.527)H-050	0.786	0.692	8.484	(5.8a)
pEC_{50} = 7.855 + 2.836(0.567)GGI8 – 2.118(0.446)ATS7e – 1.255(0.377)C-008 – 1.421(0.452)H-050	0.847	0.595	13.345	(5.9a)
pEC_{50} = 6.398 + 2.663(0.694)GGI8 –2.844(0.660)GATS1e +1.794(0.699)GATS4e + 0.628(0.425)C-029	0.765	0.721	7.414	(5.10a)
pEC_{50} = 8.523 + 2.745(0.698)LP1 –1.938(0.432)BELp8 – 2.882(0.536)MATS7m – 2.472(1.116)Hy	0.829	0.625	11.600	(5.11a)

2.2.2. CONCLUSIONS

QSAR study has been carried out on the GPR119 agonistic activity of indole-based derivatives in 0D- to 2D-Dragon descriptors. The derived QSAR models have revealed that the atomic Sandersons electronegativities weighted and charge accounting descriptors ATS7e, GATS1e, GATS4e and GGI8, molecular mass weighted descriptors, MATS7m and BELm5, and atomic polarizabilities weighted descriptors ATS7p and BELp8, and molecular topology accounting feature Lovasz-Pelikan index (LP1) played a pivotal role in rationalization of GPR119 agonistic activity of titled compounds. Hydrophilic factor (Hy) and certain structural fragments, such as CHR2X (C-008), R--CX--X (C-008) and H attached to heteroatom (H-050) are also predominant to explain GPR119 agonistic actions of indole-based derivatives. PLS analysis has also corroborated the dominance of CP-MLR identified descriptors. Applicability domain analysis revealed that the suggested model matches the high quality parameters with good fitting power and the capability of assessing external data and all of the compounds was within the applicability domain of the proposed model and were evaluated correctly.

REFERENCES

788. Danaei G, Finucane MM, Lu Y, Singh GM, Cowan MJ, Paciorek CJ, Lin JK, Farzadfar F, Khang Y-H, Stevens GA, Rao M, Ali MK, Riley LM, Robinson CA and Ezzati M. Lancet (2011) 378:31-40.

789. International Diabetes Federation, IDF diabetes atlas seventh edition 2015. http://www.diabetesatlas.org/.

790. Millatt LJ, Hanf R and Hum DW. Drug Discov World (2011) Summer:9-15.

791. Stein SA, Lamos EM and Davis SN. Expert Opin Drug Saf (2013) 12: 153-175.

792. Lovshin JA and Drucker DJ. Nat Rev Endocrinol (2009) 5:262-269.

793. Verspohl EJ. Pharmacol Therapeut (2009) 124:113-138.

794. Mikhail N. Expert Opin Invest Drugs (2008) 17:845-853.

795. Koro CE, Bowlin SJ, Bourgeois N and Fedder DO. Diabetes Care (2004) 27:17-20.

796. Shah U and Kowalski TJ. Curr Opin Drug Discov Develop (2010) 12:519-532.

797. Shah U. Curr Opin Drug Discov Develop (2009) 12:519-532.

798. Jones RM, Leonard JN, Buzard DJ and Lehmann, J. Expert Opin Ther Pat (2009) 19:1339-1359.

799. Fyfe MC, McCormack JG, Overton HA, Procter MJ and Reynet C. Expert Opin Drug Discov (2008) 3:403-413.

800. Overton HA, Fyfe MC and Reynet C. Br J Pharmacol (2008) 153:S76-S81.

801. Ohishi T and Yoshida S. Expert Opin Invest Drugs (2012) 21:321-328.

802. Overton HA, Babbs AJ, Doel SM, Fyfe MC, Gardner LS, Griffin G, Jackson HC, Procter MJ, Rasamison CM, Tang-Christensen M, Widdowson PS, Williams GM and Reynet C. Cell Metabol (2006) 3:167-175.

803. Soga T, Ohishi T, Matsui T, Saito T, MatsumotoM, Takasaki J, Matsumoto S-I, Kamohara M, Hiyama H, Yoshida S, Momose K, Ueda

Y, Matsushime H, Kobori M and Furuichi K. Biochem Biophys Res Commun (2005) 326:744-751.

804. Chu Z-L, Jones RM, He H, Carroll C, Gutierrez V, Lucman A, Moloney M, Gao H, Mondala H, Bagnol D, Unett D, Liang Y, Demarest K, Semple G, Behan DP and Leonardv J. Endocrinology (2007) 148:2601-2609.

805. Chu Z-L, Carroll C, Alfonso J, Gutierrez V, He H, Lucman A, Pedraza M, Mondala H, Gao H, Bagnol D, Chen R, Jones RM, Behan DP and Leonard J. Endocrinology (2008) 149:2038-2047.

806. Baggio LL and Drucker DJ. Ann Rev Med (2006) 57:265-281.

807. Aaboe K, Krarup T, Madsbad S and Holst JJ. Diabet Obes Metabol (2008) 10:994-1003.

808. Farilla L, Bulotta A, Hirshberg B, Calzi SL, Khoury N, Noushmehr H, Bertolotto C, Mario UD, Harlan DM, and Perfetti R. Endocrinology (2003) 144:5149-5158.

809. Trümper A, Trümper K, Trusheim H, Arnold R, Göke B and Hörsch D. Mol Endocrinol (2001) 15:1559-1570.

810. Fyfe, M, Mccormack J, Overton H, Procter M and Reynet C. American Diabetes Association 68thAnnual Scientific Sessions, San Francisco, CA, USA (2008) Abstract 297-OR.

811. Flock G, Holland D, Seino Y and Drucker DJ. Endocrinology (2011) 152:374-383.

812. Goodman ML, Dow J, Van Vliet AA, Pleszko A and Lockton JA. Diabetologia (2011) 54: Abstract OP32.

813. Roberts B, Gregoire FM, Karpf DB, Clemens E, Lavan B, Johnson J, Mcwherter CA, Martin R and Wilson M. American Diabetes Association 69th Annual Scientific Sessions, New Orleans, LA, USA (2009) Abstract 164-OR.

814. Roberts B, Karpf DB, Martin R, Lavan B, Wilson M and Mcwherter CA. American Diabetes Association 70th Annual Scientific Sessions, Orlando, FL, USA (2010) Abstract 603-P.

815. Goodman, ML, Dow J, Van Vliet AA, Hadi S, Karbiche D and Lockton

JA. Diabetes (2011) 60:Abstract A273.

816. Nunez DJ, Lewis EW, Swan S, Bush MA, Cannon C, Mcmullen SL, Collins DA and Feldman PL. American Diabetes Association 70th Annual Scientific Sessions, Orlando, FL, USA (2010) Abstract 80-OR.

817. Jones RM, Leonard JN, Buzard DJ, Lehmann J. Expert Opin Ther Pat (2009) 19:1339-1359.

818. Ritter K, Buning C, Halland N, Pöverlein C, J Med Chem. (2016) 59:3579-3592.

819. Semple G, Ren A, Fioravanti B, Pereira G, Calderon I, Choi K, Xiong Y, Shin Y-J, Gharbaoui T, Sage CR, Morgan M, Xing C, Chu Z-L, Leonard JN, Grottick AJ, Al-Shamma H, Liang Y, Demarest KT and Jones RM. Bioorg Med Chem Lett (2011) 21:3134-3141.

820. Polli JW, Hussey E, Bush M, Generaux G, Smith G, Collins D, McMullen S, Turner N and Nunez DJ. Xenobiotica (2013) 3:498-508.

821. http://clinicaltrials.gov/ct2/show/NCT01035879.

822. Matsuda D, Kobashi Y, Mikami A, Kawamura M, Shiozawa F, Kawabe K, Hamada M, Oda K, Nishimoto S, Kimura K, Miyoshi M, Takayama N, Kakinuma H and Ohtake N.Bioorg Med Chem Lett (2016) 26:3441-3446.

823. Alelyunas YW, Empfield JR, McCarthy D, Spreen RC, Bui K, Pelosi-Kilby L and Shen C. Bioorg Med Chem Lett (2010) 20:7312-7316.

824. Matsuda D, Kobashi Y, Mikami A, Kawamura M, Shiozawa F, Kawabe K, Hamada M, Nishimoto S, Kimura K, Miyoshi M, Takayama N, Kakinuma H and Ohtake N. Bioorg Med Chem (2017) 25:4339-4354.

825. Sato K, Sugimoto H, Rikimaru K, Imoto H, Kamaura H, Negoro N, Tsujihata Y, Miyashita H, Odani T and Murata T. Bioorg Med Chem (2014) 22:1649-1666.

Bibliography

BIBLIOGRAPHY

1. Brown AC and Fraser TR. Trans Roy Soc (Edinburgh) (1869) 25:151-203.
2. Craig PN. J Med Chem (1971) 14:680-684.
3. Topliss JG. J Med Chem (1972) 15:1006-1011.
4. Topliss JG. J Med Chem (1977) 20:463-469.
5. Darvas F. J Med Chem (1974) 17:799-804.
6. Bustard TM. J Med Chem (1974) 17:777-778.
7. Hansch C, Maloney PP, Fujita T and Muir RM. Nature (London) (1962) 194:178-180.
8. Hansch C and Fujita T. J Am Chem Soc (1964) 86:1616-1626.
9. Hansch C. Acc Chem Res (1969) 2:232-239.
10. Hansch C. "Drug Design", (Ariens EJ, Ed.) (1971) Vol- I. Academic Press, New York.
11. Hansch C. in "Structure-Activity Relationships" "International Encyclopedia of Pharmacology and Therapeutics" (Cavolto CJ, Ed.) Peragamon Press, Oxford.
12. Hansch C. J Med Chem (1976) 19:1-6.
13. Kubinyi H. Arzneim-Forsch (1976) 26:1991-1997.
14. Kubinyi H and Kehrhahn O-H. J Med Chem (1976) 19:1040-1049.
15. Kubinyi H. J Med Chem (1977) 20:625-629.
16. Kubinyi H. Arzneim-Forsch (1978) 28:598-601.
17. Free SM Jr and Wilson JW. J Med Chem (1964) 7:395-399.
18. Fujita T and Ban T. J Med Chem (1971) 14:148-152.
19. Nilsson NJ. "Leaning Machines" (1965) McGraw-Hill, New York.
20. Kowalski BR and Bender CF. J Am Chem Soc (1972) 94:5632-5639.
21. Kowalski BR and Bender CF. J Am Chem Soc (1973) 95:686-693.
22. Enslein K, Ralston A and Wife HS. "Mathematical Methods for Digital Computers" (1977) Vol-III. Wiley-Interscience, New York.
23. Hodes L, Hazard GF, Geran RI and Richman S. J Med Chem (1977) 20:469-475.

24. Weiner ML and Weiner PH. J Med Chem (1973) 16:655-661.
25. Cammarata A and Menon GK, J Med Chem (1976) 19:739-748.
26. Menon GK and Cammarata A. J Pharm Sci (1977) 66:304-313.
27. Hansch C, Unger SH and Forsythe AB. J Med Chem (1973) 16:1217-1222.
28. Snedecor GW and Cochran WG. "Statistical Methods" (1967) Iowa State University Press, Ames.
29. Draper NR and Smith H. "Applied Regression Analysis" (1966) Wiley, New York.
30. Wold S. Quant Struct-Act Relat (1991) 10:191-193.
31. Akaike H. "Second International Symposium on Information Theory", (Petrov B and Csáki F, Eds.) (1973) pp. 267-281, Budapest.
32. Akaike H. IEEE Trans Automat Contr (1974) AC-19:716-723.
33. Kubinyi H. Quant Struct-Act Relat (1994) 13:285-294.
34. Kubinyi H. Quant Struct-Act Relat (1994) 13:393-401.
35. Friedman J. Technical Report No. 102, (Nov. 1988; revised Aug. 1990) Laboratory for Computational Statistics, Department of Statistics, Stanford University.
36. Friedman J. The Annals of Statistics (1991) 19:1-67
37. Wu W, Walczak B, Massart DL, Heuerding S, Erni F, Last I R and Prebble KA. Chemom Intell Lab Syst (1996) 33:35-46.
38. Roy PP, Paul S, Mitra M and Roy K. Molecules (2009) 14:1660-1701.
39. Roy PP and Roy K. QSAR Comb Sci (2008) 27:302-313.
40. Rogers D and Hopfinger AJ. J Chem Inf Comput Sci (1994) 34:854-866.
41. **Roy K** and Paul S. *QSAR Comb Sci* **(2009) 28:**406-425.
42. So SS and Karplus M. J Med Chem (1997) 40: 4347-4359.
43. Afantitis A, Melagraki G, Sarimveis H, Koutentis PA, Markopoulos J and Markopoulou OI. Bioorg Med Chem (2006) 14:6686-6694.
44. Hansch C, Leo A and Elkins D. Chem Rev (1971) 71:525-616.
45. Green JP, Johnson CL and Kang S. Ann Rev Pharmacol (1974) 14:319-342.

46. Fujita T, Iwasa J and Hansch C. J Am Chem Soc (1964) 86:5175-5180.
47. Fujita T, Iwasa J and Hansch C. J Med Chem (1965) 8:150-153.
48. Hansch C, Leo A, Unger SH, Kim KH, Nikaitani D and Lien EJ. J Med Chem (1973) 16:1207-1216.
49. Hansch C and Leo A. "Substituent Constants for Correlation Analysis in Chemistry and Biology" (1979) Wiley-Interscience, New York.
50. Martin AJP. Biochem Symposia (1949) 3:4-15.
51. Haggerty WJ Jr and Murril EA. Res Develp (1974) 25(8):30-39.
52. Yamazaki M, Kakeya N, Morishita T, Kamada A and Aoki M. Chem Pharm Bull (Japan) (1970) 18:708-714.
53. Fujita T and Hansch C. J Med Chem (1967) 10:991-1000.
54. Hammett LP. "Physical Organic Chemistry" (1940) Ch 7, McGraw-Hill New York.
55. Jaffe HH. Chem Rev (1953) 53:191-261.
56. Ritchie CD and Sager WF. Prog Phys Org Chem (1964) 2:323- 400.
57. Scherrer RA and Howard SM. J Med Chem (1977), 20:53-58.
58. Swain CS and Lupton EC Jr. J Am Chem Soc (1968) 90:4328-4337.
59. Taft RW. "Steric Effects in Organic Chemistry" (1956) (Newman MS, Ed.) Wiley, New York.
60. Exner O. Coll Czech Chem Commun (1967) 32:1-23.
61. Moriguchi I. Chem Pharm Bull (Japan) (1975) 23:247-257.
62. Bondi A. J Phys Chem (1964) 68:441-452.
63. Moriguchi I, Kanada Y and Komatsu K. Chem Pharm Bull Japan (1976) 24:1799-1806.
64. Gupta SP and Prabhakar YS. J Sci Indus Res (1985) 44:189-198.
65. Kier LB, Hall LH, Murray WJ and Randic M. J Pharm Sci (1975) 64:1971-1974.
66. Kier LB, Murray WJ, Randic M and Hall LH. J Pharm Sci (1976) 65:1226-1230.
67. Murray WJ and Randic M. J Pharm Sci (1976) 65:1806-1809.

68. Murray WJ and Randic M. "Molecular Connectivity in Chemistry and Drug Research" (1976) Academic Press, New York.
69. Pullman B and Pullman A. "Quantum Biochemistry" (1963) Wiley-Interscience, New York.
70. Kier LB. "Molecular Orbital Theory in Drug Research" (1971) Academic Press, New York and London.
71. Fukui K, Yonezawa T and Nagata C. Bull Chem Soc (Japan) (1954) 27:423-427.
72. Fukui K, Yonezawa T and Shingu H. J Chem Phys, (1952) 20:722-725.
73. Dunn WJ. Eur J Med Chem-Chim Ther (1977) 12:109-113.
74. Randic M. J Am Chem Soc (1975) 97:6609-6615.
75. DRAGON software version 3.0. By Todeschini R, Consomni V and Mauri A, (2003) Milano, Italy. http://disat.unimib.it/chm/ Dragon.htm
76. ChemDraw Ultra 6.0 and Chem Bats 3D Ultra, Cambridge Soft Corporation, Cambridge, USA.
77. Todeschini R and Consomni V. “Handbook of molecular descriptors: methods and principals in medicinal chemistry”, (2000) Vol-11, (Mannhold R, Kubinyi H and Timmerman H, Eds.) Wiley, VCH.
78. Platt JR. J Chem Phys (1947) 15:419-420.
79. Hosoya H. Bull Chem Soc Jap (1971) 44:2332-2339.
80. Gutman I, Ruscic B, Trinajstic N and Wilcox CF Jr. J Chem Phys (1975) 62:3399-3405.
81. Entiger RC, Jackson DE and Snyder DA. Czech Math J (1976) 26:283-296.
82. Balaban AT. Chim Acta (1979) 53:355-375.
83. Balaban AT. Chem Phys Lett (1982) 89:399-404.
84. Balaban AT. Pure Appl Chem (1983) 55:199-206.
85. Barysz M, Jashari G, Lall RS, Srivastava AK and Trinajstic N. “Chemical Applications of Topology and Graph Theory” (1983) (King RB, Ed.) Elsevier, Amsterdam, The Netherlands.

86. Kier LB and Hall LH. “Structure-Activity Analysis, Research Studies”, (1986) Wiley, Chichester, UK.
87. Kier LB. Quant Struct-Act Relat (1987) 6:8-12.
88. Needham DE, Wei IC and Seybold PG. J Am Chem Soc (1988) 110:4186-4194.
89. Skorobogatov VA and Dobrynin AA. Commun Math Comput Chem (MATCH) (1988) 23:105-151.
90. Schultz HP. J Chem Inf Comput Sci (1989) 29:227-228.
91. Buckley F and Harary F. “Distance Matrix in Graphs”, (1990) Addison-Wesley, Redwood City, CA.
92. Schultz HP, Schultz EB and Schultz TP. J Chem Inf Comput Sci (1992) 32:69-72.
93. Plavsic D, Nikolic S, Trinajstic N and Mihalic Z. J Math Chem (1993) 12:235-250.
94. Gutman I, Yeh YN, Lee SL and Luo YL. Indian J Chem (1993) 32A:651-661.
95. Gutman I. J Chem Inf Comput Sci (1994) 34:1087-1089.
96. Voelkel A. Computers Chem (1994) 18:1-4.
97. Trinajstic N, Babic D, Nikolic S, Plavsic D, Amic D and Mihalic Z. J Chem Inf Comput Sci (1994) 34:368-376.
98. Konstantinova EV and Skorobogatov VA. J Chem Inf Comput Sci (1995) 35:472-478.
99. Pogliani L. J Phys Chem (1996) 100:18065-18077.
100. Diudea MV. J Chem Inf Comput Sci (1996) 36:535-540.
101. Randic M. J Chem Inf Comput Sci (1997) 37:1063-1071.
102. Sharma V, Goswami R and Madan AK. J Chem Inf Comput Sci (1997) 37:273-282.
103. Ren B. J Chem Inf Comput Sci (1999) 39:139-143.
104. Gupta S, Singh M and Madan AK. J Chem Inf Comput Sci (1999) 39:272-77.
105. Gramatica P, Corradi M and Consonni V. Chemosphere (2000) 41:763-777.

106. Rücker G and Rücker C. J Chem Inf Comput Sci (1993) 33:683-695.
107. Harary F. "Graph Theory", (1969) Addison-Wesley, Reading, MA.
108. Randic M, Brissey GM, Spencer RB and Wilkins CL. Computers Chem (1979) 3:5-13.
109. Randic M. Commun Math Comput Chem (MATCH) (1979) 7:3-60.
110. Randic M. J Chem Inf Compu tSci (1984) 24:164-175.
111. Randic M and Jurs PC. Quant Struct-Act Relat (1989) 8:39-48.
112. Balaban AT. "Numerical Modelling of Chemical Structures: Local Graph Invariants and Topological Indices in Graph Theory and Topology in Chemistry" (1987) (King RB and Rouvray DH, Eds.) Elsevier Amsterdam, The Netherlands.
113. Bonchev D, Mekenyan O and Trinajstic N. J Comput Chem (1981) 2:127-148.
114. Bonchev D. "Information Theoretic Indices for Characterization of Chemical Structures" (1983) Research Studies Press, Chichester, UK.
115. Koch R. Toxicol Environ Chem (1983) 6:87-96.
116. Shannon C and Weaver W. "The Mathematical Theory of Communication" (1949) University of Illinois Press, Urbana, ILL.
117. Magnuson VR, Harriss DK and Basak SC. "Topological Indices Based on Neighborhood Symmetry: Chemical and Biological Applications in Studies in Physical and Theoretical Chemistry" (1983), (King RB, Ed.), Elsevier, Amsterdam, The Netherlands.
118. Bertz SH. J Am Chem Soc (1981) 103:3599-3601.
119. Estrada E. J Chem Inf Comput Sci (1996) 36:844-849.
120. Estrada E. J Chem Inf Comput Sci (1998) 38:23-27.
121. Estrada E, Guevara N and Gutman I. J Chem Inf Comput Sci (1998) 38:428-431.
122. Markovic S and Gutman I. J Chem Inf Comput Sci (1999) 39:289-293.
123. Broto P, Moreau G and Vandycke C. Eur J Med Chem (1984) 19:66-70.
124. Moran PAP. Biometrika (1950) 37:17-23.

125. Geary RC. Incorp Statist (1954) 5:115-145.
126. Pearlman RS and Smith KM. “3D QSAR in Drug Design” (1998) Vol-II. (Kubinyi H, Folkers G and Martin YC, Eds.) Kluwer/ESCOM, Dordrecht, The Netherlands.
127. Burden FR. J Chem Inf Comput Sci (1989) 29:225-227.
128. Gálvez J, Garcìa R, Salabert MT and Soler R. J Chem Inf Comput Sci (1994) 34:520-525.
129. Lovasz L and Pelikan J. Period Math Hung (1973) 3:175-182.
130. Balaban AT, Ciubotariu D and Medeleanu M. J Chem Inf Comput Sci (1991) 31:517-523.
131. Randic M. New J Chem (1995) 19:781-791.
132. Randic M. J Chem Inf Comput Sci (1995) 35:373-382.
133. Randic M and Razinger M. J Chem Inf Comput Sci (1995) 35:594-606.
134. Jug K. Croat Chem Acta (1984) 57:941-953.
135. Bird CW. Tetrahedron (1986) 42:89-92.
136. Mekenyan O, Peitchev D, Bonchev D, Trinajstic N and Bangov IP. Arzneim Forsch (1986) 36:176-183.
137. Bogdanov B, Nikolic S and Trinajstic N. J Math Chem (1989) 3: 299-309.
138. Nikolic S, Trinajstic N, Mihalic Z and Carter S. Chem Phys Lett (1991) 179:21-28.
139. Arteca GA. “Molecular Shape Descriptors in Reviews in Computational Chemistry” (1991) Vol-9, (Lipkowitz KB and Boyd D, Eds.) VCH Publishers, New York.
140. Mihalic Z, Nikolic S and Trinajstic N. J Chem Inf Comput Sci (1992) 32:28-37.
141. Randic M, Kleiner AF and de Alba LM. J Chem Inf Comput Sci (1994) 34:277-286.
142. Krygowski TM, Ciesielski A, Bird CW and Kotschy A. J Chem Inf Comput Sci (1995) 35:203-210.
143. Bath PA, Poirrette AR, Willett P and Allen F. J Chem Inf Comput Sci (1995)

35:714-716.

144. Katritzky AR, Mu L, Lobanov VS and Karelson M. J Phys Chem (1996) 100:10400-10407.

145. Krygowski TM, Cyranski M, Ciesielski A, Swirska B and Leszczynski P. J Chem Inf Comput Sci (1996) 36:1135-1141.

146. Hemmer MC, Steinhauer V and Gasteiger J. Vibrat Spect (1999) 19:151-164.

147. Randic M and Krilov G. Int J Quant Chem (1999) 75:1017-1026.

148. Schuur JH, Selzer P and Gasteiger J. J Am Chem Soc (1996) 36:334-344.

149. Todeschini R, Lasagni M and Marengo E. J Chemom (1994) 8:263-273.

150. Todeschini R and Gramatica P. "3D QSAR in Drug Design" (1998) Vol.-II, (Kubinyi H, Folkers G and Martin YC, Eds.) Kluwer/ESCOM Dordrecht, The Netherlands.

151. Consonni V, Todeschini R and Pavan M. J Chem Inf Comput Sci (2002) 42:682-692.

152. Consonni V, Todeschini R, Pavan M and Gramatica P. J Chem Inf Comput Sci (2002) 42:693-705.

153. Viswanadhan VN, Ghose AK, Revankar GR and Robins RK. J Chem Inf Comput Sci (1989) 29:163-172.

154. Stanton DT and Jurs PC. Anal Chem (1990) 62:2323-2329.

155. Kaliszan R, Osmialowski K, Tomellini SA, Hsu S-H, Fazio SD and Hartwick RA. Chromatographia (1985) 20:705-708.

156. Osmialowski K, Halkiewicz J and Kaliszan R. J Chromat (1986) 361:63-69.

157. Clare BW and Supuran CT. J Pharm Sci (1994) 83:768-773.

158. Karelson M, Lobanov VS and KatritzkyAR. Chem Rev (1996) 96:1027-1043.

159. Prabhakar YS. QSAR Comb Sci (2003) 22:583-595.

160. So SS and Karplus M. J Med Chem (1997) 40:4347-4359.

161. Afantitis A, Melagraki G, Sarimveis H, Koutentis PA, Markopoulos J and Markopoulou OI. Bioorg Med Chem (2006) 14:6686-6694.

162. Wold S. Technometrics (1978) 20:397-405.
163. Gramatica P. QSAR Comb Sci (2007) 26:694-701.
164. Martens H and Naes T. "Multivariate Calibration" (1989) John Wiley and Sons Inc., Chichester.
165. Höskuldsson A. J Chemometrics (1988) 2:211-228.
166. Eriksson L, Johansson E, Kettaneh N and Wold S. "Multi- and megavariate data analysis: Principles and applications" (2001) Umeå: Umetrics.
167. Stahle L and Wold S. "Multivariate data analysis and experimental design in biomedical research. In: Progress in Medical Chemistry" (1988) Vol-25, (Eillis GP and West WB, Eds.) Elsevier Science Publishers, BV.
168. Geladi P and Kowalski B.Acta (1986) 185:1-17.
169. Curto JD and Pinto JC. Int Statist Rev (2007) 75:114-121.
170. Johnson RA and Wichern DW. "Applied Multivariate Statistical Analysis" (1992) Prentice-Hall, New York.
171. BUDYPLS (BUlk Data division bY PLS) Algorithm, "A novel network strategy of in situ PLS regression directed recursive data mining for chemo- and bio-informatics" by Prabhakar YS, Central Drug Research Institute, Lucknow 226 001, India.
172. Al-Lawati JA. Om Med J (2017) 32:177-179.
173. Xiao JB and Högger P. Curr Med Chem (2015) 22:23-38.
174. Cho NH, Shaw JE, Karuranga S, Huang Y, da Rocha Fernandes JD, Ohlrogge AW and Malanda B. Diabetes Res Clin Pract (2018) 138:271-281.
175. American Diabetes Association. Diabetes Care (2015) 40:S11-S24.
176. Volpe CMO, Villar-Delfino PH, Dos Anjos PMF and Nogueira-Machado JA. Cell Death Dis (2018) 9:119.
177. Harrison, L.C. In Clinical Immunology (Robert RR, William TS, Anthony JF, Thomas AF, Harry WS and Cornelia MW, Eds.). Elsevier (2019) pp. 957-966.
178. Goedeke L, Perry RJ and Shulman GI. Ann Rev Pharmacol Toxicol (2019) 59:65-87.

179. Chung YR, Ha KH, Kim HC, Park SJ, Lee K and Kim DJ. Diabetes Metab J (2019) 43:640-648.
180. American Diabetes Association. Diabetes Care (2010) 33:S62-S69.
181. Reddy VS, Sahay RK and Bhada SK. J Indian Acad Clin Med (2000) 1:245-251.
182. Moller DE. Nature (2001) 414:821-827.
183. McKinlay J and Marceau L. Lancet (2000) 356:757-761.
184. American Diabetes Association. Diabetes Care (2000) 23:77-79.
185. Definition, Diagnosis and Classification of Diabetes Mellitus and Its Complications, World Health Organization, Department of Non Communicable Disease Surveillance, Geneva, Switzerland, 1999.
186. Jovanovic L and Pettitt DJ. J Am Med Assoc (2001) 286:2516-2518.
187. Coustan D. In Diabetes in America. (Harris MI, Cowie CC, Stern MP, Boyko EJ, Reiber GE and Bennett PH, Eds.), 2nd edition (1995) pp. 703-717. US Government Printing Office, Washington, DC, USA.
188. Kronenberg HM, Melmed S, Polonsky KS and Larsen PR. Williams Textbook of Endocrinology, 11th edition (2008) Saunders, Philadelphia, Pa, USA.
189. Ovalle F, Vaughan III TB, Sohn JE and Gower B. Am J Med Sci (2008) 335:298-303.
190. Trout KK and Scheiner G. Endocrinol Rev (2008) pp. 27-29.
191. National Institute for Clinical Excellence, Technology Appraisal Guidance 60. Guidance on the Use of Patient Education Models for Diabetes, NICE, 2003.
192. Grundy SM. Nat Rev Drug Discov (2006) 5:295-309.
193. Haffner SM, Lehto S, Rönnemaa T, Pyörälä K, Laakso M. N Engl J Med (1998) 339:229-234.
194. Meier M and Hummel M. Vasc Health Risk Manag (2009) 5:859-871.
195. Best JH, Hoogwerf BJ, Herman WH, Pelletier EM, Smith DB, Wenten M and Hussein MA. Diabetes Care (2011) 34:90-95.
196. Bjorntorp P. Nutrition (1997) 13:795-803.
197. Overman RA, Yeh JY and Deal CL. Arthritis Care Res (2013) 65:294-298.

198. Seckl JR and Walker BR. Endocrinology (2001) 142:1371-1376.
199. Stewart PM and Krozowski ZS. Vitam Horm (1999) 57:249-324.
200. Chapman KE and Seckl JR. Neurochem Res (2008) 33:624-636.
201. Albiston AL, Obeyesekere VR, Smith RE and Krozowski ZS. Mol Cell Endocrinol (1994) 105:R11-R17.
202. Masuzaki H, et al. Science (2001) 294:166-170.
203. Masuzaki H, Yamamoto H, Kenyon CJ, Elmquist JK, Morton NM, Paterson JM, Shinyama H, Sharp MG, Fleming S, Mullins JJ, Seckl JR and Flier JS. J Clin Invest (2003) 112:83-90.
204. Morgan SA, McCabe EL, Gathercole LL, Hassan-Smith ZK, Larner DP, Bujalska IJ, Stewart PM, Tomlinson JW and Lavery GG. Proc Natl Acad Sci USA (2014) 111:E2482-E2491.
205. Wake DJ, Rask E, Livingstone DE, Söderberg S, Olsson T and Walker BR. J Clin Endocrinol Metab (2003) 88:3983-3988.
206. Tomlinson JW, Moore J, Cooper MS, Bujalska I, Shahmanesh M, Burt C, Strain A, Hewison M, Stewart PM. Endocrinology (2001) 142:1982-1989.
207. Hirosumi J, Tuncman G, Chang L, Görgün CZ, Uysal KT, Maeda K, Karin M and Hotamisligil GS. Nature (2002) 420:333-336.
208. Han MS, Jung DY, Morel C, Lakhani SA, Kim JK, Flavell RA and Davis RJ. Science (2013) 339:218-222.
209. Kaneto H Nakatani Y, Miyatsuka T, Kawamori D, Matsuoka TA, Matsuhisa M, Kajimoto Y, Ichijo H, Yamasaki Y and Hori M. Nat Med (2004) 10:1128-1132.
210. Ijaz A, Tejada T, Catanuto P, Xia X, Elliot SJ, Lenz O, Jauregui A, Saenz MO, Molano RD, Pileggi A, Ricordi C and Fornoni A. Kidney Int (2009) 75:381-388.
211. Murtagh J, McArdle E, Gilligan E, Thornton L, Furlong F and Martin F. J Cell Biol (2004) 166:133-143.
212. Qi AQ, Qiu J, Xiao L and Chen YZ. J Neurosci Res (2005) 80:510-517.
213. Furst R, Zahler S and Vollmar AM. Endocrinology (2008) 149:3635-3642.

214. Peng K, Pan Y, Li J, Khan Z, Fan M, Yin H, Tong C, Zhao Y, Liang G and Zheng C. Sci Rep (2016) 6:37160.
215. Alberti L, Girola A, Gilardini L, Conti A, Cattaldo S, Micheletto G and Invitti C. Int J Obes (London) (2007) 31:1826-1831.
216. Desbriere,R, Vuaroqueaux V, Achard V, Boullu-Ciocca S, Labuhn M, Dutour A and Grino M. Obesity (2006) 14:794-798.
217. Michailidou Z, Jensen MD, Dumesic DA, Chapman KE, Seckl JR, Walker BR and Morton NM. Obesity (2007) 15:1155-1163.
218. Munoz R, Carvajal C, Escalona A, Boza C, Perez G, Ibanez L and Fardella C. Obes Surg (2009) 19:764-770.
219. Paulsen SK, Pedersen SB, Fisker S, and Richelsen, B. Obesity (2007) 15:1954-1960.
220. Tomlinson JW, Finney J, Gay C, Hughes BA, Hughes SV and Stewart PM. Diabetes (2008) 57:2652-2660.
221. Wamil M and Seckl JR. Drug Discov Today (2007) 12:504-520.
222. Masuzaki H, Paterson J, Shinyama H, Morton NM, Mullins JJ, Seckl JR and Flier JS. Science (2001) 294:2166-2170.
223. Morton NM, Paterson JM, Masuzaki H, Holmes MC, Staels B, Fievet C, Walker BR, Flier JS, Mullins JJ and Seckl JR. Diabetes (2004) 53:931-938.
224. Edgerton DS, Basu R, Ramnanan CJ, Farmer TD, Neal D, Scott M, Jacobson P, Rizza RA and Cherrington AD. Am J Physiol Endocrinol Metab (2010) 298:E1019-E1026.
225. Winnick JJ, Ramnanan CJ, Saraswathi V, Roop J, Scott M, Jacobson P, Jung P, Basu R, Cherrington AD and Edgerton DS. Am J Physiol Endocrinol Metab (2013) 304:E747-E756.
226. Sundbom M, Kaiser C, Bjorkstrand E, Castro VM, Larsson C, Selen G, Nyhem CS and James SR. BMC Pharmacol (2008) 8:3.
227. Walker BR, Connacher AA, Lindsay RM, Webb DJ and Edwards CR. J Clin Endocrinol Metab (1995) 80:3155-3159.
228. Taylor A, Irwin N, McKillop AM, Flatt PR and Gault VA. Biol Chem (2008)

389:441-445.

229. Stewart PM, Wallace AM, Valentino R, Burt D, Shackleton CH and Edwards CR. Lancet (1987) 2:821-824.

230. Anderson A and Walker BR. Drugs (2013) 73:1385-1393.

231. Hermanowski-Vosatka A, Balkovec JM, Cheng K, Chen HY, Hernandez M, Koo GC, Le Grand CB, Li Z, Metzger JM, Mundt SS, Noonan H, Nunes CN, Olson SH, Pikounis B, Ren N, Robertson N, Schaeffer JM, Shah K, Springer MS, Strack AM, Strowski M, Wu K, Wu T, Xiao J, Zhang BB, Wright SD and Thieringer R. J Exp Med (2005) 202:517-527.

232. Rosenstock J, Banarer S, Fonseca VA, Inzucchi SE, Sun W, Yao W, Hollis G, Flores R, Levy R, Williams WV, Seckl JR and Huber R. Diabetes Care (2010) 33:1516-1522.

233. Feig PU, Shah S, Hermanowski-Vosatka A, Plotkin D, Springer MS, Donahue S, Thach C, Klein EJ, Lai E and Kaufman KD. Diabetes Obes Metab (2011) 13:498-504.

234. Shah S, Hermanowski-Vosatka A, Gibson K, Ruck RA, Jia G, Zhang J, Hwang PM, Ryan NW, Langdon RB and Feig PU. J Am Soc Hypertens (2011) 5:166-176.

235. Xu Z, Tice CM, Zhao W, Cacatian S, Ye YJ, Singh SB, Lindblom P, McKeever BM, Krosky PM, Kruk BA, Berbaum J, Harrison RK, Johnson JA, Bukhtiyarov Y, Panemangalore R, Scott BB, Zhao Y, Bruno JG, Togias J, Guo J, Guo R, Carroll PJ, McGeehan GM, Zhuang L, He W and Claremon DA. J Med Chem (2011) 54:6050-6062.

236. Hamilton BS, Himmelsbach F, Nar H, Schuler-Metz A, Krosky P, Guo J, Guo R, Meng S, Zhao Y, Lala DS. Zhuang L, Claremon DA and McGeehan GM. Eur J Pharmacol (2015) 746:50-55.

237. Ahren B. Nat Rev Drug Discov (2009) 8:369-385.

238. Mohler ML, He Y, Wu Z, Hwang DJ and Miller DD. Med Res Rev (2009) 29:125-195.

239. Fredriksson R, Hoglund PJ, Gloriam DE, Lagerstrom MC and Schioth HB.

FEBS Lett (2003) 554:381-388.

240. Shah U and Kowalski TJ. Vitam Horm (2010) 84:415-448.

241. Bonini JA and Borowsky BE. (2001) p. US6221660.

242. Bonini JA and Borowsky BE. (2002) p. US6468756.

243. Jones RM. (2004) p. WO2004065380.

244. Takeda S, Kadowaki S, Haga T, Takaesu H and Mitaku S. FEBS Lett (2002) 520:97-101.

245. Davey J. Expert Opin Ther Targets (2004) 8:165-170.

246. Griffin G. (2006) p.US7083933.

247. Fyfe MC. Expert Opin Drug Discov (2008) 3:403-413.

248. Fu J. Nature (2003) 425:90-93.

249. Ohishi T. (2003) p. EP1338651.

250. Soga T, Ohishi T, Matsui T, Saito T, Matsumoto M, Takasaki J, Matsumoto S, Kamohara M, Hiyama H, Yoshida S, Momose K, Ueda Y, Matsushime H, Kobori M and Furuichi K. Biochem Biophys Res Commun (2005) 326:744-751.

251. Chu ZL, Jones RM, He H and Carroll C. Endocrinology (2007) 148:2601-2609.

252. Ohishi T and Yoshida S. Expert Opin Investig Drugs (2012) 21:321-328.

253. Chu Z-L. (2006) p. WO2006076231

254. Drucker DJ, Jin T, Asa SL, Young TA and Brubaker PL. Mol Endocrinol (1994) 8:1646-1655.

255. Dhayal S and Morgan NG. Drug News Perspect (2010) 23:418-424.

256. Overton HA, Babbs AJ, Doel SM, Fyfe MC, Gardner LS, Griffin G, Jackson HC, Procter MJ, Rasamison CM, Tang-Christensen M, Widdowson PS, Williams GM and Reynet C. Cell Metab (2006) 3:167-175.

257. Overton HA, Fyfe MC, Reynet C. Br J Pharmacol (2008) 153 (Suppl 1):S76-S81.

258. Rodriguez de Fonseca F, Navarro M, Gomez R, Escuredo L, Nava F, Fu J, Murillo-Rodriguez E, Giuffrida A, LoVerme J, Gaetani S, Kathuria S, Gall C

and Piomelli D. Nature (2001) 414:209-212.

259. Yang Y, Chen M, Georgeson KE and Harmon CM. Am J Physiol Regul Integr Comp Physiol (2007) 292:R235-241.

260. Proulx K, Cota D, Castaneda TR, Tschop MH, D'Alessio DA, Tso P, Woods SC and Seeley RJ. Am J Physiol Regul Integr Comp Physiol (2005) 89:R729-737.

261. Campbell JE and Drucker DJ. Cell Metab (2013) 17:819-37.

262. International Textbook of Diabetes Mellitus. (DeFronzo RA, Ferrannini E, Zimmet P and Alberti KGMM. Eds.). (2015) Oxford:Wiley-Blackwell Ltd.

263. Li Y, Li L and Hölscher C. Rev Neurosci (2016) 27:689-711.

264. João AL, Reis F and Fernandes R. Obes Rev (2016) 17:553-72.

265. Holst JJ, Vilsbøll T and Deacon CF. Mol Cell Endocrinol (2009) 297:127-136.

266. Watterson KR, Hudson BD, Ulven T and Milligan G. Front Endocrinol (2014) 5:137.

267. Poulsen Ll, Siersbæk M and Mandrup S. Semin Cell Dev Biol (2012) 23:631-639.

268. Kebede MA, Alquier T, Latour MG and Poitout V. Diabetes Obes Metab (2009) 11:10-20.

269. Fyfe MC, McCormack JG, Overton HA, Procter MJ and Reynet C. Expert Opin Drug Discov (2008) 3:403-413.

270. Chu ZL, Carroll C, Alfonso J, Gutierrez V, He H, Lucman A, Pedraza M, Mondala H, Gao H, Bagnol D, Chen R, Jones RM, Behan DP and Leonard J. Endocrinology (2008) 149:2038-2047.

271. Hothersall JD, Bussey CE, Brown AJ, Scott JS, Dale I and Rawlins P. Eur J Pharmacol (2015) 762:430-442.

272. Ritter K, Buning C, Halland N, Pöverlein C and Schwink L. J Med Chem (2016) 59:3579-3592.

273. Nunez DJ, Bush MA, Collins DA, McMullen SL, Gillmor D, Apseloff G, Atiee G, Corsino L, Morrow L and Feldman PL. PLoS One (2014) 9:e92494.

274. Saraiva FK and Sposito AC. Cardiovasc Diabetol (2014) 13:142.
275. Candeias EM, Sebastião IC, Cardoso SM, Correia SC, Carvalho CI, Plácido AI, Santos MS, Oliveira CR, Moreira PI and Duarte AI. World J Diabetes (2015) 6:807-827.
276. Jones RM, Leonard JN, Buzard DJ and Lehmann J. Expert Opin Ther Pat (2009) 19:1339-1359.
277. Ritter K, Buning C, Halland N, Pöverlein C and Schwink L. J Med Chem (2016) 59:3579-3592.
278. Semple G, Ren A, Fioravanti B, Pereira G, Calderon I, Choi K, Xiong Y, Shin YJ, Gharbaoui T, Sage CR, Morgan M, Xing C, Chu ZL, Leonard JN, Grottick AJ, Al-Shamma H, Liang Y, Demarest KT and Jones RM. Bioorg Med Chem Lett (2011) 21:3134-3141.
279. Polli JW, Hussey E, Bush M, Generaux G, Smith G, Collins D, McMullen S, Turner N and Nunez DJ. Xenobiotica. (2013) 43:498-508.
280. http://clinicaltrials.gov/ct2/show/NCT01035879.
281. Matsuda D, Kobashi Y, Mikami A, Kawamura M, Shiozawa F, Kawabe K, Hamada M, Oda K, Nishimoto S, Kimura K, Miyoshi M, Takayama N, Kakinuma H and Ohtake N. Bioorg Med Chem Lett (2016) 26:3441-3446.
282. Alelyunas YW, Empfield JR, McCarthy D, Spreen RC, Bui K, Pelosi-Kilby L and Shen C. Bioorg Med Chem Lett (2010) 20:7312-7316.
283. Makishima M, Okamoto AY, Repa JJ, Tu H, Learned RM, Luk A, Hull MV, Lustig KD, Mangelsdorf DJ and Shan B. Science (1999) 284:1362-1365.
284. Lefebvre P, Cariou B, Lien F, Kuipers F and Staels B. Physiol Rev (2009) 89:147-191.
285. Ahlberg J, Angelin B, Bjorkhem I and Einarsson K. Gastroenterology (1977) 73:1377-1382.
286. Guo C, Xie S, Chi Z, Zhang J, Liu Y, Zhang L, Zheng M, Zhang X, Xia D, Ke Y, Lu L and Wang D. Immunity (2016) 45:802-816.
287. Benoit B, Meugnier E, Castelli M, Chanon S, Vieille-Marchiset A, Durand C, Bendridi N, Pesenti S, Monternier PA, Durieux AC, Freyssenet D, Rieusset J,

Lefai E, Vidal H and Ruzzin J. Nat Med (2017) 23:990-996.

288. Li T and Chiang JY. Pharmacol Rev (2014) 66:948-983.

289. Shapiro H, Kolodziejczyk AA, Halstuch D and Elinav E. J Exp Med (2018) 215:383-396.

290. Katsuma S, Hirasawa A and Tsujimoto G. Biochem Biophys Res Commun (2005) 329:386-390.

291. Kawamata Y, Fujii R, Hosoya M, Harada M, Yoshida H, Miwa M, Fukusumi S, Habata Y, Itoh T, Shintani Y, Hinuma S, Fujisawa Y and Fujino M. J Biol Chem (2003) 278:9435-9440.

292. Wang YD, Chen WD and Huang W. Histol Histopathol (2008) 23:621-627.

293. Sato H, Genet C, Strehle A, Thomas C, Lobstein A, Wagner A, Mioskowski C, Auwerx J and Saladin R. Biochem Biophys Res Commun (2007) 362:793-798.

294. Lo SH, Li Y, Cheng KC, Niu CS, Cheng JT and Niu HS. Naunyn Schmiedebergs Arch Pharmacol (2017) 390:1097-1104.

295. Wang LY, Cheng KC, Li Y, Niu CS, Cheng JT and Niu HS. Biomed Pharmacother (2017) 95:599-604.

296. Li Y, Cheng KC, Niu CS, Lo SH, Cheng JT and Niu HS. Drug Des Devel Ther (2017) 11:1127-1134.

297. Baggio LL and Drucker DJ. Gastroenterology (2007) 132:2131-2157.

298. Trabelsi MS, Daoudi M, Prawitt J, Ducastel S, Touche V, Sayin SI, Perino A, Brighton CA, Sebti Y, Kluza J, Briand O, Dehondt H, Vallez E, Dorchies E, Baud G, Spinelli V, Hennuyer N, Caron S and Bantubungi K. Nat Commun (2015) 6:7629.

299. Kumar DP, Rajagopal S, Mahavadi S, Mirshahi F, Grider JR, Murthy KS and Sanyal AJ. Biochem Biophys Res Commun (2012) 427:600-605.

300. Renga B, Mencarelli A, Vavassori P, Brancaleone V and Fiorucci S. Biochim Biophys Acta (2010) 1802:363-372.

301. Kumar DP, Asgharpour A, Mirshahi F, Park SH, Liu S, Imai Y, Nadler JL, Grider JR, Murthy KS and Sanyal AJ. J Biol Chem (2016) 291:6626-6640.

302. Alejandro EU, Gregg B, Blandino-Rosano M, Cras-Meneur C, Bernal-Mizrachi E. Mol Aspects Med (2015) 42:19-41.
303. Zarrinpar A and Loomba R. Aliment Pharmacol Ther (2012) 36:909-921.
304. Watanabe M, Houten SM, Mataki C, Christoffolete MA, Kim BW, Sato H, Messaddeq N, Harney JW, Ezaki O, Kodama T, Schoonjans K, Bianco AC and Auwerx J. Nature (2006) 439:484-489.
305. Perino A, Pols TW, Nomura M, Stein S, Pellicciari R and Schoonjans K. J Clin Invest (2014) 124:5424-5436.
306. Worthmann A, John C, Ruhlemann MC, Baguhl M, Heinsen FA, Schaltenberg N, Heine M, Schlein C, Evangelakos I, Mineo C, Fischer M, Dandri M, Kremoser C, Scheja L, Franke A, Shaul PW and Heeren J. Nat Med (2017) 23:839-849.
307. Wang XX, Edelstein MH, Gafter U, Qiu L, Luo Y, Dobrinskikh E, Lucia S, Adorini L, D'Agati VD, Levi J, Rosenberg A, Kopp JB, Gius DR, Saleem MA and Levi M. J Am Soc Nephrol (2016), 27:1362-1378.
308. Pierre JF, Martinez KB, Ye H, Nadimpalli A, Morton TC, Yang J, Wang Q, Patno N, Chang EB and Yin DP. Am J Physiol Gastrointest Liver Physiol (2016) 311:G286-G304.
309. Briere DA, Ruan X, Cheng CC, Siesky AM, Fitch TE, Dominguez C, Sanfeliciano SG, Montero C, Suen CS, Xu Y, Coskun T and Michael MD. PLoS One (2015) 10:e0136873.
310. Bhat BG, LeeAnne M, Dean PP, Judit M, Shufang Z, et al. J Pharmacol Clin Res (2017) 3:555623.
311. Lasalle M, Hoguet V, Hennuyer N, Leroux F, Piveteau C, Belloy L, Lestavel S, Vallez E, Dorchies E, Duplan I, Sevin E, Culot M, Gosselet F, Boulahjar R, Herledan A, Staels B, Deprez B, Tailleux A and Charton J. J Med Chem (2017) 60:4185-4211.
312. Murphy MP. Biochem J (2009) 417:1-13.
313. McBride HM, Neuspiel M and Wasiak S. Curr Biol. (2006) 16:R551-560.
314. Balaban RS, Nemoto S and Finkel T. Cell (2005) 120:483-95.

315. Raha S and Robinson BH. Trends Biochem Sci (2000) 25:502-508.
316. Garg AK and Aggarwal BB. Mol Immunol (2002) 39:509-517.
317. Janssen-Heininger YM, Poynter ME and Baeuerle PA. Free Radic Biol Med (2000) 28:1317-1327.
318. Fontana L, Partridge L and Longo VD. Science (2010) 328:321-326.
319. Colman RJ, Anderson RM, Johnson SC, Kastman EK, Kosmatka KJ, Beasley TM, Allison DB, Cruzen C, Simmons HA, Kemnitz JW and Weindruch R. Science (2009) 325:201-214.
320. Fontana L, Meyer TE, Klein S and Holloszy JO. Proc Natl Acad Sci USA (2004) 101:6659-6663.
321. Meyer TE, Kovács SJ, Ehsani AA, Klein S, Holloszy JO and Fontana L. J Am Coll Cardiol (2006) 47:398-402.
322. Imai S, Armstrong CM, KaeberleinM and Guarente L. Nature (2000) 403:795-800.
323. Cohen HY, Miller C, Bitterman KJ, Wall NR, Hekking B, Kessler B, Howitz KT, Gorospe M, de Cabo R and Sinclair DA. Science (2004) 305:390-392.
324. Bordone L, Cohen D, Robinson A, Motta MC, van Veen E, Czopik A, Steele AD, Crowe H, Marmor S, Luo J, Gu W and Guarente L. Aging Cell (2007) 6:759-767.
325. Boily G, Seifert EL, Bevilacqua L, He XH, Sabourin G, Estey C, Moffat C, Crawford S, Saliba S, Jardine K, Xuan J, Evans M, Harper ME, McBurney MW. PLoS One (2008) 3:e1759.
326. Michan S and Sinclair D. Biochem J (2007) 404:1-13.
327. Guarente L and Franklin H. N Engl J Med (2011) 364:2235-2244.
328. Kitada M, Kume S, Kanasaki K, Takeda-Watanabe A and Koya D. Curr Drug Targets (2013) 14:622-636.
329. Gomes P, Fleming Outeiro T and Cavadas C. Trends Pharmacol Sci (2015) 36:756-768.
330. McDonnell E, Peterson BS, Bomze HM and Hirschey MD. Trends Endocrinol Metab (2015) 26:486-492.

331. Kuang J, Chen L, Tang Q, Zhang J, Li Y and He J. Front Physiol (2018) 9:135.
332. Schenk S, Saberi M and Olefsky JM. J Clin Invest (2008) 118:2992-3002.
333. de Luca C and Olefsky JM. FEBS Lett (2008) 582:97-105.
334. Gerber PA and Rutter GA. Antioxid Redox Signal (2017) 26:501-518.
335. Morgan D, Oliveira-Emilio HR, Keane D, Hirata AE, Santos da Rocha M, Bordin S, Curi R, Newsholme P and Carpinelli AR. Diabetologia (2007) 50:359-369.
336. Arkan MC, Hevener AL, Greten FR, Maeda S, Li ZW, Long JM, Wynshaw-Boris A, Poli G, Olefsky J and Karin M. Nat Med (2005) 11:191-198.
337. Olefsky JM and Glass CK. Annu Rev Physiol (2010) 72:219-46.
338. Solinas G and Karin M. FASEB J (2010) 24:2596-2611.
339. Archuleta TL, Lemieux AM, Saengsirisuwan V, Teachey MK, Lindborg KA, Kim JS and Henriksen EJ. Free Radic Biol Med (2009) 47:1486-1493.
340. Evans JL, Maddux BA and Goldfine ID. Antioxid Redox Signal (2005) 7:1040–1052.
341. Evans JL, Goldfine ID, Maddux BA and Grodsky GM. Endocr Rev (2002) 23:599-622.
342. Bloch-Damti A and Bashan N. Antioxid Redox Signal (2005) 7:1553-1567.
343. Petersen KF, Dufour S, Befroy D, Garcia R and Shulman GI. N Engl J Med (2004) 50:664-671.
344. Morino K, Petersen KF, Dufour S, Befroy D, Frattini J, Shatzkes N, Neschen S, White MF, Bilz S, Sono S, Pypaert M and Shulman GI. J Clin Invest (2005) 115:3587-3593.
345. Morino K, Petersen KF and Shulman GI. Diabetes (2006) 55(Suppl. 2):S9-15.
346. Petersen KF, Befroy D, Dufour S, Dziura J, Ariyan C, Rothman DL, Rothman DL, DiPietro L, Cline GW and Shulman GI. Science (2003) 300:1140-1142.
347. Kitada M and Koya D. Diabetes Metab J (2013) 37:315-325.
348. Milne JC, Lambert PD, Schenk S, Carney DP, Smith JJ, Gagne DJ, Jin L, Boss O, Perni RB, Vu CB, Bemis JE, Xie R, Disch JS, Ng PY, Nunes JJ,

Lynch AV, Yang H, Galonek H, Israelian K, Choy W, Iffland A, Lavu S, Medvedik O, Sinclair DA, Olefsky JM, Jirousek MR, Elliott PJ, Westphal CH. Nature (2007) 450:712-716.

349. Smoliga JM, Baur JA and Hausenblas HA. Mol Nutr Food Res (2011) 55:1129-1141.

350. Tao Y, Huang C, Huang Y, Hong L, Wang H, Zhou Z and Qiu Y. Cardiovasc Toxicol (2015) 15:217-223.

351. Lang A, Anand R, Altinoluk-Hambüchen S, Ezzahoini H, Stefanski A, Iram A, Bergmann L, Urbach J, Böhler P, Hänsel J, Franke M, Stühler K, Krutmann J, Scheller J, Stork B, Reichert AS and Piekorz RP. Aging (2017) 9:2163-2189.

352. Singh CK, Chhabra G, Ndiaye MA, Garcia-Peterson LM, Mack NJ and Ahmad N. Antioxid Redox Signal (2018) 28:643-661.

353. Zhou L, Wang F, Sun R, Chen X, Zhang M, Xu Q, Wang Y, Wang S, Xiong Y, Guan KL, Yang P, Yu H and Ye D. EMBO Rep (2016) 17:811-822.

354. Wang Y, Liu Q, Huan Y, Li R, Li C, Sun S, Guo N, Yang M, Liu S and Shen Z. Exp Cell Res (2018) 371:205-213.

355. Ma Y and Fei X. Exp Ther Med (2018) 16:1417-1425.

356. Heerspink HJL, Perkins BA, Fitchett DH, Husain M and Cherney DZ. Circulation (2016) 134:752-772.

357. Vallon V and Thomson SC. Diabetologia (2017) 60:215-225.

358. Norton L, Shannon CE, Fourcaudot M, Hu C, Wang N, Ren W, Song J, Abdul-Ghani M, DeFronzo RA, Ren J and Jia W.Diabetes Obes Metab (2017) 19:1322-1326.

359. Solini A, Rossi C, Mazzanti CM, Proietti A, Koepsell H and Ferrannini E. Diabetes Obes Metab (2017) 19:1289-1294.

360. Wang XX, Levi J, Luo Y, Myakala K, Herman-Edelstein M, Qiu L, Wang D, Peng Y, Grenz A, Lucia S, Dobrinskikh E, D'Agati VD, Koepsell H, Kopp JB, Rosenberg AZ and Levi M. J Biol Chem (2017) 292:5335-5348.

361. Song P, Onishi A, Koepsell H and Vallon V. Exp Opin Ther Targets (2016)

20:1109-1125.
362. Scheen AJ. Drugs (2015) 75:33-59.
363. UK Prospective Diabetes Study (UKPDS) Group. Lancet (1998) 352:854-865.
364. Bailey C and Day C. Br J Diabetes Vasc Dis (2010) 10:193-199.
365. Day C and Bailey AV. Br J Diabetes Vasc Dis (2013) 13:2-6.
366. Vallon V and Thomson SC. Nat Rev Nephrol (2017) 13:517-518.
367. Wanner C, Inzucchi SE, Lachin JM, Fitchett D, von Eynatten M, Mattheus M, Johansen OE, Woerle HJ, Broedl UC and Zinman B. N Engl J Med (2016) 375:323-334.
368. Zinman B, Wanner C, Lachin JM, Fitchett D, Bluhmki E, Hantel S, Mattheus M, Devins T, Johansen OE, Woerle HJ, Broedl UC and Inzucchi SE. N Engl J Med (2015) 373:2117-2128.
369. Neal B, Perkovic V, Mahaffey KW, de Zeeuw D, Fulcher G, Erondu N, Shaw W, Law G, Desai M and Matthews DR. N Engl J Med (2017) 377:644-657.
370. American Diabetes Association. Diabetes Care (2018) 41:S55-S64.
371. Khunti K, Davies M, Majeed A, Thorsted BL, Wolden ML and Paul SK. Diabetes Care (2015) 38:316-322.
372. American Diabetes Association. Diabetes Care (2018) 41:S73-S85.
373. De Ferranti SD, De Boer IH, Fonseca V, Fox CS, Golden SH, Lavie CJ, Magge SN, Marx N, McGuire DK, Orchard TJ, Zinman B and Eckel RH. Diabetes Care (2014) 37:2843-2863.
374. Henry RR, Thakkar P, Tong C, Polidori D and Alba M. Diabetes Care (2015) 38:2258-2265.
375. Dandona P, Mathieu C, Phillip M, Hansen L, Griffen SC, Tschöpe D, Thorén F, Xu J and Langkilde AM. Lancet Diabetes Endocrinol (2017) 5:864-876.
376. Garg SK, Henry RR, Banks P, Buse JB, Davies MJ, Fulcher GR, Pozzilli P, Gesty-Palmer D, Lapuerta P, Simó R, Danne T, McGuire DK, Kushner JA, Peters A and Strumph P. N Engl J Med (2017) 377:2337-2348.
377. Sands AT, Zambrowicz BP, Rosenstock J, Lapuerta P, Bode BW, Garg SK, Buse JB, Banks P, Heptulla R, Rendell M, Cefalu WT and Strumph P.

Diabetes Care (2015) 38:1181-1188.

378. Grempler R, Thomas L, Eckhardt M, Himmelsbach F, Sauer A, Sharp DE, Bakker RA, Mark M, Klein T and Eickelmann P. Diabetes Obes Metab (2012) 14:83-90.
379. Ogawa W and Sakaguchi K. J Diabetes Investig (2016) 7:135-138.
380. Qiu H, Novikov A and Vallon V. Diabetes Metab Res Rev (2017) 33:e2886.
381. Peters AL, Buschur EO, Buse JB, Cohan P, Diner JC and Hirsch IB. Diabetes Care (2015) 38:1687-1693.
382. US Food and Drug Administration. FDA Drug Saf Commun (2015) 2-4.
383. Fadini GP, Bonora BM and Avogaro A. Diabetologia (2017) 60:1385-1389.
384. Blau JE, Tella SH, Taylor SI and Rother KI. Diabetes Metab Res Rev (2017) 33:e2924.
385. Jerums G, Premaratne E, Panagiotopoulos S and MacIsaac RJ. Diabetologia (2010) 53:2093-2104.
386. Magee GM, Bilous RW, Cardwell CR, Hunter SJ, Kee F and Fogarty DG. Diabetologia (2009) 52:691-697.
387. Vallon V, Blantz RC and Thomson S. J Am Soc Nephrol (2003) 14:530-537.
388. Vallon V, Richter K, Blantz RC, Thomson S and Osswald H. J Am Soc Nephrol (1999) 10:2569-2576.
389. Thomson SC, Rieg T, Miracle C, Mansoury H, Whaley J, Vallon V and Singh P. AJP Regul Integr Comp Physiol (2012) 302:R75-83.
390. Vallon V, Rose M, Gerasimova M, Satriano J, Platt KA, Koepsell H, Cunard R, Sharma K, Thomson SC and Rieg T. AJP Ren Physiol (2013) 304:F156-167.
391. Vallon V and Thomson SC. Annu Rev Physiol (2012) 74:351-375.
392. Vallon V, Gerasimova M, Rose MA, Masuda T, Satriano J, Mayoux E, Koepsell H, Thomson SC and Rieg T.AJP Ren Physiol (2014) 306:F194-204.
393. Cherney DZI, Perkins BA, Soleymanlou N, Maione M, Lai V, Lee A, Fagan NM, Woerle HJ, Johansen OE, Broedl UC and von Eynatten M. Circulation

(2014) 129:587-597.
394. Scheen AJ and Delanaye P. Diabetes Metab (2017) 43:99-109.
395. Scheen AJ. Diabetes Res Clin Pract (2016) 121:204-214.
396. Bolinder J, Ljunggren O, Kullberg J, Johansson L, Wilding J, Langkilde AM, Sugg J and Parikh S. J Clin Endocrinol Metab (2012) 97:1020-1031.
397. Yokono M, Takasu T, Hayashizaki Y, Mitsuoka K, Kihara R, Muramatsu Y, Miyoshi S, Tahara A, Kurosaki E, Li Q, Tomiyama H, Sasamata M, Shibasaki M, Uchiyama Y. Eur J Pharmacol (2014) 727:66-74.
398. Pessoa TD, Campos LCG, Carraro-Lacroix L, Girardi AC and Malnic G. J Am Soc Nephrol (2014) 25:2028-2039.
399. Layton AT and Vallon V. Am J Physiol Renal Physiol (2018) 314:F969-F984.
400. Ferrannini E, Mark M and Mayoux E. Diabetes Care (2016) 39:1108-1114.
401. Morgan NG and Dhayal S. Biochem Pharmacol (2009) 78:1419-1427.
402. Civelli O. Trends Pharmacol Sci (2005) 26:15-19.
403. Oh DY and Lagakos WS. Curr Opin Clin Nutr Metab Care (2011) 14: 322-327.
404. Yonezawa T, Kurata R, Yoshida K, Murayama MA, Cui X and Hasegawa A. Curr Med Chem (2013) 20:3855-3871.
405. Talukdar S, Olefsky JM and Osborn O. Trends Pharmacol Sci (2011) 32:543-550.
406. Mohammad S. Curr Drug Targets (2015) 16:771-775.
407. Negoro N, Sasaki S, Mikami S, Ito M, Suzuki M, Tsujihata Y, Ito R, Harada A, Takeuchi K, Suzuki N, Miyazaki J, Santou T, Odani T, Kanzaki N, Funami M, Tanaka T, Kogame A, Matsunaga S, Yasuma T and Momose Y. ACS Med Chem Lett (2010) 1:290-294.
408. Vettor R, Granzotto M, De Stefani D, Trevellin E, Rossato M, Farina MG, Milan G, Pilon C, Nigro A, Federspil G, Vigneri R, Vitiello L, Rizzuto R, Baratta R and Frittitta L. J Clin Endocrinol Metab (2008) 93:3541-3550.
409. Kebede M, Alquier T, Latour MG, Semache M, Tremblay C and Poitout V.

Diabetes (2008) 57:2432-2437.

410. Edfalk S, Steneberg P and Edlund H. Diabetes (2008) 57:2280-2287.

411. Schnell S, Schaefer M and Schofl C. Mol Cell Endocrinol (2007) 263:173-180.

412. Tomita T, Masuzaki H, Noguchi M, Iwakura H, Fujikura J, Tanaka T, Ebihara K, Kawamura J, Komoto I, Kawaguchi Y, Fujimoto K, Doi R, Shimada Y, Hosoda K, Imamura M and Nakao K. Biochem Biophys Res Commun (2005) 338:1788-1790.

413. Steneberg P, Rubins N, Bartoov-Shifman R, Walker MD and Edlund H. Cell Metab (2005) 1:245-258.

414. Fujiwara K, Maekawa F and Yada T. Am J Physiol Endocrinol Metab (2005) 289:E670-677.

415. Briscoe CP, Tadayyon M, Andrews JL, Benson WG, Chambers JK, Eilert MM, Ellis C, Elshourbagy NA, Goetz AS, Minnick DT, Murdock PR, Sauls HR Jr, Shabon U, Spinage LD, Strum JC, Szekeres PG, Tan KB, Way JM, Ignar DM, Wilson S and Muir AI. J Biol Chem (2003) 278:11303-11311.

416. Itoh Y, Kawamata Y, Harada M, Kobayashi M, Fujii R, Fukusumi S, Ogi K, Hosoya M, Tanaka Y, Uejima H, Tanaka H, Maruyama M, Satoh R, Okubo S, Kizawa H, Komatsu H, Matsumura F, Noguchi Y, Shinohara T, Hinuma S, Fujisawa Y and Fujino M. Nature (2003) 422:173-176.

417. Nagasumi K, Esaki R, Iwachidow K, Yasuhara Y, Ogi K, Tanaka H, Nakata M, Yano T, Shimakawa K, Taketomi S, Takeuchi K, Odaka H and Kaisho Y. Diabetes (2009) 58:1067-1076.

418. Briscoe CP, Peat AJ, McKeown SC, Corbett DF, Goetz AS, Littleton TR, McCoy DC, Kenakin TP, Andrews JL, Ammala C, Fornwald JA, Ignar DM, Jenkinson S. Br J Pharmacol (2006) 148:619-628.

419. Tomita T, Hosoda K, Fujikura J, Inagaki N and Nakao K. Front Endocrinol (Lausanne) (2014) 5:152.

420. Xiong Y, Swaminath G, Cao Q, Yang L, Guo Q, Salomonis H, Lu J, Houze JB, Dransfield PJ, Wang Y, Liu JJ, Wong S, Schwandner R, Steger F, Liu L,

Baribault H, Coberly S, Miao L, Zhang J, Lin DC and Schwarz M. Mol Cell Endocrinol (2013) 369:119-129.

421. Parker HE, Habib AM, Rogers GJ, Gribble FM and Reimann F. Diabetologia (2009) 52:289-298.

422. Habib AM, Richards P, Rogers GJ, Reimann F and Gribble FM. Diabetologia (2013) 56:1413-1416.

423. Yamashima T. Prog Lipid Res (2012) 51:221-231.

424. Boneva NB, Kikuchi M, Minabe Y and Yamashima T. J Pharmacol Sci (2011) 116:163-172.

425. Boneva NB and Yamashima T. Hippocampus (2012) 22:896-905.

426. Ma D, Zhang M, Larsen CP, Xu F, Hua W, Yamashima T, Mao Y and Zhou L. Brain Res (2010) 1330:1-8.

427. Ma D, Tao B, Warashina S, Kotani S, Lu L, Kaplamadzhiev DB, Mori Y, Tonchev AB and Yamashima T. Neurosci Res (2007) 58:394-401.

428. Khan MZ and He L. Neuropharmacology 2015.

429. Yamashima T. Prog Lipid Res (2015) 58:40-50.

430. Nakamoto K, Nishinaka T, Sato N, Mankura M, Koyama Y and Tokuyama S. Yakugaku Zasshi (2014) 134:397-403.

431. Nakamoto K and Tokuyama S. Nihon Shinkei Seishin Yakurigaku Zasshi (2012) 32:233-237.

432. Honore JC, Kooli A, Hamel D, Alquier T, Rivera JC, Quiniou C, Hou X, Kermorvant-Duchemin E, Hardy P, Poitout V and Chemtob S. Arterioscler Thromb Vasc Biol (2013) 33:954-961.

433. Skov V, Glintborg D, Knudsen S, Tan Q, Jensen T, Kruse TA, Beck-Nielsen H and Højlund K. PLoS One (2008) 3:e2466.

434. Skov V, Glintborg D, Knudsen S, Jensen T, Kruse TA, Tan Q, Brusgaard K, Beck-Nielsen H and Højlund K. Diabetes (2007) 56:2349-2355.

435. Timmers S, Konings E, Bilet L, Houtkooper RH, van de Weijer T, Goossens GH, Hoeks J, van der Krieken S, Ryu D, Kersten S, Moonen-Kornips E,

Hesselink MKC, Kunz I, Schrauwen-Hinderling VB, Blaak E, Auwerx J and Schrauwen P. Cell Metab (2011) 14:612-622.

436. Tateno C, Miya F, Wake K, Kataoka M, Ishida Y, Yamasaki C, Yanagi A, Kakuni M, Wisse E, Verheyen F, Inoue K, Sato K, Kudo A, Arii S, Itamoto T, Asahara T, Tsunoda T and Yoshizato K. Lab Invest (2013) 93:54-71.

437. Affo S, Dominguez M, Lozano JJ, Sancho-Bru P, Rodrigo-Torres D, Morales-Ibanez O, Moreno M, Millán C, Loaeza-del-Castillo A, Altamirano J, García-Pagán JC, Arroyo V, Ginès P, Caballería J, Schwabe RF and Bataller R. Gut (2013) 62:452-460.

438. van Dijk SJ, Feskens EJ, Bos MB, Hoelen DW, Heijligenberg R, Bromhaar MG, de Groot LC, de Vries JH, Müller M and Afman LA. Am J Clin Nutr (2009) 90:1656-1664.

439. She X, Rohl CA, Castle JC, Kulkarni AV, Johnson JM and Chen R. BMC Genomics (2009) 10:269.

440. Hardy OT, Perugini RA, Nicoloro SM, Gallagher-Dorval K, Puri V, Straubhaar J and Czech MP. Surg Obes Relat Dis (2011) 7:60-67.

441. Keller P, Gburcik V, Petrovic N, Gallagher IJ, Nedergaard J, Cannon B and Timmons JA. BMC Endocr Disord (2011) 11:7.

442. Wu HT, Chen W, Cheng KC, Ku PM, Yeh CH and Cheng JT. J Nutr Biochem (2012) 23:1264-1270.

443. Suh HN, Huong HT, Song CH, Lee JH, Han HJ. Am J Physiol Cell Physiol (2008) 295:C1518-1527.

444. Ou HY, Wu HT, Hung HC, Yang YC, Wu JS and Chang CJ. Am J Physiol Endocrinol Metab (2013) 304:E668-676.

445. Lin DC, Zhang J, Zhuang R, Li F, Nguyen K, Chen M, Tran T, Lopez E, Lu JY, Li XN, Tang L, Tonn GR, Swaminath G, Reagan JD, Chen JL, Tian H, Lin YJ, Houze JB and Luo J. PLoS One (2011) 6:e27270.

446. Houze JB, Zhu L, Sun Y, Akerman M, Qiu W, Zhang AJ, Sharma R, Schmitt M, Wang Y, Liu J, Liu J, Medina JC, Reagan JD, Luo J, Tonn G, Zhang J, Lu

JY, Chen M, Lopez E, Nguyen K, Yang L, Tang L, Tian H, Shuttleworth SJ and Lin DC. Bioorg Med Chem Lett (2012) 22:1267-1270.

447. Yazaki R, Kumagai N and Shibasaki M. Org Lett (2011) 13:952-955.

448. Tan CP, Feng Y, Zhou YP, Eiermann GJ, Petrov A, Zhou C, Lin S, Salituro G, Meinke P, Mosley R, Akiyama TE, Einstein M, Kumar S, Berger JP, Mills SG, Thornberry NA, Yang L and Howard AD. Diabetes (2008) 57:2211-2219.

449. Song F, Lu S, Gunnet J, Xu JZ, Wines P, Proost J, Liang Y, Baumann C, Lenhard J, Murray WV, Demarest KT and Kuo GH. J Med Chem (2007) 50:2807-2817.

450. Stoddart LA, Brown AJ and Milligan G. Mol Pharmacol (2007) 71:994-1005.

451. Tanaka H, Yoshida S, Minoura H, Negoro K, Shimaya A, Shimokawa T and Shibasaki M. Life Sci (2014) 94:115-121.

452. Fujita T, Matsuoka T, Honda T, Kabashima K, Hirata T and Narumiya S. J Invest Dermatol (2011) 131:1660-1667.

453. Christiansen E, Hansen SV, Urban C, Hudson BD, Wargent ET, Grundmann M, Jenkins L, Zaibi M, Stocker CJ, Ullrich S, Kostenis E, Kassack MU, Milligan G, Cawthorne MA and Ulven T. ACS Med Chem Lett (2013) 4: 441-445.

454. Brown SP, Dransfield PJ, Vimolratana M, Jiao X, Zhu L, Pattaropong V, Sun Y, Liu J, Luo J, Zhang J, Wong S, Zhuang R, Guo Q, Li F, Medina JC, Swaminath G, Lin DC and Houze JB. ACS Med Chem Lett (2012) 3:726-730.

455. Gowda N, Dandu A, Singh J, Biswas S, Raghav V, Lakshmi MN, Shilpa PC, Sunil V, Reddy A, Sadasivuni M, Aparna K, Verma MK, Moolemath Y, Anup MO, Venkataranganna MV, Somesh BP and Jagannath MR. BMC Pharmacol Toxicol (2013) 14:28.

456. Doshi LS, Brahma MK, Sayyed SG, Dixit AV, Chandak PG, Pamidiboina V, Motiwala HF, Sharma SD and Nemmani KV. Metabolism 2009; 58(3):333-343.

457. Ito R, Tsujihata Y, Matsuda-Nagasumi K, Mori I, Negoro N and Takeuchi K. Br J Pharmacol (2013) 170:568-580.

458. Lead GPR40 agonist bites the dust. Nat Rev Drug Discov [News and analysis] (2014) 13(2):91.

459. Srivastava A, Yano J, Hirozane Y, Kefala G, Gruswitz F, Snell G, Lane W, Ivetac A, Aertgeerts K, Nguyen J, Jennings A and Okada K. Nature (2014) 513:124-127.

460. Luo J, Swaminath G, Brown SP, Zhang J, Guo Q, Chen M, Nguyen K, Tran T, Miao L, Dransfield PJ, Vimolratana M, Houze JB, Wong S, Toteva M, Shan B, Li F, Zhuang R and Lin DC. PLoS One (2012) 7:e46300.

461. Tanaka H, Yoshida S, Oshima H, Minoura H, Negoro K, Yamazaki T, Sakuda S, Iwasaki F, Matsui T and Shibasaki M. J Pharmacol Exp Ther (2013) 346:443-452.

462. Zhao YF, Wang L, Zha D, Qiao L, Lu L, Yu J, Qu P, Sun Q, Qiu J and Chen C. J Mol Endocrinol (2013) 51:69-77.

463. Philippe C, Wauquier F, Leotoing L, Coxam V and Wittrant Y. Exp Cell Res (2013) 319:3035-3041.

464. Christiansen E, Urban C, Merten N, Liebscher K, Karlsen KK, Hamacher A, Spinrath A, Bond AD, Drewke C, Ullrich S, Kassack MU, Kostenis E and Ulven T. J Med Chem (2008) 51:7061-7064.

465. Christiansen E, Due-Hansen ME, Urban C, Grundmann M, Schröder R, Hudson BD, Milligan G, Cawthorne MA, Kostenis E, Kassack MU and Ulven T. J Med Chem (2012) 55:6624-6628.

466. Urban C, Hamacher A, Partke HJ, Roden M, Schinner S, Christiansen E, Due-Hansen ME, Ulven T, Gohlke H and Kassack MU. Naunyn Schmiedebergs Arch Pharmacol (2013) 386:1021-1030.

467. Wagner R, Kaiser G, Gerst F, Christiansen E, Due-Hansen ME, Grundmann M, Machicao F, Peter A, Kostenis E, Ulven T, Fritsche A, Häring HU, Ullrich S. Diabetes (2013) 62:2106-2111.

468. Naik H, Vakilynejad M, Wu J, Viswanathan P, Dote N, Higuchi T and Leifke

E. J Clin Pharmacol (2012) 52:1007-1016.

469. Leifke E, Naik H, Wu J, Viswanathan P, Demanno D, Kipnes M, Vakilynejad M. Clin Pharmacol Ther (2012) 92:29-39.

470. Kaku K, Araki T and Yoshinaka R. Diabetes Care (2013) 36:245-250.

471. Burant CF, Viswanathan P, Marcinak J, Cao C, Vakilynejad M, Xie B, Leifke E. Lancet (2012) 379:1403-1411.

472. Araki T, Hirayama M, Hiroi S and Kaku K. Diabetes Obes Metab (2012) 14:271-278.

473. Molecule of the Month. TAK-875. Drug News Perspect (2010) 23:667.

474. Frumento D, Ben Nasr M, El Essawy B, D'Addio F, Zuccotti GV and Fiorina P. J Endocr Invest (2017) 40:803-814.

475. van Belle TL, Coppieters KT, and von Herrath MG. Physiol Rev (2011) 91:79-118.

476. Pociot F and Lernmark A. Lancet (2106), 387:2331-2339.

477. Rewers M and Ludvigsson L. Lancet (2016) 387:2340-2348.

478. Leete P, Willcox A, Krogvold L, Dahl-Jørgensen K, Foulis AK, Richardson SJ and Morgan NG. Diabetes (2016) 65:1362-1369.

479. Krogvold L, Wiberg A, Edwin B, Buanes T, Jahnsen FL, Hanssen KF, Larsson E, Korsgren O, Skog O and Dahl-Jørgensen K. Diabetologia (2016) 59:492–501.

480. Coppieters KT, Dotta F, Amirian N, Campbell PD, Kay TW, Atkinson MA, Roep BO and von Herrath MG. J Exp Med (2012) 209:51–60.

481. Krogvold L, Skog O, Sundstrom G, Edwin B, Buanes T, Hanssen KF, Ludvigsson J, Grabherr M, Korsgren O and Dahl-Jørgensen K. Diabetes (2015) 64:2506-2512.

482. Cobo-Vuilleumier N, Lorenzo PI, Rodriguez NJ, Herrera Gómez IG, Fuente-Martin E, López-Noriega L, Mellado-Gil JM, Romero-Zerbo SY, Baquié M, Lachaud CC, Stifter K, Perdomo G, Bugliani M, De Tata V, Bosco D, Parnaud G, Pozo D, Hmadcha A, Florido JP, Toscano MG, de Haan P, Schoonjans K, Sánchez Palazón L, Marchetti P, Schirmbeck R, Martín-

Montalvo A, Meda P, Soria B, Bermúdez-Silva FJ, St-Onge L and Gauthier BR. Nature Commun (2018) 9:1488.

483. Atkinson MA, Roep BO, Posgai A, Wheeler DCS and Peakman M. Lancet Diab Endocr (2019) 7:52-64. .

484. Menendez-Gutierrez MP, Roszer T and Ricote M. Curr Topics Med Chem (2012) 12:548-584.

485. Tyagi S, Gupta P, Saini AS, Kaushal C and Sharma S. J Adv Pharm Technol Res (2011) 2:236-240.

486. Umemoto T and Fujiki Y. Genes to Cells (2012) 17:576-596.

487. Le Menn G and Neels JG. Int J Mol Sci (2018) 19:1575. .

488. Capelli D, Cerchia C, Montanari R, Loiodice F, Tortorella P, Laghezza A, Cervoni L, Pochetti G and Lavecchia A. Sci Rep (2016) 6:34792.

489. Poulsen L, Siersbaek M and Mandrup S. Sem Cell Dev Biol (2012) 23:631-639.

490. Zoete V, Grosdidier A and Michielin O. Biochimica Biophysica Acta (2007) 1771:915-925.

491. Nettles KW and Greene GL. Ann Rev Physiol (2005) 67:309-333.

492. Ricote M and Glass CK. Biochim Biophys Acta (2007) 1771: 926-935.

493. Torchia J, Glass C and Rosenfeld MG. Curr Opin Cell Biol (1998) 10:373-383.

494. Yessoufou A and Wahli W. Swiss Medical Weekly (2010) 140:w13071.

495. Mandard S, Muller M and Kersten S. Cell Mol Life Sci (2004) 61:393-416.

496. Zingarelli B, Piraino G, Hake PW, O'Connor M, Denenberg A, Fan H and Cook JA. Am J Pathol (2010) 177:1834-1847.

497. Fajas L, Auboeuf D, Raspe E Schoonjans K, Lefebvre AM, Saladin R, Najib J, Laville M, Fruchart JC, Deeb S, Vidal-Puig A, Flier J, Briggs MR, Staels B, Vidal H and Auwerx J. J Biol Chem (1997) 272:18779-18789.

498. Ricote M, Huang JT, Welch JS and Glass CK. J Leukoc Biol (1999) 66:733-739.

499. Odegaard JI, Ricardo-Gonzalez RR, Goforth MH, Morel CR, Subramanian V, Mukundan L, Red Eagle A, Vats D, Brombacher F, Ferrante AW, Chawla A. Nature (2007) 447:1116-1120.

500. Willson TM, Brown PJ, Sternbach DD and Henke BR. J Med Chem (2000) 43:527-550.

501. Racke MK, Gocke AR, Muir M, Diab A, Drew PD and Lovett-Racke AE. J Nutr (2006) 136:700-703.

502. Aprahamian TR, Bonegio RG, Weitzner Z, Gharakhanian R and Rifkin IR. Immunology (2014) 142:363-373.

503. Fallahi P, Ferrari SM, Elia G, Nasini F, Colaci M, Giuggioli D, Vita R, Benvenga S, Ferri C and Antonelli A. Expert Rev Clin Pharmacol (2016) 9:853-861.

504. Pawlak-Adamska E, Daroszewski J, Bolanowski M, Oficjalska J, Janusz P, Szalinski M and Frydecka I. Immunogenetics (2013) 65:493-500.

505. Okamoto H, Iwamoto T, Kotake S, Momohara S, Yamanaka H and Kamatani N. Clin Exp Rheumatol (2005) 23:323-330.

506. Lima EA, de Andrade Lima MMD, Marques CDL, Duarte ALBP, da Rocha Pita I and da Rocha Pita MG. An Bras Dermatol (2013) 88:1029-1035.

507. Ramkalawan H, Wang YZ, Hurbungs A, Yang YF, Tian FF, Zhou WB, Li J, Yang H, Xiao B and Zhang W. Inflammation (2012) 35:1338-1347.

508. Varga T, Czimmerer Z and Nagy L. BBA-Mol Basis Dis (2011) 1812:1007-1022.

509. Beeson PB. Am J Med (1994) 96:457-462.

510. Dunn SE, Ousman SS, Sobel RA, Zuniga L, Baranzini SE, Youssef S, Crowell A, Loh J, Oksenberg J and Steinman L. J Exp Med (2007) 204:321-330.

511. Holm LJ, Krogvold K, Hasselby JP, Kaur S, Claessens LA, Russell MA, Mathews CE, Hanssen KF, Morgan NG, Koeleman BPC, Roep BO, Gerling IC, Pociot F, Dahl-Jørgensen K and Buschard K. Diabetologia (2018) 61:1650-1661.

512. Janani C and Ranjitha Kumari BD, Diab Metab Syndr (2015) 9: 46-50.
513. Hull CM, Peakman M and Tree TI. Diabetologia (2017) 60:1839-1850.
514. Szablewski L. Int Immunopharmacol (2014) 22:182-191.
515. Bloem SJ and Roep BO. Diabetologia (2017) 60:1185-1189.
516. Willcox A, Richardson SJ, Bone AJ, Foulis AK and Morgan NG. Clin Exp Immunol (2009) 155:173-181.
517. Choi JM and Bothwell AL. Mol Cell (2012) 33:217-222.
518. Eizirik DL, Sammeth M, Bouckenooghe T, Bottu G, Sisino G, Igoillo-Esteve M, Ortis F, Santin I, Colli ML, Barthson J, Bouwens L, Hughes L, Gregory L, Lunter G, Marselli L, Marchetti P, McCarthy MI and Cnop M. PLoS Genet (2012) 8:e1002552.
519. Ravnskjaer K, Frigerio F, Boergesen M, Nielsen T, Maechler P and Mandrup S. J Lipid Res (2010) 51:1370-1379.
520. Dillon JS, Yaney GC, Zhou Y, Voilley N, Bowen S, Chipkin S, Bliss CR, Schultz V, Schuit FC, Prentki M, Waxman DJ and Corkey BE. Diabetes (2000) 49:2012-2020.
521. Gremlich S, Nolan C, Roduit R, Burcelin R, Peyot ML, Delghingaro-Augusto V, Desvergne B, Michalik L, Prentki M and Wahli W. Endocrinology (2005) 146:375-382.
522. Ravnskjaer K, Boergesen M, Dalgaard LT and Mandrup S. J Mol Endocrinol (2006) 36:289-299.
523. Bihan H, Rouault C, Reach G, Poitout V, Staels B and Guerre-Millo M. FEBS Lett (2005) 579:2284-2288.
524. Tang T, Abbott MJ, Ahmadian M, Lopes AB, Wang Y and Sul HS. Cell Metabol (2013) 18:883-895.
525. Iglesias J, Barg S, Vallois D, Lahiri S, Roger C, Yessoufou A, Pradevand S, McDonald A, Bonal C, Reimann F, Gribble F, Debril MB, Metzger D, Chambon P, Herrera P, Rutter GA, Prentki M, Thorens B and Wahli W. J Clin Investig (2012) 122:4105-4117.
526. Nakamichi Y, Kikuta T, Ito E, Ohara-Imaizumi M, Nishiwaki C, Ishida H and

Nagamatsu S. Biochem Biophys Res Commun (2003) 306:832-836.

527. Bollheimer LC, Troll S, Landauer H, Wrede CE, Scholmerich J and Buettner R. J Mol Endocrinol (2003) 31:61-69.

528. Ito E, Ozawa S, Takahashi K, Tanaka T, Katsuta H, Yamaguchi S, Maruyama M, Takizawa M, Katahira H, Yoshimoto K, Nagamatsu S and Ishida H. Biochem Biophys Res Commun (2004) 324:810-814.

529. Ravnskjaer K, Boergesen M, Rubi B, Larsen JK, Nielsen T, Fridriksson J, Maechler P and Mandrup S. Endocrinol (2005) 146:3266-3276.

530. Wang X, Zhou L, Shao L, Qian L, Fu X, Li G, Luo T, Gu Y, Li F, Li J, Zheng S and Luo M. Life Sci (2007) 81:160-165.

531. Gupta D, Jetton TL, Mortensen RM, Duan SZ, Peshavaria M and Leahy JL. J Biol Chem (2008) 283:32462-32470.

532. Shimabukuro M, Zhou YT, Lee Y and Unger RH. J Biol Chem (1998) 273:3547-3550.

533. Higa M, Zhou YT, Ravazzola M, Baetens D, Orci L and Unger RH. Proc Natl Acad Sci (1999) 96:11513-11518.

534. Finegood DT, McArthur MD, Kojwang D, Thomas MJ, Topp BG, Leonard T and Buckingham RE. Diabetes (2001) 50:1021-1029.

535. Jia DM, Tabaru A, Nakamura H, Fukumitsu KI, Akiyama T and Otsuki M. Metabolism (2000) 49:1167-1175.

536. Jia DM and Otsuki M. Pancreas (2002) 24:303-312.

537. Matsui J, Terauchi Y, Kubota N, Takamoto I, Eto K, Yamashita T, Komeda K, Yamauchi T, Kamon J, Kita S, Noda M and Kadowaki T. Diabetes (2004) 53:2844-2854.

538. Kim EK, Kwon KB, Koo BS, Han MJ, Song MY, Song EK, Han MK, Park JW, Ryu DG and Park BH. Int J Biochem Cell Biol (2007) 39:1260-1275.

539. Shen X, Yang L, Yan S, Zheng H, Liang L, Cai X and Liao M. Mol Cell Endocrinol (2015) 412:1-11.

540. Lin CY, Gurlo T, Haataja L, Hsueh WA and Butler PC. J Endocrinol Metab (2005) 90:6678-6686.

541. Hull RL, Shen Z-P, Watts MR, Kodama K, Carr DB, Utzschneider KM, Zraika S, Wang F and Kahn SE. Diabetes (2005) 54:2235-2244.
542. Akiyama M, Hatanaka M, Ohta Y, Ueda K, Yanai A, Uehara Y, Tanabe K, Tsuru M, Miyazaki M, Saeki S, Saito T, Shinoda K, Oka Y and Tanizawa Y. Diabetologia (2009) 52:653-663.
543. Evans-Molina C, Robbins RD, Kono T, Tersey SA, Vestermark GL, Nunemaker CS, Garmey JC, Deering TG, Keller SR, Maier B and Mirmira RG. Mol Cell Biol (2009) 29:2053-2067.
544. Pearson JA, Wong FS and Wen L. J Autoimmun (2016) 66:76-88.
545. Reed JC and Herold KC. Nature Rev Endocrinol (2015) 11:308-314.
546. Yaacob NS, Goh KS and Norazmi MN. Exp Toxic Pathol (2012) 64:127-131.
547. Yaacob NS, Kaderi MA and Norazmi MN. J Clin Immunol (2009) 29:595-602.
548. Beales PE and Pozzilli P. Diab Metab Res Rev (2002) 18:114-117.
549. Ogawa J, Takahashi S, Fujiwara T, Fukushige J, Hosokawa T, Izumi T, Kurakata S and Horikoshi H. Life Sci (1999) 65:1287-1296.
550. Fu Z, Zhen W, Yuskavage J and Liu D. Br J Nutr (2011) 105:1218-1225.
551. Zhang S, Yang X, Luo J, Ge X, Sun W, Zhu H, Zhang W, Cao J and Hou Y. Nutr Cancer (2014) 66:315-324.
552. Castro CN, Barcala Tabarrozzi AL, Winnewisser J, Gimeno ML, Antunica Noguerol M, Liberman AC, Paz DA, Dewey RA and Perone MJ. Clin Exp Immunol (2014) 177:149-160.
553. Mazidi M, Karimi E, Meydani M, Ghayour-Mobarhan M and Ferns GA. World J Methodol (2016) 6:112-117.
554. Weiss L, Zeira M, Reich S, Har-Noy M, Mechoulam R, Slavin S and Gallily R. Autoimmun (2006) 39:143-151.
555. O'Sullivan SE. Br J Pharmacol (2016) 173:1899-1910.
556. Bi X, Li F, Liu S, Jin Y, Zhang X, Yang T, Dai Y, Li X and Zhao AZ. J Endocrinol Invest (2017) 127:1757-1771.
557. Nevius E, Srivastava PK and Basu S. Mucosal Immunol (2012) 5:76-86.

558. Kang JH, Goto T, Han IS, Kawada T, Kim YM and Yu R. Obesity (2010) 18:780-787.

559. Arany E, Strutt B, Romanus P, Remacle C, Reusens B and Hill DJ. Diabetologia (2004) 47:1831-1837.

560. Bonfleur ML, Borck PC, Ribeiro RA, Caetano LC, Soares GM, Carneiro EM and Balbo SL. Life Sci (2015) 135:15-21.

561. Soares FL, de Oliveira ML, Teixeira LG, Menezes Z, Pereira SS, Alves AC, Batista NV, de Faria AM, Cara DC, Ferreira AV and Alvarez-Leite JI. J Nutr Biochem (2013) 24:1105-1111.

562. Funda DP, Kaas A, Bock T, Tlaskalova-Hogenova H and Buschard K. Diab Metabol Res Rev (1999) 15:323-327.

563. Antvorskov JC, Josefsen K, Haupt-Jorgensen M, Fundova P, Funda DP and Buschard K. J Diabet Res (2016) 2016:7.

564. Haupt-Jorgensen M, Larsen J, Josefsen K, Jørgensen TZ, Antvorskov JC, Hansen AK and Buschard K. Diab Metab Res Rev (2018) 34:e2987.

565. Scheen AJ and Lefebvre PJ. Diabetes Care (1999) 22:1568-1577.

566. Ahmad F, Li PM, Meyerovitch J and Goldstein BJ. J Biol Chem (1995) 270:20503-20508.

567. Zinker BA, Rondinone CM, Trevillyan JM, Gum RJ, Clampit JE, Waring JF, Xie N, Wilcox D, Jacobson P, Frost L, Kroeger PE, Reilly RM, Koterski S, Opgenorth TJ, Ulrich RG, Crosby S, Butler M, Murray SF, McKay RA, Bhanot S, Monia BP and Jirousek MR. Proc Nat Acad Sci USA (2002) 99:11357-11362.

568. Johnson TO, Ermolieff J, and Jirousek MR. Nat Rev Drug Discov (2002) 1:696-709.

569. Zhang S and Zhang ZY. Drug Discov Today (2007) 12:373-381.

570. Zhang ZY and Lee SY. Expert Opin Investig Drugs (2003) 12:223-233.

571. Bandyopadhyay D, Kusari A, Kenner KA, Liu F, Chernoff J, Gustafson TA and Kusari J. J Biol Chem (1997) 272:1639-1645.

572. Goldstein BJ, Bittner-Kowalczyk A, White MF and Harbeck M. J Biol Chem (2000) 275:4283-4289.

573. Lund IK, Hansen JA, Andersen HS, Moller NPH and Billestrup N. J Mol Endocrinol (2005) 34:339-351.

574. Byon JCH, Kusari AB and Kusari J. Mol Cell Biochem (1998) 182:101-108.

575. Kenner KA, Anyanwu E, Olefsky JM and Kusari J. J Biol Chem (1996) 271:19810-19816.

576. Meshkani R, Taghikhani M, Al-Kateb H, Larijani B, Khatami S, Sidiropoulos GK, Hegele RA and Adeli K. Clin Chem (2007) 53:1585-1592.

577. Elchebly M, Payette P, Michaliszy E, Cromlish W, Collins S, Loy AL, Normandin D, Cheng A, Himms-Hagen J, Chan CC, Ramachandran C, Gresser MJ, Tremblay ML and Kennedy BP.Science (1999) 283:1544-1548.

578. Klaman LD, Boss O, Peroni OD, Kim JK, Martino JL, Zabolotny JM, Moghal N, Lubkin M, Kim YB, Sharpe AH, Stricker-Krongrad A, Shulman GI, Neel BG and Kahn BB. Mol Cell Biol (2000) 20:5479-5489.

579. Bence KK, Delibegovic M, Xue B, Gorgun CZ, Hotamisligil GS, Neel BG and Kahn BB. Nat Med (2006) 12:917-924.

580. Zhao H, Liu G, Xin Z, Serby MD, Pei Z, Szczepankiewicz BG, Hajduk PJ, Abad-Zapatero C, Hutchins CW, Lubben TH, Ballaron SJ, Haasch DL, Kaszubska W, Rondinone CM, Trevillyan JM and Jirousek MR. Bioorgan Med Chem Lett (2004) 14:5543-5546.

581. Leung KW, Posner BI and Just G. Bioorg Med Chem Lett (1999) 9:353-356.

582. Andersen HS, Olsen OH, Iversen LF, Sorensen ALP, Mortensen SB, Christensen M, Branner S, Hansen TK, Lau JF, Jeppesen L, Moran EJ, Su J, Bakir F, Judge L, Shahbaz M, Collins T, Vo T, Newman MJ, Ripka WC and Møller NP. J Med Chem (2002) 45:4443-4459.

583. Wilson DP, Wan ZK, Xu WX, Kirincich SJ, Follows BC, Joseph-McCarthy D, Foreman K, Moretto A, Wu J, Zhu M, Binnun E, Zhang YL, Tam M, Erbe DV, Tobin J, Xu X, Leung L, Shilling A, Tam SY, Mansour TS and Lee J. J Med Chem (2007) 50:4681-4698.

584. Liu S, Zeng LF, Wu L, Yu X, Xue T, Gunawan AM, Long YQ and Zhang ZY. J Am Chem Soc (2008) 130:17075-17084.
585. Jiang CS, Liang LF and Guo YW. Acta Pharmacol Sin (2012) 33:1217-1245.
586. Yamazaki H, Sumilat DA, Kanno S, Ukai K, Rotinsulu H, Wewengkang DS, Ishikawa M, Mangindaan REP and Namikoshi M. J Nat Med (2013) 67:730-735.
587. Li Y, Zhang Y, Shen X and Guo YW. Bioorg Med Chem Lett (2009) 19:390-392.
588. Fantus IG, Deragon G, Lai R and Tang S. Mol Cell Biochem (1995) 153:103-112.
589. Fantus IG and Tsiani E. Mol Cell Biochem (1998) 182:109-119.
590. Meyerovitch J, Farfel Z, Sack J and Shechter Y. J Biol Chem (1987) 262:6658-6662.
591. Meyerovitch J, Backer J and Kahn CR. J Clin Investig (1989) 84:976-983.
592. Myers MP, Andersen JN, Cheng A, Tremblay ML, Horvath CM, Parisien JP, Salmeen A, Barford D and Tonks NK. J Biol Chem (2001) 276:47771-47774.
593. Cheng A, Uetani N, Simoncic PD, Chaubey VP, Lee-Loy A, McGlade CJ, Kennedy BP and Tremblay ML. Dev Cell (2002) 2:497-503.
594. Zabolotny JM, Bence-Hanulec KK, Stricker-Krongrad A, Haj F, Wang Y, Minokoshi Y, Kim YB, Elmquist JK, Tartaglia LA, Kahn BB and Neel BG. Dev Cell (2002) 2:489-495.
595. Ahmad F, Considine RV, Bauer TL, Ohannesian JP, Marco CC and Goldstein BJ. Metabolism (1997) 46:1140-1145.
596. Cheung A, Kusari J, Jansen D, Bandyopadhyay D, Kusari A and Bryer-Ash M. J Lab Clin Med (1999) 134:115-123.
597. Wu X, Hardy VE, Joseph JI, Jabbour S, Mahadev K, Zhu L and Goldstein BJ. Metabolism (2003) 52:705-712.
598. Brown-Shimer S, Johnson KA, Lawrence JB, Johnson C, Bruskin A, Green NR and Hill DE. Proc Natl Acad Sci USA (1990) 87:5148-5152.

599. Gerich JE. Mayo Clin Proc (2003) 78:447-456.
600. Creutzfeldt W. Diabetologia (1979) 16:75-85.
601. Nauck MA, Bartels E, Orskov C, Ebert R and Creutzfeldt W. J Clin Endocrinol Metab (1993) 76:912-917.
602. Nauck MA and Meier JJ. Diabetes Obes Metab (2018) 20(Suppl 1):5-21.
603. Nauck M, Stöckmann F, Ebert R and Creutzfeldt W. Diabetologia (1986) 29:46-52.
604. Nauck MA, Kleine N, Orskov C, Holst JJ, Willms B and Creutzfeldt W. Diabetologia (1993) 36:741-744.
605. Drucker DJ and Nauck MA. Lancet (2006) 368:1696-1705.
606. Mentlein R, Gallwitz B and Schmidt WE. Eur J Biochem (1993) 214:829-835.
607. Holst JJ and Deacon CF. Diabetes (1998) 47:1663-1670.
608. Deacon CF. Front Endocrinol (Lausanne) (2019) 10:80.
609. Drucker DJ. Cell Metab (2006) 3:153-165.
610. Sesti G, Avogaro A, Belcastro S, Bonora BM, Croci M and Daniele G. Acta Diabetol (2019) 56:605-617.
611. Tomovic K, Lazarevic J, Kocic G, Deljanin-Ilic M and Anderluh M. Med Res Rev (2019) 39:404-422.
612. Nabeno M, Akahoshi F, Kishida H, Miyaguchi I, Tanaka Y and Ishii S. Biochem Biophys Res Commun (2013) 434:191-196.
613. Patel BD and Ghate MD. Eur J Med Chem (2014) 74:574-605.
614. Kim SH, Yoo JH, LeeWJ and Park CY. Diabetes Metab J (2016) 40:339-353.
615. White WB, Bakris GL, Bergenstal RM, Cannon CP, Cushman WC, Fleck P, Heller S, Mehta C, Nissen SE, Perez A, Wilson C and Zannad F. Am Heart J (2011) 162:620-626.e1.
616. White WB, Cannon CP, Heller SR, Nissen SE, Bergenstal RM, Bakris GL, Perez AT, Fleck PR, Mehta CR, Kupfer S, Wilson C, Cushman WC and Zannad F. N Engl J Med. (2013) 369:1327-1335.

617. Mosenzon O, Raz I, Scirica BM, Hirshberg B, Stahre CI, Steg PG, Davidson J, Ohman P, Price DL, Frederich B, Udell JA, Braunwald E and Bhatt DL. Diabetes Metab Res Rev (2013) 29:417-426.

618. Scirica BM, Bhatt DL, Braunwald E, Steg PG, Davidson J, Hirshberg B Ohman P, Frederich R, Wiviott SD, Hoffman EB, Cavender MA, Udell JA, Desai NR, Mosenzon O, McGuire DK, Ray KK, Leiter LA and Raz I. N Eng J Med (2013) 369:1317-1326.

619. Green JB, Bethel MA, Paul SK, Ring A, Kaufman KD, Shapiro DR, Califf RM and Holman RR. Am Heart J (2013) 166:983-989.e7.

620. Green JB, BethelMA, Armstrong PW, Buse JB, Engel SS, Garg J, Josse R, Kaufman KD, Koglin J, Korn S, Lachin JM, McGuire DK, Pencina MJ, Standl E, Stein PP, Suryawanshi S, Van de Werf F, Peterson ED and Holman RR. N Engl J Med (2015) 373:232-242.

621. Rosenstock J, Perkovic V, Alexander JH, Cooper ME, Marx N, Pencina MJ, Toto RD, Wanner C, Zinman B, Baanstra D, Pfarr E, Mattheus M, Broedl UC, Woerle HJ, George JT, von Eynatten M and McGuire DK. Cardiovasc Diabetol (2018) 17:39.

622. Rosenstock J, Perkovic V, Johansen OE, Cooper ME, Kahn SE, Marx N, Alexander JH, Pencina M, Toto RD, Wanner C, Zinman B, Woerle HJ, Baanstra D, Pfarr E, Schnaidt S, Meinicke T, George JT, von Eynatten M and McGuire DK. JAMA (2019) 321:69-79.

623. Bailey CJ and Marx N. Diabetes Obes Metab (2019) 21:3-14.

624. Scirica BM, Braunwald E, Raz I, Cavender MA, Morrow DA, Jarolim P, Udell JA, Mosenzon O, Im K, Umez-Eronini AA, Pollack PS, Hirshberg B, Frederich R, Lewis BS, McGuire DK, Davidson J, Steg PG, Bhatt DL. Circulation (2014) 130:1579-1588.

625. Udell JA, Bhatt DL, Braunwald E, Cavender MA, Mosenzon O, Steg PG, Davidson JA, Nicolau JC, Corbalan R, Hirshberg B, Frederich R, Im K, Umez-Eronini AA, He P, McGuire DK, Leiter LA, Raz I and Scirica BM. Diabetes Care (2015) 38:696-705.

626. McGuire DK, Van de Werf F, Armstrong PW, Standl E, Koglin J, Green JB, Bethel MA, Cornel JH, Lopes RD, Halvorsen S, Ambrosio G, Buse JB, Josse RG, Lachin JM, Pencina MJ, Garg J, Lokhnygina Y, Holman RR and Peterson ED. JAMA Cardiol (2016) 1:126-135.

627. Gallwitz B, Rosenstock J, Rauch T, Bhattacharya S, Patel S, von Eynatten M, Dugi KA and Woerle HJ. Lancet (2012) 380:475-483.

628. Marx N, Rosenstock J, Kahn SE, Zinman B, Kastelein JJ, Lachin JM, Espeland MA, Bluhmki E, Mattheus M, Ryckaert B, Patel S, Johansen OE and Woerle HJ. Diab Vasc Dis Res (2015) 12:164-174.

629. Davies MJ, D'Alessio DA, Fradkin J, Kernan WN, Mathieu C, Mingrone G, Rossing P, Tsapas A, Wexler DJ and Buse JB. Diabetes Care (2018) 41:2669-2701.

630. Davies MJ, D'Alessio DA, Fradkin J, Kernan WN, Mathieu C, Mingrone G, Rossing P, Tsapas A, Wexler DJ and Buse JB.Diabetologia (2018) 61:2461-2498.

631. American Diabetes Association. Diabetes Care (2019) 42(Suppl 1):S90-S102.

632. Inzucchi SE, Bergenstal RM, Buse JB, Diamant M, Ferrannini E, Nauck M, Peters AL, Tsapas A, Wender R and Matthews DR. Diabetologia (2012) 55:1577-1596.

633. Inzucchi SE, Bergenstal RM, Buse JB, Diamant M, Ferrannini E, Nauck M, Peters AL, Tsapas A, Wender R and Matthews DR. Diabetologia (2015) 58:429-442.

634. Nauck MA, Kahle M, Baranov O, Deacon CF and Holst JJ. Diabetes Obes Metab (2017) 19:200-207.

635. Tahrani AA, Bailey CJ, Del Prato S and Barnett AH. Lancet (2011) 378:182-197.

636. Rines AK, Sharabi K, Tavares CD and Puigserver P. Nat Rev Drug Discov (2016) 15:786-804.

637. Perseghin G. J Clin Endocrinol Metab (2010) 95:4871-4873.

638. Petersen MC, Vatner DF and Shulman GI. Nat Rev Endocrinol (2017) 13:572-587.
639. Agius L. Expert Opin Ther Pat (2014) 24:1155-1159.
640. Nakamura A and Terauchi Y. J Diabetes Investig (2015) 6:124-132.
641. Choi JM, Seo MH, Kyeong HH, Kim E and Kim HS. Proc Natl Acad Sci USA (2013) 110:10171-10176.
642. Raimondo A, Rees MG and Gloyn AL. Curr Opin Lipidol (2015) 26:88-95.
643. Matschinsky FM and Wilson DF. Front Physiol (2019) 10:148.
644. Agius L. Biochem J (2008) 414:1-18.
645. Chakera AJ, Steele AM, Gloyn AL, Shepherd MH, Shields B, Ellard S and Hattersley AT. Diabetes Care (2015) 38:1383-1392.
646. Steele AM, Shields BM, Wensley KJ, Colclough K, Ellard S and Hattersley AT. JAMA (2014) 311:279-286.
647. Amed S and Oram R. Can J Diabetes (2016) 40:449-454.
648. Ping F, Wang Z and Xiao X. J Diabetes Investig (2019) 10:1454-1462.
649. Doliba NM, Fenner D, Zelent B, Bass J, Sarabu R and Matschinsky FM. Diabetes Obes Metab (2012) 14(Suppl 3):109-119.
650. Lu M, Li P, Bandyopadhyay G, Lagakos W, Dewolf WE Jr, Alford T, Chicarelli MJ, Williams L, Anderson DA, Baer BR, McVean M, Conn M, Véniant MM and Coward P. PLoS One (2014) 9:e88431.
651. Nakamura A, Shimazaki H, Ohyama S, Eiki J and Terauchi Y. J Diabetes Investig (2011) 2:276-279.
652. Nakamura A, Terauchi Y, Ohyama S, Kubota J, Shimazaki H, Nambu T, Takamoto I, Kubota N, Eiki J, Yoshioka N, Kadowaki T and Koike T. Endocrinology (2009) 150:1147-1154.
653. Nakamura A, Togashi Y, Orime K, Sato K, Shirakawa J, Ohsugi M, Kubota N, Kadowaki T and Terauchi Y. Diabetologia (2012) 55:1745-1754.
654. Grimsby J, Sarabu R, Corbett WL, Haynes NE, Bizzarro FT, Coffey JW, Guertin KR, Hilliard DW, Kester RF, Mahaney PE, Marcus L, Qi L, Spence

CL, Tengi J, Magnuson MA, Chu CA, Dvorozniak MT, Matschinsky FM and Grippo JF. Science (2003) 301:370-373.

655. Filipski KJ and Pfefferkorn JA. Expert Opin Ther Pat (2014) 24:875-891.

656. Sarabu R, Berthel SJ, Kester RF and Tilley JW. Expert Opin Ther Pat (2011) 21:13-33.

657. Pfefferkorn JA. Expert Opin Drug Discov (2013) 8:319-330.

658. Egan A and Vella A. Expert Opin Investig Drugs (2019) 28:741-747.

659. Matschinsky FM. Trends Pharmacol Sci (2013) 34:90-99.

660. Amin NB, Aggarwal N, Pall D, Paragh G, Denney WS, Le V, Riggs M and Calle RA. Diabetes Obes Metab (2015) 17:751-759.

661. Brouwers M, Jacobs C, Bast A, Stehouwer CDA and Schaper NC. Trends Mol Med (2015) 21:583-594.

662. Scheen AJ. Lancet Diabetes Endocrinol (2018) 6:591-593.

663. Tornovsky-Babeay S, Dadon D, Ziv O, Tzipilevich E, Kadosh T, Schyr-Ben Haroush R, Hija A, Stolovich-Rain M, Furth-Lavi J, Granot Z, Porat S, Philipson LH, Herold KC, Bhatti TR, Stanley C, Ashcroft FM, Veld PI, Saada A, Magnuson MA, Glaser B and Dor Y. Cell Metab (2014) 19:109-121.

664. Arden C, Petrie JL, Tudhope SJ, Al-Oanzi Z, Claydon AJ, Beynon RJ, Towle HC and Agius L. Diabetes (2011) 60:3110-3120.

665. Arden C, Tudhope SJ, Petrie JL, Al-Oanzi ZH, Cullen KS, Lange AJ, Towle HC and Agius L. Biochem J (2012) 443:111-123.

666. Baggio LL and Drucker DJ. Gastroenterology (2007) 132:2131-2157.

667. DeFronzo RA. Diabetes (2009) 58:773-795.

668. Elrick H, Stimmler L, Hlad CJ Jr and Arai Y. J Clin Endocrinol Metab (1964) 24:1076-1082.

669. Drucker DJ and Nauck MA. Lancet (2006) 368:1696-1705.

670. Meier JJ and Nauck MA. Curr Opin Investig Drugs (2004) 5:402-410.

671. Nauck M, Stockmann F, Ebert R and Creutzfeldt W. Diabetologia (1986) 29: 46-52.

672. Calanna S, Christensen M, Holst JJ, Laferrère B, Gluud LL, Vilsbøll T and Knop FK. Diabetes Care (2013) 36:3346-3352.
673. Nauck MA and Meier JJ. Lancet Diabetes Endocrinol (2016) 4:525-536.
674. Drucker DJ. Nat Clin Pract End Met (2005) 1:22-31.
675. Nauck MA, Kleine N, Orskov C, Holst JJ, Willms B and Creutzfeldt W. Diabetologia (1993) 36:741-744.
676. Zander M, Madsbad S, Madsen JL and Holst JJ. Lancet (2002) 359:824-830.
677. Gupta V. Indian J Endocrinol Metab (2013) 17:413-421.
678. Byetta (exenatide) prescribing information. Wilmington, Del., AstraZeneca, February 2015.
679. Victoza (liraglutide) prescribing information. Plainsboro, N.J., Novo Nordisk, March 2015.
680. Bydureon (exenatide extended release) prescribing information. Wilmington, Del., AstraZeneca, September 2015.
681. Tanzeum (albiglutide) prescribing information. Wilmington, Del., GlaxoSmithKline, May 2015.
682. Trulicity (dulaglutide) prescribing information. Indianapolis, Ind., Eli Lilly And Company, March 2015.
683. Adlyxin (lixisenatide) prescribing information. Bridgewater, N.J., Sanofi-Aventis, July 2016.
684. Kayaniyil S, Lozano-Ortega G, Bennett HA, Johnsson K, Shaunik A, Grandy S and Kartman B. Diabetes Ther (2016) 7:27-43.
685. Fox CS, Golden SH, Anderson C, Bray GA, Burke LE, de Boer IH, Deedwania P, Eckel RH, Ershow AG, Fradkin J, Inzucchi SE, Kosiborod M, Nelson RG, Patel MJ, Pignone M, Quinn L, Schauer PR, Selvin E and Vafiadis DK. Diabetes Care (2015) 38:1777-1803.
686. Handelsman Y, Bloomgarden ZT, Grunberger G, Umpierrez G, Zimmerman RS, Bailey TS, Blonde L, Bray GA, Cohen AJ, Dagogo-Jack S, Davidson JA, Einhorn D, Ganda OP, Garber AJ, Garvey WT, Henry RR, Hirsch IB, Horton ES, Hurley Dl, Jellinger PS, Jovanovič L, Lebovitz HE, LeRoith D,

Levy P, McGill GB, Mechanick JI, Mestman JH, Moghissi ES, Orzeck EA, Pessah-Pollack R, Rosenblit PD, Vinik AI, Wyne K and Zangeneh F. Endocr Pract (2015) 21(Suppl. 1):1-87.

687. Acosta A, Camilleri M, Burton D, O'Neill J, Eckert D, Carlson P and Zinsmeister AR. Physiol Rep (2015) 3:e12610

688. U.S. Food and Drug Administration. Silver Spring, Md., U.S. Food and Drug Administration, 2008.

689. Gaede P, Vedel P, Larsen N, Jensen GV, Parving HH and Pedersen O. N Engl J Med (2003) 348:383-393.

690. Saraiva FK and Sposito AC. Cardiovasc Diabetol (2014) 13:142-156.

691. Monami M, Dicembrini I, Nardini C, Fiordelli I and Mannucci E. Diabetes Obes Metab (2014) 16:38-47.

692. Best JH, Hoogwerf BJ, Herman WH, Best JH, Hoogwerf Bj, Herman WH, Pelletier EM, Smith DB, Wenten M and Hussein MA. Diabetes Care (2011) 34:90-95.

693. Pfeffer MA, Claggett B, Diaz R, Dickstein K, Gerstein HC, Køber LV, Lawson FC, Ping L, Wei X, Lewis EF, Maggioni AP, McMurray JJV, Probstfield JL, Riddle MC, Solomon SD and Tardif J-C. N Engl J Med (2015) 373:2247-2257.

694. Marso SP, Daniels GH, Brown-Frandsen K, Kristensen P, Mann JFE, Nauck MA, Nissen SE, Pocock S, Poulter NR, Ravn LS, Steinberg WM, Stockner M, Zinman B, Bergenstal RM and Buse JB. N Engl J Med (2016) 375:311-322.

695. Holman RR, Bethel MA, George J, Sourij H, Doran Z, Keenan J, Khurmi NS, Mentz RJ, Oulhaj A, Buse JB, Chan JC, Iqbal N, Kundu S, Maggioni AP, Marso SP, Öhman P, Pencina MJ, Poulter N, Porter LE, Ramachandran A, Zinman B and Hernandez AF. Am Heart J (2016) 174:103-110.

696. Garber AJ, Abrahamson MJ, Barzilay JI, Blonde L, Bloomgarden ZT, Bush MA, Dagogo-Jack S, DeFronzo RA, Einhorn D, Fonseca VA, Garber JR, Garvey WT, Grunberger G, Handelsman Y, Henry RR, Hirsch IB, Jellinger

PS, McGill JB, Mechanick JI, Rosenblit PD and Umpierrez GE. Endocr Pract (2016) 22:84-113.

697. American Diabetes Association. Diabetes Care (2016) 39(Suppl. 1):S1-S112.

698. Cernea S and Raz I. Diabetes Care (2011) 34(Suppl. 2):S264-S271.

699. Wideman RD and Kieffer TJ. Trends Endocrinol Metab (2009) 20:280-286.

700. Triplitt C, Wright A and Chiquette E. Pharmacotherapy (2006) 26:360-374.

701. Gallwitz B. Treat Endocrinol (2005) 4:361-370.

702. Holst JJ. Gastroenterology (1994) 107:1048-1055.

703. Drucker DJ. Diabetes (1998) 47:159-169.

704. Vilsboll T and Holst JJ. Diabetologia (2004) 47:357-366.

705. Elahi D, McAloon-Dyke M, Fukagawa NK, Meneilly GS, Sclater AL, Minaker KL, Habener JF and Andersen DK. Regul Pept (1994) 51:63-75.

706. Nauck MA, Bartels E, Oerskov C, Ebert R and Creutzfedlt WJ. J Clin Endocrinol Metabol (1993) 76:912- 917.

707. Gentilella R, Bianchi C, Rossi A and Rotella CM. Diabetes Obes Metab (2009) 11:544-556.

708. Gonzalez C, Beruto V, Keller G, Santoro S and Di Girolamo G. Expert Opin Invest Drugs (2006) 15:887-895.

709. Holst JJ. Diabetes Metab Res and Rev (2002) 18:430-441.

710. Hansen L, Deacon CF, Orskov C and Holst JJ. Endocrinology (1999) 140: 5356-5363.

711. Nagakura T, Yasuda N, Yamazaki K, Ikuta H, Yoshikawa S, Asano O and Tanaka I. Biochem Biophys Res Commun (2001) 284:501-506.

712. Marguet D, Baggio L, Kobayashi T, Bernard A, Pierres M, Nielsen PF, Ribel U, Watanabe T, Drucker DJ and Wagtmann N. Proc Natl Acad Sci USA (2000) 97:6874-6879.

713. Durinx C, Lambeir A, Bosmans E, Falmagne J, Berghmans R, Haemers A, Scharpe S and De Meester I. Eur J Biochem (2000) 267:5608-5613.

714. Iwaki-Egawa S, Watanabe Y, Kikuya Y and Fujimoto Y. J Biochem (1998) 124:428-433.

715. Villhauer EB, Brinkman JA, Naderi GB, Burkey BF, Dunning BE, Prasad K, Mangold BL, Russell ME and Hughes TE. J Med Chem (2003) 46:2774-2789.

716. Kim D, Wang L, Beconi M, Eiermann GJ, Fisher MH, He H, Hickey GJ, Kowalchick JE, Leiting B, Lyons K, Marsilio F, McCann ME, Patel RA, Petrov A, Scapin G, Patel SB, Roy RS, Wu JK, Wyvratt MJ, Zhang BB, Zhu L, Thornberry NA and Weber AE. J Med Chem (2005) 48:141-151.

717. Augeri DJ, Robl JA, Betebenner DA, Magnin DR, Khanna A, Robertson JG, Wang A, Simpkins LM, Taunk P, Huang Q, Han SP, Abboa-Offei B, Cap M, Xin L, Tao L, Tozzo E, Welzel GE, Egan DM, Marcinkeviciene J, Chang SY, Biller SA, Kirby MS, Parker RA and Hamann LG. J Med Chem (2005) 48: 5025-5037.

718. Feng J, Zhang Z, Wallace MB, Stafford JA, Kaldor SW, Kassel DB, Navre M, Shi L, Skene RJ, Asakawa T, Takeuchi K, Xu R, Webb DR and Gwaltney SL. J Med Chem (2007) 50:2297-2300.

719. Havale SH and Pal M. Bioorg Med Chem (2009) 17:1783-1802.

720. Brigance RP, Meng W, Fura A, Harrity T, Wang A, Zahler R, Kirby MS and Hamman LG. Bioorg Med Chem Lett (2010) 20:4395-4398.

721. Meng W, Brigance RP, Chao HJ, Fura A, Harrity T, Marcinkkeviciene J, O'Connor SP, Tamura JK, Xie D, Zhang Y, Klei HE, Kish K, Weigett CA, Turdi H, Wang A, Zahler R, Kirby MS and Hamann LG. J Med Chem (2010) 53:5620-5628.

722. Yaron A and Naider F. Crit Rev Biochem Mol Biol (1993) 28:31-81.

723. Mentlein R, Gallwitz B and Schmidt WE. Eur J Biochem (1993) 214: 829-835.

724. Kieffer TJ, McIntosh CHS and Pederson RA. Endocrinology (1995) 136:3585-3596.

725. Deacon CF, Nauck MA, Toft-Nielsen M, Pridal L, Willms B and Holst JJ. Diabetes (1995) 44:1126-1131.

726. Ashworth DM, Atrash B, Baker GR, Baxter AJ, Jenkins PD, Jones DM and Szelke M. Bioorg Med Chem Lett (1996) 6:2745-2748.

727. Ashworth DM, Atrash B, Baker GR, Baxter AJ, Jenkins PD, Jones DM and Szelke M. Bioorg Med Chem Lett (1996) 6:1163-1166.

728. Ahren B, Gomis R, Standl E, Mills D and Schweizer AA. 64th Scientific Sessions of the American Diabetes Association. Orlando, FL. (2004) Late-breaking 7-LB (Abstract).

729. Tsu H, Chen X, Chen C-T, Lee S-J, Chang C-N, Kao K-H, Coumar MS, Yeh Y-T, Chien C-H, Wang H-S, Lin K-T, Chang Y-Y, Wu S-H, Chen Y-S, Lu I-L, Wu S-Y, Tsai T-Y, Chen W-C, Hsieh H-P, Chao Y-S and Jiaang W-T. J Med Chem (2006) 49:373-380.

730. Kim D, Wang L, Beconi M, Eiermann GJ, Fisher MH, He H, Hickey GJ, Kowalchick JE, Leiting B, Lyons K, Marsilio F, McCann ME, Patel RA, Petrov A, Scapin G, Patel SB, Sinha Roy R, Wu JK, Wyvratt MJ, Zhang BB, Zhu L, Thornberry NA and Weber AE. J Med Chem (2005) 48:141-151.

731. Edmondson SE, Fisher MH, Kim D, Maccoss M, Parmee ER, Weber AE and Xu J. US Patent 6699871B2, March 2, 2004.

732. Villhauer EB, Brinkman JA, Naderi GB, Burkey BF, Dunning BE, Prasad K, Mangold BL, Russell ME and Hughes TE. J Med Chem (2003) 46: 2774-89.

733. Magnin DR, Robl JA, Sulsky RB, Augeri DJ, Huang Y, Simpkins LM, Taunk PC, Betebenner DA, Robertson JG, Abboa-Offei BE, Wang A, Cap M, Xin L, Tao L, Sitkoff DF, Malley MF, Gougoutas JZ, Khanna A, Huang Q, Han S, Parker RA and Hamann LG. J Med Chem (2004) 47:2587-2598.

734. Kim D, Kowalchick JE, Edmondson SD, Mastracchio A, Xu JM, Eiermann GJ, Leiting B, Wu JK, Pryor K, Patel RA, He H, Lyons KA, Thornberry NA and Weber AE. Bioorg Med Chem Lett (2007) 17:3373-3377.

735. Xu J, Ok HO, Gonzalez EJ, Colwell LF Jr, Habulihaz B, He H, Leiting B, Lyona KA, Marsilio F, Patel RA, Wu JK, Thornberry NA, Weber AE and Parmee ER. Bioorg Med Chem Lett (2004) 14:4759-4762.

736. Brockunier L, He J, Colwell LF Jr, Habulihaz B, He H, Leiting B, Lyons KA, Marsilio F, Patel R, Teffera Y, Wu JK, Thornberry NA, Weber AE, Parmee ER. Bioorg Med Chem Lett (2004) 14:4763-4766.

737. Kim D, Kowalchick JE, Brockunier LL, Parmee ER, Eiermann GJ, Fisher MH, He H, Leiting B, Lyons K, Scapin G, Patel SB, Petrov A, Pryor KD, Roy RS, Wu JK, Zhang X, Wyvratt MJ, Zhang BB, Zhu L, Thornberry N and Weber AE. J Med Chem (2008) 51:589-602.

738. Hemmer MC, Steinhauer V and Gasteiger J. Vib Spectrosc (1999) 19:151-164.

739. Silverman BD. J Chem Inform Comput Sci (2000) 40:1470-1476.

740. Schuur JH, Selzer P and Gasteiger J. J Am Chem Soc (1996) 36:334-344.

741. (a) Consonni V, Todeschini R and Pavan M. J Chem Inform Comput Sci (2002) 42:682-692. (b) Consonni V, Todeschini R, Pavan M and Gramatica P. J Chem Inform Comput Sci (2002) 42:693-705.

742. (a) Todeschini R, Lasagni M and Marengo E. J Chemom (1994) 8:263-273. (b) Todeschini R and Gramatica P. 3D QSAR in drug design. (Kubinyi H, Folkers G and Martin YC Eds.), Kluwer/ESCOM, Dordrecht (The Netherlands) (1998) Vol 2:355-380.

743. American Diabetes Association. Diabetes Care (2004) 27:S5.

744. Porte D Jr and Schwartz MW. Science (1996) 272:699-700.

745. Saltiel AR. Cell (2001)104:517-529.

746. Zimmet P, Alberti KG and Shaw J. Nature (2001) 414:782-787.

747. Kopelman PG and Hitman GA. Lancet (1998) 352:SIV5

748. King H, Aubert RE and Herman WH. Diabetes Care (1998) 21:1414-1431.

749. Moller DE. Nature (2001) 414:821-827.

750. Mangelsdorf DJ and Evans RM. Cell (1995) 83:841-850.

751. Isseman I and Green S. Nature (1990) 347:645-650.

752. Berger J and Moller DE. Annu Rev of Med (2002) 53:409-435.

753. Staels B and Fruchart J-C. Diabetes (2005) 54:2460-2470.

754. Berger JP, Akiyama TE and Meinke PT. Trends Pharmacol Sci (2005) 26:244-251.

755. Willson TM, Brown PJ, Sternbach DD and Henke BR. J Med Chem (2000) 43:527-550.

756. Henke BR, J Med Chem (2004) 47:4118-4127.

757. Lehmann JM, Moore MB, Smith-Oliver TA, Wilkinson WO, Willson TM and Kliewer SA. J Biol Chem (1995) 270:12953-2956.

758. Willson TM, Cobb JE, Cowan DJ, Wiethe RW, Correa ID, Prakash SR, Beck KD, Moore LB, Kliewer SA and Lehmann JM. J Med Chem (1996) 39:665-668.

759. Lohray BB, Lorhay VB, Bajji AC, Kalchar S, Poondra RR, Padakanti S, Chakrabarti R, Vikramadithyan RK, Misra P, Juluri S, Mamidi RNVS and Rajagopalan R. J Med Chem (2001) 44:2675-2678.

760. Sauerberg P, Petterson I, Jeppesen L, Bury PS, Mogensen JP, Wassermann K, Brand CL, Sturis J, Wöldike HF, Fleckner J, Anderson AST, Mortensen SB, Anders-Svensson L, Rasmussen HB, Lehmann SV, Polivka Z, Sindelar K, Panajotova V, Ynddal L and Wulff EM. J Med Chem (2002) 45:789-804.

761. Ebdrup S, Petterson I, Rasmussen HB, Deussen HJ, Frost-Jensen A, Mortensen SB, Fleckner J, Pridal L, Nygaard L and Sauerberg P. J Med Chem (2003) 46:1306-1317.

762. Devasthale PV, Chen S, Jeon Y, Qu F, Shao C, Wang W, Zhang H, Farreley D, Golla R, Grover G, Harrity T, Ma Z, Moore L, Seethala R, Ren J, Cheng L, Sleph P, Sun W, Tieman A, Wetterau JR, Chang SY, Doweyko A, Chandrasena G, Humphreys JR, Sasseville VG, Ryono DE, Biller SA, Selan F, Harihanan N and Cheng PTW. J Med Chem (2005) 48:2248-2250.

763. Koyama H, Miller DJ, Boueres JK, Desai RC, Jones AB, Berger JP, MacNaul KL, Kelly LJ, Doebber TW, Wu MS, Zhou G, Wang PR, Ippolito MC, Chao YS, Agrawal AK, Franklin R, Heck JV, Wright SD, Moller DE and Sahoo SP. J Med Chem (2004) 47:3255-3263.

764. Shi GQ, Dropinski JF, McKeever BM, Xu S, Becker JW, Berger JP, MacNaul KL, Elbrecht A, Zhou G, Doebber TW, Wang P, Chao YS, Forrest M, Heck JV, Moller DE and Jones AB. J Med Chem (2005) 48:4457-4468.

765. Gervois P, Fruchart J-C and Staels B. Int J Clinic Pract (2004) 58:22-29.

766. Parmenon C, Viaud-Massuard M-C, Guillard J, Dacquet C, Ktorza A and Caignard D-H. (2006). WO2006/079719.

767. Parmenon C, Guillard J, Caignard D-H, Hennuyer N, Staels B, Audinot-Bouchez V, Boutin J-A, Ktorza A, Dacquet C and Viaud-Massuard M-C. Bioorg Med Chem Lett (2008) 18:1617-1622.

768. Parmenon C, Guillard J, Caignard D-H, Hennuyer N, Staels B, Audinot-Bouchez V, Boutin J-A, Dacquet C, Ktorza A, Viaud-Massuard M-C. Bioorg Med Chem Lett (2009) 19:2683-2687.

769. Golbraikh A and Tropsha A. J Mol Graph Model (2002) 20:269-276.

770. Ohkubo Y. Kishikawa H, Araki E, Miyata T, Lsami S, Motoyoshi S, Kojima Y, Furuyoshi N and Shichiri M. Diabs Res Clinic Pract (1995) 28:103-117.

771. Haffner SM, Alexander CM, Cook TJ, Boccuzzi SJ, Musliner TA, Pedersen R, Kjekshus J and Pyörälä K. Arch Int Med, (1999) 159:2661-2667.

772. Pyörälä K, Pedersen TR, KJekshus J, Faergeman O, Olsson AG and Thorgeirsson G. Diabetes Care (1997) 20:614-620.

773. Sacks FM, Pfeffer MA, Moye LA, Rouleau JL, Rutherford ID, Cole TG, Brown L, Warnica JW, Arnold JMO, Wun CC, Davis BR and Braunwald E. New Eng J Med (1996) 335:1001-1009.

774. Vamecq J and Latruffe N. Curr Opin Endocrinol Diabetes (2000) 7:8-18.

775. Evans RM, Barish GD and Wang Y-X. Nat. Med. (2004) 10:355-361.

776. Cho N and Momose Y. Curr Topic Med Chem (2008) 8:1483-1507.

777. Rikimaru K, Wakabayashi T, Abe H, Imoto H, Maekawa T, Ujikawa O, Murase K, Matsuo T, Matsumoto M, Nomura C, Tsuge H, Arimura N, Kawakami K, Sakamoto J, Funami M, Mol CD, Snell GP, Bragstad KA, Sang B-C; Dougan DR, Tanaka T, Katayama N, Horiguchi Y, Momose Y. Bioorg Med Chem (2012) 20:714-733.

778. Momose Y, Maekawa T, Yamano T, Kawada M, Odaka H, Ikeda H and Sohda T. J Med Chem (2002) 45:1518-1534.

779. Henke BR. Prog Med Chem (2004) 42:1-53.

780. Loviscach M, Rehman N, Carter L, Mudaliar S, Mohadeen P, Ciaraldi TP, Veerkamp JH and Henry RR. Diabetlogia (2000) 43:304-311.

781. Minoura H, Takeshita S, Ita M, Hirosumi J, Mabuchi M, Kawamura I, Nakajima S, Nakayama O, Kayakiri H, Oku T, Okubo-Suzuki A, Fukagawa M, Kojo H, Hanioka K, Yamasaki N, Imoto T, Kobayashi Y and Mutoh S. Eur J Pharmacol (2004) 494:273-281.

782. Minoura H, Takeshita S, Yamamoto T, Mabuchi M, Hirosumi J, Takakura S, Kawamura I, Seki J, Manda T, Ita M and Mutoh S. Eur J Pharmacol (2005) 519:182-190.

783. Kyowa Hakko Kirin Company Limited. Yanagisawa A, Otsubo W, Ueno K, Suzuki M, Matsubara M, Saeki S, Yao K, Hamaguchi A and Aratake S. (2008). PCT Int. Appl. WO2008/096820.

784. Kowa Company Limited. Miura T, Onogi K, Gomi N, Araki T, Tagashira J, Sekimoto R, Ishida R, Aoki H and Ohgiya T. (2011). PCT Int. Appl. WO2011/024468.

785. Hopkins CR, O'Neil SV, Laufersweiler MC, Wang Y, Pokross M, Mekel M, Evdokimov A, Walter R, Kontoyianni M, Petrey ME, Sabatakos G, Roesgen JT, Richardson E and Demuth TP. Bioorg Med Chem Lett (2006) 16:5659-5663.

786. Etgen GJ and Mantlo N. Curr Topic Med Chem (2003) 3:1649.

787. Rikimaru K, Wakabayashi T, Abe H, Tawaraishi T, Imoto H, Yonemori J, Hirose H, Murase K, Matsuo T, Matsumoto M, Nomura C, Tsuge H, Arimura N, Kawakami K, Sakamoto J, Funami M, Mol CD, Snell GP, Bragstad KA, Sang B-C, Dougan DR, Tanaka T, Katayama N, Horiguchi Y and Momose Y. Bioorg Med Chem (2012) 20:3332-3358.

788. Danaei G, Finucane MM, Lu Y, Singh GM, Cowan MJ, Paciorek CJ, Lin JK, Farzadfar F, Khang Y-H, Stevens GA, Rao M, Ali MK, Riley LM, Robinson CA and Ezzati M. Lancet (2011) 378:31-40.

789. International Diabetes Federation, IDF diabetes atlas seventh edition 2015. http://www.diabetesatlas.org/.

790. Millatt LJ, Hanf R and Hum DW. Drug Discov World (2011) Summer: 9-15.

791. Stein SA, Lamos EM and Davis SN. Expert Opin Drug Saf (2013) 12:153-75.

792. Lovshin JA and Drucker DJ. Nat Rev Endocrinol (2009) 5:262-269.

793. Verspohl EJ. Pharmacol Therapeut (2009) 124:113-138.

794. Mikhail N. Expert Opin Invest Drugs (2008) 17:845-853.

795. Koro CE, Bowlin SJ, Bourgeois N and Fedder DO. Diabetes Care (2004) 27: 17-20.

796. Shah U and Kowalski TJ. Curr Opin Drug Discov Develop (2010) 12:519-32.

797. Shah U. Curr Opin Drug Discov Develop (2009) 12:519-532.

798. Jones RM, Leonard JN, Buzard DJ and Lehmann, J. Expert Opin Ther Pat (2009) 19:1339-1359.

799. Fyfe MC, McCormack JG, Overton HA, Procter MJ and Reynet C. Expert Opin Drug Discov (2008) 3:403-413.

800. Overton HA, Fyfe MC and Reynet C. Br J Pharmacol (2008) 153:S76-S81.

801. Ohishi T and Yoshida S. Expert Opin Invest Drugs (2012) 21:321-328.

802. Overton HA, Babbs AJ, Doel SM, Fyfe MC, Gardner LS, Griffin G, Jackson HC, Procter MJ, Rasamison CM, Tang-Christensen M, Widdowson PS, Williams GM and Reynet C. Cell Metabol (2006) 3:167-175.

803. Soga T, Ohishi T, Matsui T, Saito T, MatsumotoM, Takasaki J, Matsumoto S-I, Kamohara M, Hiyama H, Yoshida S, Momose K, Ueda Y, Matsushime H, Kobori M and Furuichi K. Biochem Biophys Res Commun (2005) 326:744-751.

804. Chu Z-L, Jones RM, He H, Carroll C, Gutierrez V, Lucman A, Moloney M, Gao H, Mondala H, Bagnol D, Unett D, Liang Y, Demarest K, Semple G, Behan DP and Leonardv J. Endocrinology (2007) 148:2601-2609.

805. Chu Z-L, Carroll C, Alfonso J, Gutierrez V, He H, Lucman A, Pedraza M, Mondala H, Gao H, Bagnol D, Chen R, Jones RM, Behan DP and Leonard J. Endocrinology (2008) 149:2038-2047.

806. Baggio LL and Drucker DJ. Ann Rev Med (2006) 57:265-281.

807. Aaboe K, Krarup T, Madsbad S and Holst JJ. Diabet Obes Metabol (2008) 10: 994-1003.

808. Farilla L, Bulotta A, Hirshberg B, Calzi SL, Khoury N, Noushmehr H, Bertolotto C, Mario UD, Harlan DM, and Perfetti R. Endocrinology (2003) 144:5149-5158.

809. Trümper A, Trümper K, Trusheim H, Arnold R, Göke B and Hörsch D. Mol Endocrinol (2001) 15:1559-1570.

810. Fyfe, M, Mccormack J, Overton H, Procter M and Reynet C. American Diabetes Association 68thAnnual Scientific Sessions, San Francisco, CA, USA (2008) Abstract 297-OR.

811. Flock G, Holland D, Seino Y and Drucker DJ. Endocrinology (2011) 152:374-383.

812. Goodman ML, Dow J, Van Vliet AA, Pleszko A and Lockton JA. Diabetologia (2011) 54: Abstract OP32.

813. Roberts B, Gregoire FM, Karpf DB, Clemens E, Lavan B, Johnson J, Mcwherter CA, Martin R and Wilson M. American Diabetes Association 69th Annual Scientific Sessions, New Orleans, LA, USA (2009) Abstract 164-OR.

814. Roberts B, Karpf DB, Martin R, Lavan B, Wilson M and Mcwherter CA. American Diabetes Association 70th Annual Scientific Sessions, Orlando, FL, USA (2010) Abstract 603-P.

815. Goodman, ML, Dow J, Van Vliet AA, Hadi S, Karbiche D and Lockton JA. Diabetes (2011) 60:Abstract A273.

816. Nunez DJ, Lewis EW, Swan S, Bush MA, Cannon C, Mcmullen SL, Collins DA and Feldman PL. American Diabetes Association 70th Annual Scientific Sessions, Orlando, FL, USA (2010) Abstract 80-OR.

817. Jones RM, Leonard JN, Buzard DJ, Lehmann J. Expert Opin Ther Pat (2009) 19:1339-1359.

818. Ritter K, Buning C, Halland N, Pöverlein C, J Med Chem. (2016) 59:3579-3592.

819. Semple G, Ren A, Fioravanti B, Pereira G, Calderon I, Choi K, Xiong Y, Shin Y-J, Gharbaoui T, Sage CR, Morgan M, Xing C, Chu Z-L, Leonard JN, Grottick AJ, Al-Shamma H, Liang Y, Demarest KT and Jones RM. Bioorg Med Chem Lett (2011) 21:3134-3141.

820. Polli JW, Hussey E, Bush M, Generaux G, Smith G, Collins D, McMullen S, Turner N and Nunez DJ. Xenobiotica (2013) 3:498-508.

821. http://clinicaltrials.gov/ct2/show/NCT01035879.

822. Matsuda D, Kobashi Y, Mikami A, Kawamura M, Shiozawa F, Kawabe K, Hamada M, Oda K, Nishimoto S, Kimura K, Miyoshi M, Takayama N, Kakinuma H and Ohtake N.Bioorg Med Chem Lett (2016) 26:3441-3446.

823. Alelyunas YW, Empfield JR, McCarthy D, Spreen RC, Bui K, Pelosi-Kilby L and Shen C. Bioorg Med Chem Lett (2010) 20:7312-7316.

824. Matsuda D, Kobashi Y, Mikami A, Kawamura M, Shiozawa F, Kawabe K, Hamada M, Nishimoto S, Kimura K, Miyoshi M, Takayama N, Kakinuma H and Ohtake N. Bioorg Med Chem (2017) 25:4339-4354.

825. Sato K, Sugimoto H, Rikimaru K, Imoto H, Kamaura H, Negoro N, Tsujihata Y, Miyashita H, Odani T and Murata T. Bioorg Med Chem (2014) 22:1649-1666.

Published Papers

LIST OF PUBLICATIONS

1. Parihar R and Sharma BK. QSAR study on the dipeptide-derived inhibitor of dipeptidyl peptidase IV: The (2*S*)-Cyanopyrrolidine analogues. Br J Med Health Res (2020) 7(06):107-126.
2. Parihar R and Sharma BK. QSAR rationales for the dipeptidyl peptidase-4 (DPP-4) inhibitors: The Imidazolopyrimidine amides. GSC Biol Pharma Sci (2020) 11(03):130-144.
3. Parihar R, Jahan A and Sharma BK. 3D-descriptors in modeling the dipeptidyl peptidase IV (DPP-4) inhibition activity: Derivatives of β-Aminoamide bearing subsituted triazolopiperazines. J Innov Pharma Biol Sci (2020) 7(3):39-53.
4. Parihar R and Sharma BK. QSAR rationales for the PPARα/γ agonistic activity of 4,4-Dimethyl-1,2,3,4-tetrahydroquinoline derivatives. World J Biol Pharm Health Sci (2020) 3(1):38-53.
5. Parihar R and Sharma BK. CP-MLR derived qsar study on the PPARγ agonists: The benzylpyrazole acylsulfonamide derivatives. GSC Adv Res Rev (2020) 4(02):9-22.
6. Parihar R, Jahan A and Sharma BK. CP-MLR derived QSAR rationales for the PPARγ agonistic activity of the pyridyloxybenzene-acylsulfonamide derivatives. GSC Biol Pharma Sci (2020) 12(1):273-285.
7. Parihar R and Sharma BK. CP-MLR/PLS derived QSAR rationales for the GPR119 agonistic activity of the indole-based derivatives. Am J PharmTech Res (2020) 10(4):237-256.
8. Sharma BK and Parihar R. QSAR study on the hGPR119 agonistic activity of triazolopyridines. Am J Pharm Health Res (2020) 8(08):1-17.

BJMHR
British Journal of Medical and Health Research
Journal home page: www.bjmhr.com

QSAR study on the Dipeptide-Derived Inhibitor of Dipeptidyl Peptidase IV: The (2S)-Cyanopyrrolidine analogues

Raghuraj Parihar[1], Brij Kishore Sharma[1]*
Department of Chemistry, Government College, Bundi-323 001, India

ABSTRACT

The inhibition activity of (2*S*)-cyanopyrrolidine analogues for dipeptidyl peptidase IV has been quantitatively analyzed in terms of topological 0D-, 1D- and 2D-descriptors based on molecular graph theory. Statistically sound models have been obtained between the activity and various DRAGON descriptors through combinatorial protocol-multiple linear regression (CP-MLR) computational procedure. Amongst the large number of such derived models, the most significant ones have only been discussed to draw meaningful conclusions. Additionally the inhibition activity for DPP8 enzyme, reported for a limited number of such congeners, has also been correlated with such descriptors. From the final statistically significant models, it appeared that the mode of actions of titled compounds were different for DPP IV and DPP8 enzyme systems. Applicability domain analysis carried out for DPP IV inhibitors revealed that the suggested model matches the high quality parameters with good fitting power and the capability of assessing external data and all of the compounds was within the applicability domain of the proposed model and were evaluated correctly.

Keywords: QSAR, DPP IV inhibitors, Combinatorial protocol in multiple linear regression (CP-MLR) analysis, Dragon descriptors, (2*S*)-cyanopyrrolidine analogues.

*Corresponding Author Email: bksharma_sikar@rediffmail.com
Received 05 June 2020, Accepted 19 June 2020

Please cite this article as: Sharma BK *et al*., QSAR study on the Dipeptide-Derived Inhibitor of Dipeptidyl Peptidase IV: The (2S)-Cyanopyrrolidine analogues . British Journal of Medical and Health Research 2020.

INTRODUCTION

The most potent known insulinotropic hormone is glucagon-like peptide-1 (GLP-1) [1-3]. This hormone, containing 30 amino acids, is produced by L-cells of the intestinal mucosa that results from tissue specific processing of the proglucagon gene [4,5]. The stimulation of insulin secretion, inhibition of glucagons release [6-9] and slow down of gastric emptying [10-13] is due to the active form of GLP-1. These beneficial effects are helpful in controlling the glucose homeostasis in patients with type 2 diabetes [14-16]. However, GLP-1 is rapidly degraded by plasma DPP-IV and is lacking of oral activity; its use as a therapeutic agent is, therefore, restricted. In view of this a small molecule, as the inhibitor of DPP-IV, may extend the duration of action of GLP-I and result in the beneficial effects of this hormone for a long period of time. Through human clinical trials, it was shown that inhibition of DPP-IV may improve glucose tolerance in diabetic patients and healthy volunteers and leads to a new strategy for the treatment of type 2 diabetes [17-21]. DPP-IV is a serine protease, able to cleave the N-terminal dipeptide having preference for L-proline or L-alanine at the penultimate position [22-25]. A large number of DPP-IV inhibitors resemble the P2-P1 dipeptidyl substrate cleavage product. The simplest inhibitors are the compounds which are not having a carbonyl functionality of the proline residue, e.g., aminoacyl pyrrolidines and thiazolidines, possessing moderate inhibition activity for DPP-IV. Replacement of hydrogen with an electrophilic nitrile group at the 2-position of the pyrrolidine, in some compounds, elicited a 1000-fold increase in potency compared to the unsubstituted pyrrolidines [26].

One of the potent and stable representatives of the nitrile class is cyclohexylglycine-(2*S*)-cyanopyrrolidine, having a K_i value of 1.4 nM and an excellent chemical stability $t_{1/2}$ ~ 48 h at pH 7.4[27]. Another class, similar to proline inhibitors, was synthesized with diverse N-substituted glycines in the P2 site [17]. In this class, the side chain was moved from the α-carbon to the terminal nitrogen, led to two potent derivatives which have showed greater efficacy in clinical trial [28]. From this study, it was concluded that (2*S*)-cyanopyrrolidine derivatives with *N*-substituted glycine in the P2 site are more selective for DPP-IV than α-carbon-substituted glycine. An interesting study has recently been reported to develop a new pharmacophore in the P2 site with N-substituted glycine [29]. Initially, the P2 site amine extension was designed using β-alanine as building block and it was coupled the C-terminal with various substituted amines to generate a novel pharmacophore in the P2 site. Then, the N-terminal of the β-alanine derivative was combined with the P1 site α-bromoacetyl (2*S*)-cyanopyrrolidide to design 2-[3-[[2-[(2*S*)-2-cyano-1-pyrrolidinyl]-2-oxoethyl]amino]-1-oxopropyl]-based DPP-IV inhibitors. The structure-activity relationships of several series (I–III) of these DPP-IV inhibitors were explored to discover the potent and selective DPP-IV inhibitors. Series I, II, and III, being N-

substituted glycine derivatives include, respectively, the bicyclic ring system, monocyclic piperazine ring, and phenylalkyl groups. These compounds were tested for inhibition of DPP-IV, DPP8, and DPP-II. The activity was evaluated in terms of the concentration of a compound required to bring out 50% inhibition of the enzyme concerned. The aim of present communication is to establish the quantitative relationships between the reported activities and molecular descriptors unfolding the substitutional changes in titled compounds.

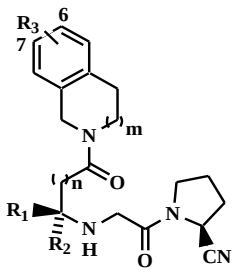
MATERIALS AND METHOD

Data set

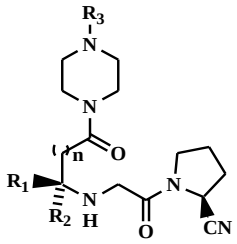
The reported twenty five (2*S*)-cyanopyrrolidine analogues, belonging to series I, II, and III are considered to formulate the data set for present study [29]. The structural variations and the activity values, expressed as IC_{50}(nM), of the reported analogues are given in Table 1. Since the activity variation for DPP-II is very small, therefore, inhibition profiles for DPP-IV and DPP8 have only been considered for quantitative analysis.

For the purpose of modeling study, the data set has been divided into training and test sets. One fifth of the compounds, from this data set, have been included in the test set for the validation of derived models while remaining compounds were used to derive the model correlating biological activity with descriptors unfolding molecular structures. The test-set, containing 5 compounds out of the 25 active ones, was generated in the SYSTAT [30] using the single linkage hierarchical cluster procedure involving the Euclidean distances of the activity. The selection of test set from the cluster tree was done in such a way to keep the test compounds at a maximum possible distance from each other. In this way, the identified test set will further ensure the statistical significance and reasonable predictability of derived models. As the leave-one-out (LOO) procedure has been applied to each model, therefore, corresponding to test set the derived model would be validated both internally and externally. The training and test set compounds are listed in Table 1.

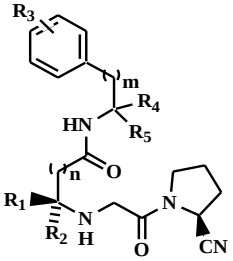
Table 1: **General structures and structural variations of (2*S*)-Cyanopyrrolidine analogues.**



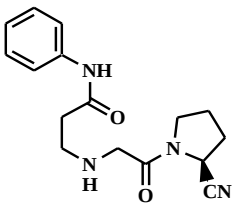
Series I (1-12)



Series II (13-16)



Series III (17-24)



25

Cpd.	n	m	R_1	R_2	R_3	R_4	R_5	IC_{50}(nM)[a]	
								DPP-IV	DPP8

1.	0	1	H	H	H			3236	4169
2.	1	1	H	H	H			116	3583
3.	1	1	H	H	6,7-diOMe			651	3340
4.[b]	1	1	H	H	6-F			83	1700
5.	1	0	H	H	H			132	2121
6.	2	1	H	H	H			428	1407
7.	1	1	CH_3	H	H			54	5346
8.[b]	1	1	*i*-Pr	H	H			811	41859
9.	1	1	CH_3	CH_3	H			49	>100000
10.	1	1	CH_3	CH_3	6-F			30	>100000
11.	1	1	CH_3	CH_3	6,8-F_2			22	>100000
12.	1	0	CH_3	CH_3	H			15	>100000
13.	1		H	H	CO(3,5-$F_2C_6H_3$)			676	202
14.	1		H	H	$SO_2C_6H_4$-4-$NHCOCH_3$			418	3416
15.	1		H	H	nicotinonitrile			629	2000
16.	1		H	H	benzothiazole			527	2117
17.	1	0	H	H	H	H	H	452	10744
18.	1	0	H	H	4-NO_2	H	H	317	2387
19.[b]	1	0	H	H	H	H	ethyl[c]	447	21961
20.	1	0	H	H	3,5-F_2	H	H	369	5532
21.	1	0	H	H	H	H	*i*-Pr[c]	784	12847
22.[b]	1	0	H	H	H	CH_3	CH_3	119	8338
23.	1	0	CH_3	CH_3	H	CH_3	CH_3	1108	>100000
24.	1	1	H	H	H	H	H	564	2592
25.[b]								298	855

[a]IC_{50} represents the concentration of a compound required to bring out 50% inhibition of DPP-IV and DPP8, taken from ref [29]; [b]compound of test set; [c]The stereochemistry at the benzylic carbon is not defined (mixture of diastereomers).

Theoretical molecular descriptors

The structures of the compounds were drawn in 2D ChemDraw [31], converted into 3D modules and subjected to energy minimization in the MOPAC using the AM1 procedure for closed shell system to ensure a well defined conformer relationship among the compounds under investigation. DRAGON software [32] was then used to compute the molecular descriptors of titled compounds. This software offers several hundreds of descriptors corresponding to 0D-, 1D-, and 2D-descriptor modules. The modules include ten different classes, namely, the constitutional (CONST), the topological (TOPO), the molecular walk counts (MWC), the BCUT descriptors (BCUT), the Galvez topological charge indices (GALVEZ), the 2D-autocorrelations (2D-AUTO), the functional groups (FUNC), the atom-centered fragments (ACF), the empirical descriptors (EMP), and the properties describing descriptors (PROP). For each of these classes the DRAGON software computes a large number of descriptors which are characteristic to the molecules under multi-descriptor environment. A total number of 484 descriptors, belonging to above classes, have been computed to generate most appropriate models describing the biological activity. The combinatorial protocol in multiple linear regression (CP-MLR) [33] method, discussed below, has been used further for developing

statistical significant models divulging quantitative structure-activity relationship (QSAR). Before the application of CP-MLR procedure, all those descriptors which are intercorrelated beyond 0.90 and showing a correlation of less than 0.1 with the biological endpoints (descriptor vs. activity, $r < 0.1$) were excluded. This has reduced the total dataset of the compounds from 484 to 90 and 84 descriptors as relevant ones for the DPP-IV and DPP8 inhibitory activity, respectively. The descriptors, in all above models, have been scaled between the intervals 0 to 1 [34] to ensure that a descriptor will not dominate simply because it has larger or smaller pre-scaled value compared to the other descriptors. In this way, the scaled descriptors would have equal potential to influence the QSAR models.

Model development

The CP-MLR software is developed for the selection of most appropriate descriptors from a pool, which are subsequently used to develop statistical significant models in a systematic manner. Its procedural aspects and implementation are discussed in some of our publications [35-40]. The thrust of this software is implicated mainly on its embedded "filters" which have been interfaced with multiple linear regression (MLR) to extract diverse structure-activity models, each having unique combination of descriptors from the dataset under investigation. In this procedure, the contents and number of variables to be evaluated are mixed according to the predefined confines and the 'filters' are significance evaluators of the variables in regression at different stages of model development. Of these, filter-1 is set in terms of inter-parameter correlation cutoff criteria for variables to stay as a subset (filter-1, default value 0.3 and upper limit ≤ 0.79). In this, if two variables are correlated higher than a predefined cutoff value the respective variable combination is forbidden and will be rejected. The second filter is in terms of t-values of regression coefficients of variables associated with a subset (filter-2, default value 2.0). Here, if the ratio of regression coefficient and associated standard error of any variable is less than a predefined cutoff value then the variable combination will be rejected. Since successive additions of variables to multiple regression equation will increase successive multiple correlation coefficient (r) values, square-root of adjusted multiple correlation coefficient of regression equation, r-bar, has been used to compare the internal explanatory power of models with different number of variables. Accordingly, a filter has been set in terms of predefined threshold level of r-bar (filter-3, default value 0.79) to decide the variables' 'merit' in the model formation. Finally, to exclude false or artificial correlations, the external consistency of the variables of the model have been addressed in terms of cross-validated Q^2 criteria from the leave-one-out (LOO) cross-validation procedure as default option (filter-4, default threshold value $0.3 \leq Q^2 \leq 1.0$). All these filters make the variable selection process efficient and lead to unique solution. In order to collect the descriptors with higher information content and explanatory power, the threshold of filter-3 may be incremented successively with

increasing number of descriptors (per equation) by considering the r-bar value of the preceding optimum model as the new threshold for next generation.

Model validation

The subdivision of data set into training set and test set have been used, respectively, for model development and external prediction. Goodness of fit of the models was assessed by examining the multiple correlation coefficient (r), the standard deviation (s) and the F-ratio between the variances of calculated and observed activities (F). The internal validation of derived model was ascertained through the cross-validated index, Q^2, from leave-one-out (Q^2_{LOO}) and leave-five-out (Q^2_{L5O}) procedures. The LOO method creates a number of modified data sets by taking away one compound from the parent data set in such a way that each observation has been removed once only. Then one model is developed for each reduced data set, and the response values of the deleted observations are predicted from these models.

The external validation or predictive power of derived model is based on test set compounds. The index r^2_{Test}, representing the squared correlation coefficient between the observed and predicted data of the test-set, has been used to infer the same. A value greater than 0.5 of r^2_{Test} suggests that the model obtained from training set has a reliable predictive power.

Y-randomization

Chance correlations, if any, associated with the CP-MLR models were explored through randomization test [41,42] by repeated scrambling of the biological response. The data sets with scrambled response vector have been reassessed by multiple regression analysis (MRA). The resulting regression equations, if any, with correlation coefficients better than or equal to the one corresponding to the unscrambled response data were counted. Every model has been subjected to 100 such simulation runs. This has been used as a measure to express the percent chance correlation of the model under scrutiny.

Applicability domain

The utility of a QSAR model is based on its accurate predictive ability for new congeners. A model is valid only within its training domain, and new compounds must be assessed as belonging to this domain before the model is applied. The applicability domain is assessed by the leverage values for each compound [43,44]. A Williams plot (the plot of standardized residuals versus leverage values (h) can then be used for an immediate and simple graphical detection of both the response outliers (Y outliers) and structurally influential chemicals (X outliers) in the model. In this plot, the applicability domain is established inside a squared area within ± β.(standard deviations) and a leverage threshold h*. The threshold h* is generally fixed at 3(k+1)/n (n is the number of compounds in the training-set and k is the number of independent descriptors of the model) whereas β = 2 or 3. Prediction must be considered unreliable for compounds with a high leverage value (h > h*). On the other hand, when the leverage value of

a compound is lower than the threshold value, the probability of agreement between predicted and observed values is as high as that for the training set compounds.

RESULTS AND DISCUSSION

Qsar results

Initially, the DPP-IV inhibition activity of titled compounds was investigated with a variety of 0D-, 1D- and 2D-descriptors obtained from DRAGON software. For the development of QSAR models, 20 compounds were considered in training set while 05 (one-fifth of the total) compounds were included in test set for external validation of derived significant models. Though each individual descriptor class is enriched with information corresponding to the activity, different descriptors classes together have led to the models with optimum explained variance. These models were identified in CP-MLR by successively incrementing the filter-3 with increasing number of descriptors (per equation). For this the optimum r-bar value of the preceding level model has been used as the new threshold of filter-3 for the next generation.

The models, in three parameters of the descriptor pool of 90 descriptors, emerged in CP-MLR for the DPP-IV inhibition actions are given in Table 2 as Equations (i) to (vi). The signs of the regression coefficients have indicated the direction of influence of explanatory variables in above models. The positive regression coefficient associated to a descriptor will augment the activity profile of a compound while the negative coefficient will cause detrimental effect to it. The maximum number of descriptors, participated in these models ATS8p, GATS8p, GATS7e, MATS3e and MATS8e, belong to 2D-autocorrelations (2D-AUTO) class. The 2D-AUTO descriptors, ATSke, GATSke and MATSke have their origin in autocorrelation of topological structure of Broto-Moreau, of Moran and of Geary, respectively. The computation of these descriptors involves the summation of different autocorrelation functions corresponding to the different fragment lengths and lead to different autocorrelation vectors corresponding to the lengths of the structural fragments. Also a weighting component in terms of a physicochemical property has been embedded in these descriptors. As a result, these descriptors address the topology of the structure or parts thereof in association with a selected physicochemical property. In these descriptors' nomenclature, the penultimate character, a number, indicates the number of consecutively connected edges considered in its computation and is called as the autocorrelation vector of lag k (corresponding to the number of edges in the unit fragment). The very last character of the descriptor's nomenclature indicates the physicochemical property considered in the weighting component – m for atomic mass, e for atomic Sanderson electronegativity and p for atomic polarizability - for its computation.

Table 2: **CP-MLR models[a] derived in three parameters for the DPP-IV inhibition activity.**

Model	r	s	F	q^2_{LOO}	r^2_{Test}	Eq.

pIC_{50} = 5.090 + 1.709(0.280)JGI4 – 1.150(0.331)ATS8p + 2.074(0.297)GATS8p	0.934	0.242	36.897	0.806	0.279	(i)
pIC_{50} = 5.538 + 0.599(0.248)BELm1 – 0.902(0.364)GATS7e + 1.784(0.358)GATS8p	0.854	0.353	14.465	0.615	0.331	(ii)
pIC_{50} = 7.174 – 1.493(0.444)Uindex + 1.200(0.347)JGI4 – 1.492(0.442)MATS3e	0.850	0.358	13.927	0.533	0.250	(iii)
pIC_{50} = 5.939 – 0.976(0.400)GATS7e + 1.442(0.410)GATS8p + 0.668(0.325)C-002	0.842	0.367	13.021	0.562	0.331	(iv)
pIC_{50} = 6.094 – 0.663(0.327)RBN – 0.725(0.361)GATS7e + 1.848(0.374)GATS8p	0.841	0.368	12.913	0.511	0.191	(v)
pIC_{50} = 5.651 + 1.109(0.418)JGI4 + 1.124(0.524)MATS8e – 0.806(0.400)GATS7e	0.804	0.404	9.812	0.521	0.233	(vi)

[a]The models, in three parameters, emerged from CP-MLR protocol with filter-1 as 0.79, filter-2 as 2.0, filter-3 as 0.5 and filter-4 as $0.3 \leq q^2 \leq 1.0$ with a training set of 20 compounds.

Desciptors, GATS8p and MATS8e both added positively to the inhibitory activity whereas ATS8p, GATS7e and MATS3e contributed negatively to the activity advocating that higher values of descriptors GATS8p and MATS8e and lower values of descriptors ATS8p, GATS7e and MATS3e would be beneficiary to the activity. Constitutional class descriptors are dimensionless or 0D descriptors and are independent from molecular connectivity and conformations. The appeared constitutional class descriptor RBN (number of rotatable bonds) favors the least preference of rotatable bonds.

Descriptor Uindex, corresponds to Balaban U-index, is a topological class descriptor. Topological class descriptors are based on a graph representation of the molecule and are numerical quantifiers of molecular topology obtained by the application of algebraic operators to matrices representing molecular graphs and whose values are independent of vertex numbering or labeling. They can be sensitive to one or more structural features of the molecules such as size, shape, symmetry, branching and cyclicity and can also encode chemical information concering atom type and bond multiplicity. The negative contribution of descriptor Uindex suggested that a lower value of it would be supportive to the activity. The other participated descriptors are JGI4 (from the Galvez topological charge indices), BELm1 (from the modified Burden eigenvalues class, BCUT descriptors) and C-002 (from the atom-centered fragments). GALVEZ class descriptors are 2D-descriptors representing the first 10 eigenvalues of corrected adjacency matrix. BCUT descriptors are also 2D-descriptors representing positive and negative eigenvalues of the adjacency matrix weighting the diagonal elements and atoms. Atom centered fragments (ACF) are molecular descriptors based on the counting of 120 atom centered fragments, as defined by Ghose-Crippen. The 4^{th} order mean Galvez topological charge index (JGI4), the lowest eigenvalue n.1 of Burden matrix/ weighted by atomic masses (BELm1) and CH2R2 type atom centered fragment (C-002) correlated positively to the activity suggested that a higher value of these will augment the activity. However, for all the models mentioned in Table 2 the r^2_{Test} values (<0.5) are inferior to a specified value.

Considering the number of observation in the data set, models with up to four descriptors were explored. A total number of seven such models, sharing 18 descriptors among them, have been obtained through CP-MLR. The shared 18 descriptors along with their brief description, average regression coefficients and total incidences are given in Table 3. Following are the selected four-descriptor models, obtained from CP-MLR, for the DPP-IV inhibitory activity.

Table 3. **Identified descriptors[a] along with their physical meaning, average regression coefficient and incidence[b], in modeling the DPP-IV inhibitory activity.**

S. No.	Descriptor	Descriptor class	Physical meaning	Average regression coefficient (incidence)
1	RBN	Constitutional	Number of rotatable bonds	-1.302(1)
2	PJI2	Topological	2D Petitijean shape index	0.360(1)
3	BIC3	Topological	Bond information content (neighborhood symmetry of 3 order)	3.300(2)
4	BIC5	Topological	Bond information content (neighborhood symmetry of 5 order)	-2.310(1)
5	SRW09	Molecular walk counts	Self- returning walk count of order 09	1.341(1)
6	BELm1	BCUT	Lowest eigenvalue n.1 of Burden matrix/ weighted by atomic masses	0.713(1)
7	BEHv1	BCUT	Highest eigenvalue n.1 of Burden matrix/ weighted by atomic van der Waals volumes	0.821(1)
8	BELe1	BCUT	lowest eigenvalue n.1 of Burden matrix/ weighted by atomic Sanderson electro negativities	0.655(1)
9	BELp2	BCUT	lowest eigenvalue n.2 of Burden matrix/ weighted by atomic polarizabilities	-0.675(1)
10	JGI4	Galvez topological charge indices	Mean topological charge index of order 4	1.993(1)
11	ATS8p	2D autocorrelations	Broto-Moreau autocorrelation of a topological structure - lag8/ weighted by atomic polarizabilities	-0.992(2)
12	GATS7e	2D autocorrelations	Geary autocorrelation of lag-7/ weighted by atomic Sanderson electro negativities	-0.901(4)
13	GATS8p	2D autocorrelations	Moran autocorrelation of lag-8/ weighted by atomic polarizabilities	2.020(5)
14	C-024	Atom-centred fragments	R--CH--R	0.405(1)
15	C-040	Atom-centred fragments	R-C(=X)-X/ R-C#X/X-=C=X	0.853(1)

16	H-052	Atom-centred fragments	H attached to C0(sp3) with 1X attached to next C	2.214(2)
17	MR	Properties	Ghose-Crippen molecular refractivity	-1.058(1)
18	MLOGP	Properties	Moriguchi octanol-water partition coefficient (logP)	1.331(1)

[a]The descriptors are identified from the four parameter models, emerged from CP-MLR protocol with filter-1 as 0.79, filter-2 as 2.0, filter-3 as 0.809 (r-bar of the three parameter model having the highest r^2_{Test}), and filter-4 as $0.3 \leq q^2 \leq 1.0$ with a training set of 20 compounds. [b]The average regression coefficient of the descriptor corresponding to all models and the total number of its incidence. The arithmetic sign of the coefficient represents the actual sign of the regression coefficient in the models.

pIC_{50} =4.722+1.993(0.276) JGI4–1.295(0.300) ATS8p+2.163(0.265) GATS8p+ 0.405(0.173)C-024

n = 20, r = 0.952, s = 0.214, F = 36.718, q^2_{LOO} = 0.832, q^2_{L5O} = 0.805, r^2_{Test}= 0.576 (1)

pIC_{50}=5.165–0.688(0.336) ATS8p–0.690(0.267) GATS7e+2.112(0.344) GATS8p+ 1.331(0.288) MLOGP

n = 20, r = 0.921, s = 0.273, F = 21.086, q^2_{LOO} = 0.715, q^2_{L5O} = 0.704, r^2_{Test}= 0.730 (2)

pIC_{50} = 5.910 – 1.302(0.282) RBN + 3.530(0.486) BIC3 – 2.310(0.364) BIC5 + 1.897(0.247) H-052

n = 20, r = 0.920, s = 0.274, F = 20.846, q^2_{LOO} = 0.561, q^2_{L5O} = 0.713, r^2_{Test}= 0.801 (3)

pIC_{50} = 3.727 +3.070(0.467) BIC3+1.341(0.226) SRW09+0.853(0.243) C-040+ 2.530(0.335) H-052

n = 20, r = 0.907, s = 0.295, F = 17.516, q^2_{LOO} = 0.558, q^2_{L5O} = 0.582, r^2_{Test}= 0.543 (4)

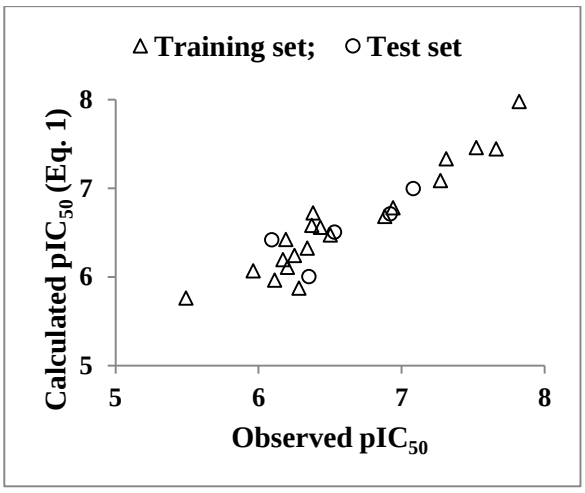
Where n and F represent respectively the number of data points and the F-ratio between the variances of calculated and observed activities. The data within the parentheses are the standard errors associated with regression coefficients. In all above equations, the F-values remained significant at 99% level. The indices q^2_{LOO} and q^2_{L5O} (> 0.5) have accounted for their internal robustness. For all above models the r^2_{Test} values, obtained greater than 0.5, specified that the selected test-set is fully accountable for their external validation.

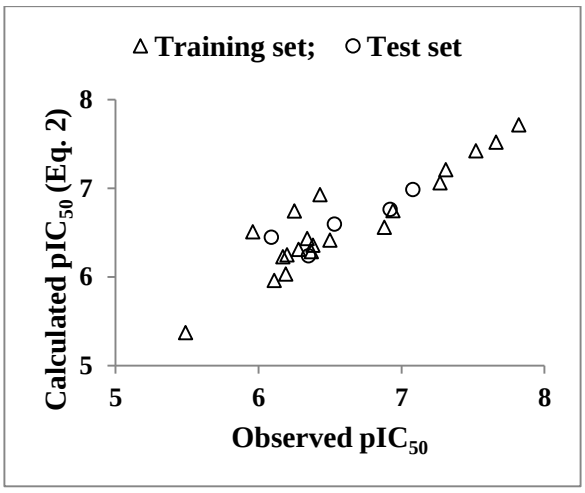
These models are able to estimate up to 90.73 percent of variance in observed activity of the compounds. The derived statistical parameters of these four models have shown that these models are significant. These models were, therefore, used to calculate the activity profiles of all the compounds and are included in Table 4 for the sake of comparison with observed ones. A close agreement between them has been observed. Additionally, the graphical display, showing the variation of observed versus calculated activities is given in Figure 1 to ensure the goodness of fit for each of these four models.

Table 4: Observed, calculated and predicted DPP-IV inhibition activities of (2*S*)-Cyanopyrrolidine analogues.

Cpd.	pIC_{50}[a]								
	Obsd[b].	Eq. (1) Calc.	Pred[c].	Eq. (2) Calc. 2	Pred[c].	Eq. (3) Calc.3	Pred[c].	Eq. (4) Calc. 4	Pred[c].
1	5.49	5.76	5.90	5.37	5.31	5.75	5.98	5.73	5.79
2	6.94	6.78	6.76	6.75	6.72	6.59	6.52	6.64	6.61
3	6.19	6.42	6.53	6.04	6.00	5.84	5.76	5.86	5.80
4[d]	7.08	7.00	-[d]	6.99	-[d]	7.26	-[d]	7.22	-[d]
5	6.88	6.68	6.65	6.56	6.50	7.03	7.11	6.84	6.82
6	6.37	6.58	6.60	6.29	6.18	6.40	6.41	6.78	6.82
7	7.27	7.09	7.07	7.06	7.04	6.93	6.85	7.27	7.27
8[d]	6.09	6.42	-[d]	6.45	-[d]	6.26	-[d]	6.32	-[d]
9	7.31	7.33	7.34	7.21	7.19	7.13	7.09	7.10	7.06
10	7.52	7.46	7.44	7.42	7.40	7.70	7.77	7.60	7.63
11	7.66	7.44	7.34	7.52	7.47	7.70	7.72	7.60	7.58
12	7.82	7.98	8.07	7.72	7.68	7.45	7.18	7.33	6.99
13	6.17	6.20	6.20	6.23	6.24	6.15	6.14	5.88	5.71
14	6.38	6.72	6.80	6.36	6.34	6.15	6.06	6.57	6.68
15	6.20	6.11	5.99	6.25	6.28	6.21	6.21	6.30	6.34
16	6.28	5.88	5.78	6.31	6.32	6.27	6.27	6.81	7.10
17	6.34	6.33	6.32	6.44	6.45	6.56	6.59	6.27	6.26
18	6.50	6.47	6.47	6.42	6.41	6.67	6.71	6.44	6.43
19[d]	6.35	6.00	-[d]	6.24	-[d]	6.14	-[d]	6.42	-[d]
20	6.43	6.56	6.59	6.93	7.16	6.88	6.99	6.55	6.57
21	6.11	5.97	5.88	5.96	5.88	6.06	6.05	5.87	5.83
22[d]	6.92	6.71	-[d]	6.76	-[d]	6.97	-[d]	6.62	-[d]
23	5.96	6.07	6.20	6.51	6.80	6.32	7.14	6.26	6.87
24	6.25	6.24	6.24	6.75	6.84	6.24	6.23	6.36	6.37
25[d]	6.53	6.51	-[d]	6.60	-[d]	6.68	-[d]	6.16	-[d]

[a]IC_{50} represents the concentration of a compound required to bring out 50% inhibition of DPP-IV and the same is expressed as pIC_{50} on molar basis; [b]Taken from ref. [29]; [c]Leave-one-out (LOO) procedure; [d]Compound included in test set.





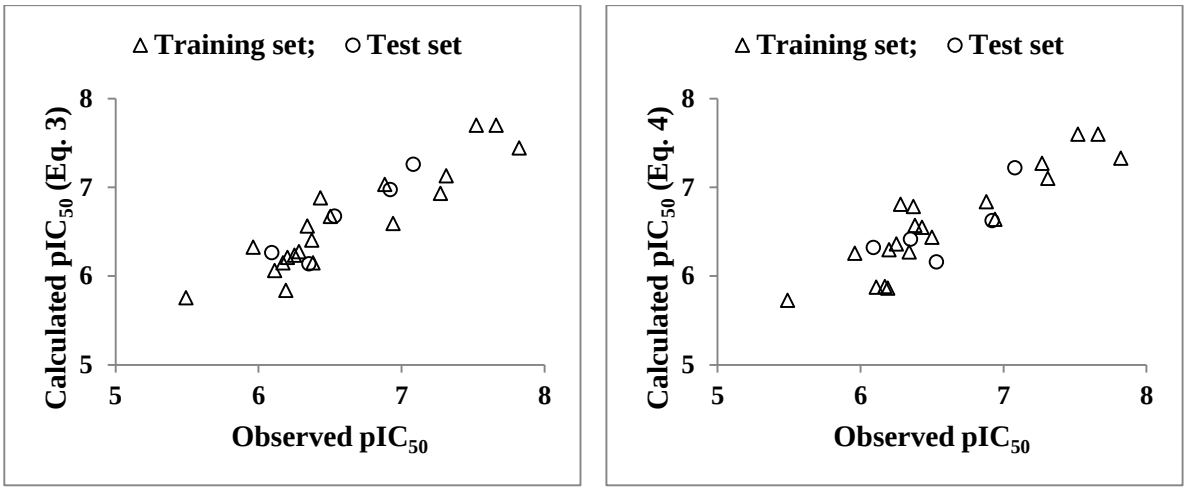


Figure 1: **Plots showing variations of observed versus calculated pIC_{50} values of training and test set compounds.**

The newly emerged descriptors C-024, H-052 and C-040 in these models are atom centered fragments and shown positive correlation to the activity. Therefore, presence of R--CH--R (descriptor C-024), H attached to C0(sp3) with 1X attached to next C (descriptor H-052) and R-C(=X)-X/R-C#X/X-=C=X (descriptor C-040) type atom centered fragments in a molecular structure would enhance the activity. Topological class descriptor BIC3 (bond information content of 3rd order neighborhood symmetry) contributed positively whereas BIC5 (bond information content of 5th order neighborhood symmetry) contributed negatively to the activity revealed that a higher value of descriptor BIC3 and a lower value of descriptor BIC5 would be beneficiary to the activity. Descriptor SRW09 represents self- returning walk count of order 09 is from molecular walk counts class. Molecular walk counts are 2D-descriptors representing self-returning walk counts of different lengths. The descriptor MLOGP is from properties class representing Moriguchi octanol-water partition coefficient (logP). It is evinced from the models that higher values of both of these descriptors (SRW09 and MLOGP) would augment the activity.

CP-MLR analysis has also been performed for the DPP-8 inhibitory activity with the descriptor pool of 84 descriptors with the same test which was used for the DPP-IV inhibitory activity. All the emerged four models in three descriptors are given below through Equations (5) to (8).

$pIC_{50} = 5.283 + 0.863(0.234)MW – 0.745(0.207)X1Av + 0.544(0.123)PJI2$

$n = 15, r = 0.885, s = 0.218, F = 13.264, q^2_{LOO} = 0.552, q^2_{L5O} = 0.578, r^2_{Test} = 0.683$ (5)

$pIC_{50} = 5.419 – 0.536(0.210)X1Av + 0.533(0.129)PJI2 + 0.497(0.148)C\text{-}040$

$n = 15, r = 0.871, s = 0.229, F = 11.623, q^2_{LOO} = 0.507, q^2_{L5O} = 0.612, r^2_{Test} = 0.761$ (6)

$pIC_{50} = 5.300 – 0.778(0.223)X1Av + 0.575(0.130)PJI2 + 0.771(0.231)BEHm8$

$n = 15, r = 0.871, s = 0.230, F = 11.525, q^2_{LOO} = 0.506, q^2_{L5O} = 0.726, r^2_{Test} = 0.646$ (7)

$pIC_{50} = 6.037 + 0.609(0.135)PJI2 – 0.601(0.254)GATS1e – 0.691(0.205)C\text{-}024$

$n = 15, r = 0.860, s = 0.238, F = 10.493, q^2_{LOO} = 0.533, q^2_{L5O} = 0.528, r^2_{Test} = 0.511$ (8)

The derived statistical parameters of these four models have shown that these models are significant and are able to explain up to 78.34 percent of variance in observed DPP8 activity of the compounds. A close agreement between observed and calculated activity has been observed and is given in Table 5 for the sake of comparison. The participated descriptors in above models suggested that higher values of molecular weight (MW, constitutional class descriptor), 2D- Petitijean shape index (PJI2, topological class), highest eigenvalue n.8 of Burden matrix weighted by atomic masses (BEHm8, BCUT descriptor) and presence of R-C(=X)-X/R-C#X/X-=C=X type atom centered fragments (descriptor C-040, atom centered fragment descriptor) would be beneficiary to DPP8 inhibitory activity. Another emerged topological class descriptor X1Av (average valence connectivity index, chi-1), 2D-AUTO class descriptor GATS1e (Geary autocorrelation of lag-1/weighted by atomic Sanderson electro negativities) advocated that a lower value of these descriptors and absence of R--CH--R type fragment (descriptor C-024) would augment the activity.

Table 5: **Observed, calculated and predicted DPP-8 inhibition activities of (2*S*)-Cyanopyrrolidine analogues.**

Cpd.	pIC_{50}[a] Obsd[b].	Eq. (5)		Eq. (6)		Eq. (7)		Eq. (8)	
		Calc.	Pred[c].	Calc. 2	Pred[c].	Calc. 3	Pred[c].	Calc. 4	Pred[c].
1	5.38	5.72	5.82	5.79	5.89	5.64	5.74	5.81	5.89
2	5.45	5.11	5.04	5.16	5.11	5.20	5.16	5.20	5.16
3	5.48	5.60	5.63	5.34	5.32	5.69	5.75	5.50	5.53
4[d]	5.77	5.88	-[d]	5.80	-[d]	6.01	-[d]	6.20	-[d]
5	5.67	5.72	5.73	5.79	5.82	5.90	5.94	5.81	5.84
6	5.85	5.61	5.55	5.62	5.56	5.79	5.77	5.81	5.81
7	5.27	5.12	5.09	5.13	5.10	5.23	5.23	5.20	5.19
8[d]	4.38	5.11	-[d]	5.02	-[d]	5.17	-[d]	5.20	-[d]
9	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
10	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
11	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
12	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
13	6.69	6.40	6.11	6.43	6.11	6.29	6.02	6.21	5.97
14	5.47	5.55	5.70	5.53	5.62	5.45	5.43	5.55	5.56
15	5.7	5.66	5.66	5.91	6.04	5.64	5.63	5.67	5.65
16	5.67	5.87	5.95	5.57	5.54	5.81	5.85	5.69	5.69
17	4.97	5.01	5.02	5.18	5.21	4.99	5.00	4.88	4.84
18	5.62	5.61	5.61	5.48	5.43	5.41	5.33	5.65	5.67
19[d]	4.66	4.92	-[d]	5.02	-[d]	4.84	-[d]	4.89	-[d]
20	5.26	5.49	5.57	5.41	5.46	5.60	5.72	5.58	5.68
21	4.89	4.97	5.00	5.01	5.05	4.88	4.87	4.89	4.90
22[d]	5.08	5.05	-[d]	5.11	-[d]	4.90	-[d]	4.89	-[d]
23	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]	-[e]
24	5.59	5.51	5.48	5.63	5.65	5.46	5.39	5.50	5.46
25[d]	6.07	5.60	-[d]	5.79	-[d]	5.65	-[d]	5.49	-[d]

[a]IC_{50} represents the concentration of a compound required to bring out 50% inhibition of DPP-IV and the same is expressed as pIC_{50} on molar basis; [b]Taken from ref. [29]; [c]Leave-one-out

(LOO) procedure; [d]Compound included in test set. [e]Compound with uncertain activity, not part of data set.

CP-MLR has also been carried out on DPP8 inhibitory activity from the pool of 90 descriptors which was used to find rationales for DPP-IV inhibitory activity. The analysis resulted into 10 models having $r^2_{Test} > 0.5$ and the highest significant four models are listed in Table 6. These are able to estimate up 84.82 percent of variance in observed DPP8 activity of the compounds. The newly appeared descriptors in above models are MR (property class descriptor), N-072 and H-047 (ACF class descriptor), X0Av and X2Av (TOPO class descriptor) and BELp2 (BCUT descriptor). Tabled Equations reveal that lower values of average valence connectivity indices (X0Av and X2Av, chi-0 and chi-2) would be advantageous to enhance the activity. On the other hand, a higher lower value of Ghose-Crippen molecular refractivity (MR) and lowest eigenvalue n.2 of Burden matrix weighted by atomic polarizabilities are incremental to the activity. Counts for certain structural fragments, H attached to C1(sp3) /C0(sp2) (descriptor H-047) and R-CO-N</>N-X=X (descriptor N-072) strongly recommend the presence of such structural features favorable to activity. Thus the descriptors identified for rationalizing the DPP-IV activity give avenues to rationalize the DPP8 inhibitory activity. From the different nature of emerged descriptors in final statistically significant models for DPP IV and DPP8 inhibition actions, it appeared that the mode of actions of titled compounds were different for DPP IV and DPP8 enzyme systems.

Table 6: **Three parameter CP-MLR models for the DPP-8 inhibition activity from the descriptor pool of DPP-IV.**

Model	r	s	F	q^2_{LOO}	r^2_{Test}	Eq.
pIC_{50} = 5.218 – 0.999(0.193)X2Av + 0.523(0.103)PJI2 + 1.075(0.230)MR	0.920	0.182	20.490	0.690	0.545	(vii)
pIC_{50} = 5.510 – 0.737(0.203)X2Av + 0.464(0.118)PJI2 + 0.613(0.161)N-072	0.896	0.208	14.982	0.590	0.755	(viii)
pIC_{50} = 5.405 – 0.865(0.217)X2Av + 0.411(0.125)PJI2 + 0.702(0.195)BELp2	0.889	0.214	13.934	0.539	0.516	(ix)
pIC_{50} = 5.221 – 1.066(0.276)X0Av + 0.410(0.125)PJI2 + 0.871(0.231)H-047	0.888	0.215	13.730	0.562	0.762	(x)

[a]The models, in three parameters, emerged from CP-MLR protocol with filter-1 as 0.79, filter-2 as 2.0, filter-3 as 0.5 and filter-4 as $0.3 \leq q^2 \leq 1.0$ with a training set of 15 compounds.

Applicability domain

On analyzing the applicability domain (AD) in the Williams plot (Figure 2) of the model based on the whole data set (Table 7), none of the compound has been identified as an obvious 'outlier' for the DPP-IV inhibitory activity if the limit of normal values for the Y outliers (response outliers) was set as 3×(standard deviation) units. None of the compounds was found to have leverage (h) values greater than the threshold leverage (h*). For both the training-set and test-set, the suggested model matches the high quality parameters with good fitting power

and the capability of assessing external data. Furthermore, all of the compounds were within the applicability domain of the proposed model and were evaluated correctly.

Table 7: **Models derived for the whole data set (n = 25) for the DPP-IV inhibitory activity in descriptors identified through CP-MLR.**

Model	r	s	F	q^2_{LOO}	Eq.
pIC_{50}=4.707+2.088(0.260)JGI4 1.513(0.230)ATS8p+2.197(0.246)GATS8p+ 0.478(0.158)C-024	0.942	0.212	39.987	0.815	(1a)
pIC_{50}=5.216–0.847(0.249)ATS8p–0.805(0.207)GATS7e+ 2.143(0.298)GATS8p + 1.365(0.254)MLOGP	0.920	0.249	27.603	0.760	(2a)
pIC_{50}=5.902–1.229(0.233)RBN+3.431(0.415)BIC3– 2.257(0.315)BIC5 + 1.866(0.217)H-052	0.919	0.250	27.370	0.658	(3a)
pIC_{50}=3.827+2.966(0.412)BIC3+1.301(0.208)SRW09+0.809(0.223)C-040+ 2.471(0.305)H-052	0.895	0.283	20.219	0.582	(4a)

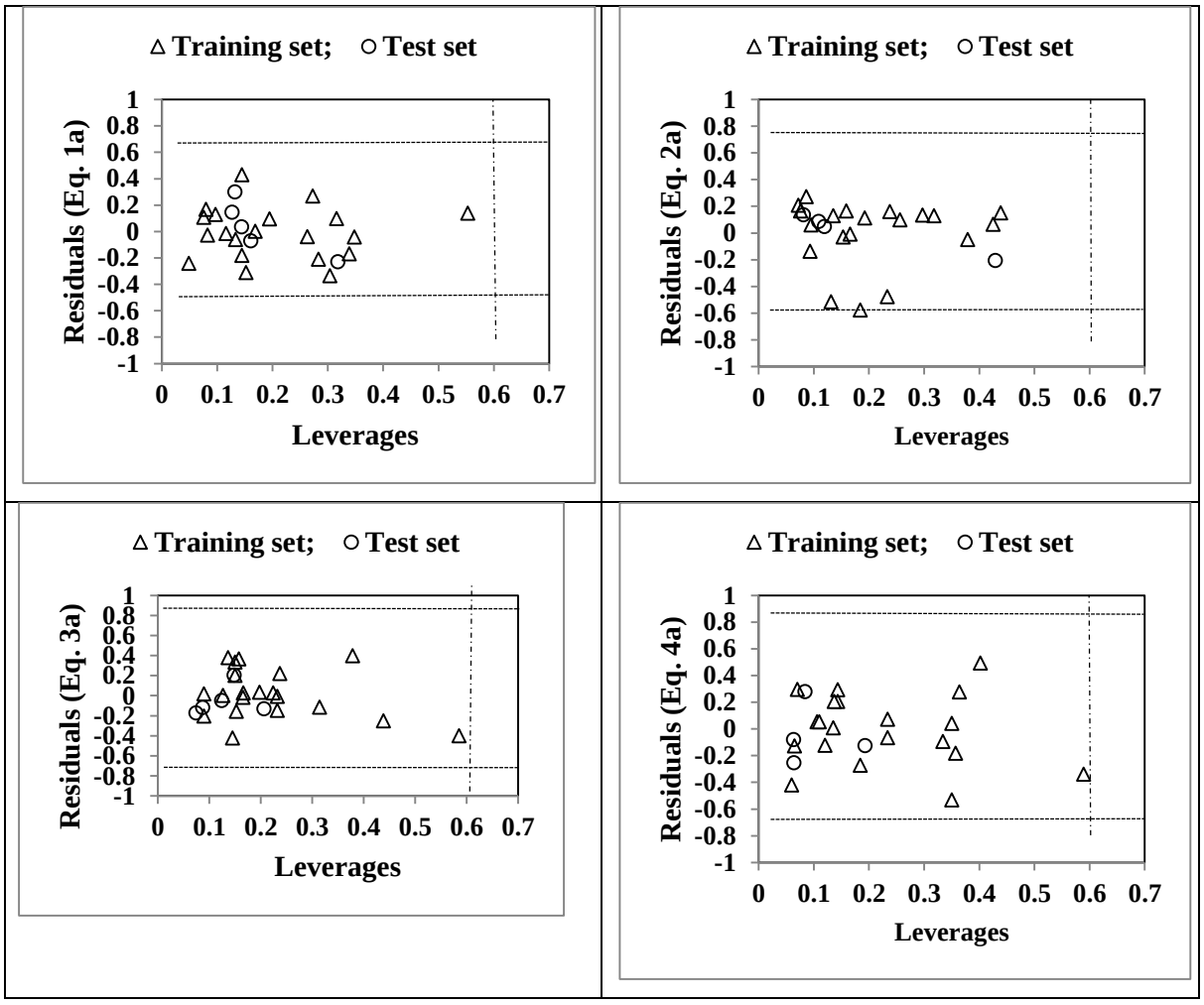


Figure 2. **Williams plot for the training-set and test- set for inhibition activity of DPP4 for the compounds in Table 1. The horizontal dotted line refers to the residual limit (±3×standard deviation) and the vertical dotted line represents threshold leverage h* (= 0.6).**

CONCLUSION

This study has provided a rational approach for the development of (2*S*)-cyanopyrrolidine analogues as DPP-IV inhibitors. The descriptors identified in CP-MLR analysis have highlighted the role of atomic properties in respective lags of 2D-autocorrelations (ATS8p, GATS8p and GATS7e), 4th order mean Galvez topological charge index (JGI4), 3rd and 5th order bond information content of neighborhood symmetry (BIC3 and BIC5) and 9th order self returning walk-count (SRW09) to explain the biological actions of (2*S*)-cyanopyrrolidine analogues as DPP IV inhibitors. Certain structural features or fragments (RBN, C-024, C-040 and H-052) in molecular structures in addition to hydrophobicity (MLOGP) of a molecule have also shown prevalence to optimize the DPP IV inhibitory activity of titled compounds. Applicability domain analysis revealed that the suggested model for DPP IV inhibitory activity matches the high quality parameters with good fitting power and the capability of assessing external data and all of the compounds was within the applicability domain of the proposed model and were evaluated correctly.

REFERENCES

1. Holst, J. J. Glucagon-like peptide-1 (GLP-1): A newly discovered gastrointestinal hormone. Gastroenterology 1994, 107, 1848-1855.
2. Orskov C. Glucagon-like peptide-1, a new hormone of the enteroinsular axis. Diabetologia 1992, 35, 701-711.
3. Drucker, D. J. Glucagon-Like Peptides. Diabetes 1998, 47, 159-169.
4. Drucker, D. J. Minireview: The glucagon-like peptides. Endocrinology 2001, 142, 521-527.
5. Ross, S. A.; Gulve, E. A.; Wang, M. Chemistry and biochemistry of type 2 diabetes. Chem. Rev. 2004, 104, 1255-1282.
6. Ahren, B.; Larsson, H.; Holst, J. J. Effects of glucagon-like peptide-1 on islet function and insulin sensitivity in noninsulin dependent diabetes mellitus. J. Clin. Endocrinol. Metab. 1997, 82, 473-478.
7. Rachman, J.; Barrow, B. A.; Levy, J. C.; Turner, R. C. Near normalization of diurnal glucose concentrations by continuous administration of glucagon-like peptide 1 (GLP-1) in subjects with NIDDM. Diabetologia 1997, 40, 205-211.
8. Nauck, M. A.; Wollschlager, D.; Werner, J.; Holst, J. J.; Orskov, C.; Creutzfeldt, W.; Willms, B. Effects of subcutaneous glucagon-like peptide 1 (GLP-1 [7-36 amide]) in patients with NIDDM. Diabetologia 1996, 39, 1546-1553.
9. Drucker, D. J. Biological actions and therapeutic potential of the glucagon-like peptides. Gastroenterology 2002, 122, 531-544.

10. Wettergren, A.; Schjoldager, B.; Mortensen, P. E.; Myhre, J.; Christiansen, J.; Holst, J. J. Truncated GLP-1 (proglucagon 78-107-amide) inhibits gastric and pancreatic functions in man. Dig. Dis. Sci. 1993, 38, 665-673.
11. Nauck, M. A.; Niedereichholz, U.; Ettler, R.; Holst, J. J.; Orskov, C.; Ritzel, R.; Schmiegel, W. H. Glucagonlike peptide-1 inhibition of gastric emptying outweighs its Insulinotropic effects in healthy humans. Am. J. Physiol. 1997, 273, E981-E988.
12. Nauck, M. A. Is GLP-1 an incretin hormone? Diabetologia 1999, 42, 373-379.
13. Egan, J.; Clocquet, A. R.; Elahi, D. The insulinotropic effect of acute exendin-4-administered to humans: Comparison of nondiabetic state to type 2 diabetes. J. Clin. Endocrinol. Metab. 2002, 87, 1282-1290.
14. Gutniak, M.; Orskov, C.; Holst, J. J.; Ahren, B.; Efendic, S. Antidiabetogenic effect of glucagon-like peptide-1 (7-36)amide in normal subjects and patients with diabetes mellitus. N. Engl. J. Med. 1992, 326, 1316-1322.
15. Nauck, M. A.; Kleine, N.; Orskov, C.; Holst, J. J.; Willms, B.; Creutzfeldt, W. Normalization of fasting hyperglycaemia by exogenous glucagon-like peptide 1 (7-36 amide) in type 2 (non-insulin-dependent) diabetic patients. Diabetologia 1993, 36, 741-744.
16. Gutniak, M. K.; Linde, B.; Holst, J. J.; Efendic, S. Subcutaneous injection of the incretin hormone glucagon like peptide 1 abolishes postprandial glycemia in NIDDM. Diabetes Care 1994, 17, 1039-1044.
17. Villhauer, E. B.; Brinkman, J. A.; Naderi, G. B.; Burkey, B. F.; Dunning, B. E.; Prasad, K.; Mangold, B. L.; Russell. M. E.; Hughes, T. E. 1-[[(3-Hydroxy-1-adamantyl)amino]acetyl]-2-cyano-(S)-pyrrolidine: A potent, selective, and orally bioavailable dipeptidyl peptidase IV inhibitor with antihyperglycemic properties. J. Med. Chem. 2003, 46, 2774-2789.
18. Deacon, C. F.; Ahren, B.; Holst, J. J. Inhibitors of dipeptidyl peptidase IV: A novel approach for the prevention and treatment of Type 2 diabetes? Expert. Opin. Investig. Drugs 2004, 13, 1091-1102.
19. Kim, D.; Wang, L.; Beconi, M.; Eiermann, G. J.; Fisher, M. H.; He, H.; Hickey, G. J.;Kowalchick, J. E.; Leiting, B.; Lyons, K.; Marsilio, F.; McCann, M. E.; Patel, R. A.; Petrov, A.; Scapin, G.; Patel, S. B.; Roy, R. S.; Wu, J. K.; Wyvratt, M. J.; Zhang, B. B.; Zhu, L.; Thornberry, N. A.; Weber, A. E. (2R)-4-Oxo-4-[3-(trifluoromethyl)-5,6-dihydro[1,2,4]-triazolo-[4,3-a]pyrazin-7(8H)-yl]-1-(2,4,5-trifluorophenyl)butan-2-amine: A potent, orally active dipeptidyl peptidase IV inhibitor for the treatment of type 2 diabetes. J. Med. Chem. 2005, 48, 141-151.

20. Ahren, B.; Landin-Olsson, M.; Jansson, P. A.; Svensson, M.; Holmes, D.; Schweizer, A. Inhibition of dipeptidyl peptidase-4 reduces glycemia, sustains insulin levels, and reduces glucagon levels in type 2 diabetes. J. Clin. Endocrinol. Metab. 2004, 89, 2078-2084.
21. Weber, A. E., Kim, D.; Beconi, M.; Eiermann, G.; Fisher, M.; He, H.; Hickey, G.; Jaspal, S.; Leiting, B.; Lyons, K.; Marsilio, F.; McCann, P.; Moller, D. E.; Patel, R.; Petrov, A.; Pryor, K.; Sinha Roy, R.; Wu, J. K.; Wyvratt, M.; Zhang, B. B.; Thornberry, N. A. MK- 0431 is a potent, selective dipeptidyl peptidase IV inhibitor for the treatment of type 2 diabetes. Presented at the American Diabetes Association 64th Scientific Session, Orlando, FL, June 4-8, 2004, Abstract 633-P.
22. Yaron, A.; Naider, F. Proline-Dependent Structural and Biological Properties of Peptides and Proteins. Crit. Rev. Biochem. Mol. Biol.1993, 28, 31-81.
23. Mentlein, R.; Gallwitz, B.; Schmidt, W. E. Dipeptidyl-peptidase IV hydrolyses gastric inhibitory polypeptide, glucagon-like peptide-1(7-36)amide, peptide histidine methionine and is responsible for their degradation in human serum. Eur. J. Biochem. 1993, 214, 829-835.
24. Kieffer, T. J.; McIntosh, C. H. S.; Pederson, R. A. Degradation of glucose-dependent Insulinotropic polypeptide and truncated glucagon-like peptide 1 in vitro and in vivo by dipeptidyl peptidase IV. Endocrinology 1995, 136, 3585-3596.
25. Deacon, C. F.; Nauck, M. A.; Toft-Nielsen, M.; Pridal, L.; Willms, B.; Holst, J. J. Both subcutaneously and intravenously administered glucagon-like peptide I are rapidly degraded from the NH2-terminus in type II diabetic patients and in healthy subjects. Diabetes 1995, 44, 1126-1131.
26. Ashworth, D. M.; Atrash, B.; Baker, G. R.; Baxter, A. J.; Jenkins, P. D.; Jones, D. M.; Szelke, M. 4-Cyanothiazolidides as very potent, stable inhibitors of dipeptidyl peptidase IV. Bioorg. Med. Chem. Lett. 1996, 6, 2745-2748.
27. Ashworth, D. M.; Atrash, B.; Baker, G. R.; Baxter, A. J.; Jenkins, P. D.; Jones, D. M.; Szelke, M. Bioorg. 2-Cyanopyrrolidides as potent, stable inhibitors of dipeptidyl peptidase IV. Med. Chem. Lett. 1996, 6, 1163-1166.
28. Ahren, B.; Gomis, R.; Standl, E.; Mills, D.; Schweizer, A. A. prolonged efficacy of LAF237 in patients with Type 2 diabetes (T2DM) inadequately controlled with metformin. 64th Scientific Sessions of the American Diabetes Association. Orlando, FL, 2004, Late-breaking 7-LB (Abstract).
29. Tsu, H.; Chen, X.; Chen, C.-T.; Lee, S.-J.; Chang, C.-N.; Kao, K.-H.; Coumar, M. S.;Yeh, Y.-T.; Chien,C.-H.; Wang, H.-S.; Lin, K.-T.; Chang, Y.-Y.; Wu, S.-H.; Chen, Y.-S.; Lu, I-L.; Wu, S.-Y.; Tsai, T.-Y.; Chen, W.-C.; Hsieh, H.-P.; Chao, Y.-S.; Jiaang

W.-T. 2- [3-[2-[(2S)-2-Cyano-1-pyrrolidinyl]-2-oxoethylamino]-3-methyl-1-oxobutyl]-1,2,3,4-tetrahydroisoquinoline: A Potent, Selective, and Orally Bioavailable Dipeptide-Derived Inhibitor of Dipeptidyl Peptidase IV. J.Med. Chem. 2006, 49, 373-380.

30. SYSTAT, Version 7.0; SPSS Inc 444 North Michigan Avenue, Chicago IL, 60611.
31. Chemdraw ultra 6.0 and Chem3D ultra, Cambridge Soft Corporation, Cambridge, USA. http://www.cambridgesoft.com
32. Dragon software (version 1.11-2001) by Todeschini R.; Consonni V. Milano, Italy. http//www.talete.mi.it/dragon.htm
33. Prabhakar, Y. S. A combinatorial approach to the variable selection in multiple linear regression: analysis of Selwood et al. Data Set-a case study. QSAR Comb. Sci. 2003, 22,583-595.
34. Golbraikh, A.; Tropsha, A. Beware of q2! J. Mol. Graph. Model. 2002, 20, 269-276.
35. Sharma, S.; Prabhakar, Y. S.; Singh, P.; Sharma, B. K. QSAR study about ATP-sensitive potassium channel activation of cromakalim analogues using CP-MLR approach, Eur J Med Chem 2008, 43, 2354-2360.
36. Sharma, S.; Sharma, B. K.; Sharma, S. K.; Singh, P.; Prabhakar, Y. S. Topological Descriptors in modeling the agonistic activity of human A3 adenosine receptor ligands: The derivatives of 2-Chloro-N^6-substituted-4'-thioadenosine-5'-uronamide, Eur. J. Med. Chem. 2009, 44, 1377-1382.
37. Sharma, B. K.; Pilania, P.; Singh, P.; Prabhakar, Y. S. Combinatorial protocol in multiple linear regression/partial least-squares directed rationale for the caspase-3 inhibition activity of isoquinoline-1,3,4-trione derivatives. SAR QSAR Environ. Res. 2010, 21, 169-185.
38. Sharma, B. K.; Singh, P.; Sarbhai, K.; Prabhakar, Y. S. A quantitative structure-activity relationship study on serotonin 5-HT6 receptor ligands: Indolyl and piperidinyl sulphonamides, SAR QSAR Environ. Res. 2010, 21, 369-388.
39. Sharma, B. K.; Pilania, P.; Sarbhai, K.; Singh, P.; Prabhakar, Y. S. Chemometric descriptors in modeling the carbonic anhydrase inhibition activity of sulfonamide and sulfamate derivatives. Mol. Divers. 2010, 14, 371-384.
40. Sharma, B. K.; Singh, P.; Shekhawat, M.; Sarbhai, K.; Prabhakar, Y. S. Modelling of serotonin reuptake inhibitory and histamine H_3 antagonistic activity of piperazine and diazepane amides: QSAR rationales for co-optimization of the activity profiles. SARQSAR Environ. Res. 2011, 22, 365-383

41. So, S. –S.; Karplus, M. Three-dimensional quantitative structure activity relationships from molecular similarity matrices and genetic neural networks. 1. Method and validations. J. Med. Chem. 1997, 40, 4347-4359.
42. Prabhakar, Y. S.; Solomon, V. R.; Rawal, R. K.; Gupta, M. K.; Katti, S. B. CP-MLR/PLS directed structure–activity modeling of the HIV-1 RT inhibitory activity of 2,3-diaryl-1,3-thiazolidin-4-ones. QSAR Comb. Sci. 2004, 23, 234-244.
43. Gramatica, P. Principles of QSAR models validation: internal and external. QSAR Comb.Sci. 2007, 26, 694-701.
44. Eriksson, L.; Jaworska, J.; Worth, A. P.; Cronin, M. T. D.; McDowell, R. M.; Gramatica, P. Methods for reliability and uncertainty assessment and for applicability evaluations of classification and regression-based QSARs. Env. Health Persp. 2003, 111, 1361-1375.

GSC Biological and Pharmaceutical Sciences, 2020, 11(03), 130-144

Available online at GSC Online Press Directory

GSC Biological and Pharmaceutical Sciences

e-ISSN: 2581-3250, CODEN (USA): GBPSC2

Journal homepage: https://www.gsconlinepress.com/journals/gscbps

(RESEARCH ARTICLE)

QSAR rationales for the dipeptidyl peptidase-4 (DPP-4) inhibitors: The imidazolopyrimidine amides

Parihar Raghuraj and Sharma Brij Kishore *

Department of Chemistry, Government College, Bundi-323 001 (Rajasthan), India.

Publication history: Received on 09 June 2020; revised on 18 June 2020; accepted on 19 June 2020

Article DOI: https://doi.org/10.30574/gscbps.2020.11.3.0169

Abstract

The DPP4 inhibition activity of imidazolopyrimidine amides has been quantitatively analyzed in terms of chemometric descriptors. The statistically validated QSAR models provided rationales to explain the inhibition activity of these congeners. The descriptors identified through CP-MLR analysis have highlighted the role of mean electrotopological state (Ms), number of double bonds in molecular structure (nDB), 2D Petitijean shape index (PJI2), Moran autocorrelation of lag-2/weighted by atomic polarizabilities (MATS2p), Moran autocorrelation of lag-6 and lowest eigenvalue n.5 of Burden matrix /weighted by atomic Sanderson electronegativities (MATS6e and BELe5), lowest eigenvalue n.3 and highest eigenvalue n.1 of Burden matrix/weighted by atomic van der Waals volumes (BELv3 and BEHv1). In addition to these 2nd order mean Galvez topological charge index (JGI2), number of ring tertiary C(sp3) (nCrHR) and R--CR--X type structural fragments (C-028) have also shown prevalence to model the inhibitory activity.

From statistically validated models, positive contribution of descriptors Ms, PJI2, JGI2, MATS2p, BELe5, BELv3 and BEHv1 suggested that higher values of these are conducive in improving the DPP4 inhibition activity. On the other hand, negative contribution of descriptors nDB, C-028, nCrHR and MATS6e advocated that absence of number of double bonds (nDB), R--CR--X type structural fragment (C-028), number of ring tertiary C(sp3) (nCrHR) and lower value of descriptor MATS6e would be advantageous. PLS analysis has confirmed the dominance of the CP-MLR identified descriptors and applicability domain analysis revealed the acceptable predictability of suggested models. All the compounds are within the applicability domain of the proposed models and were evaluated correctly.

Keywords: Quantitative structure-activity relationship (QSAR); DPP-4 inhibitors; Combinatorial protocol in multiple linear regression (CP-MLR) analysis; Chemometric descriptors; Imidazolopyrimidine amides.

1. Introduction

Therapeutics based on Glucogon-like peptide-1 (GLP-1[a]) is among the novel and promising targets to cure type 2 diabetes [1-3]. The active and natural form of GLP-1, the incretin hormone GLP-1 (7-36), is secreted from intestinal L-cells after the intake of meals. The stimulation of insulin secretion, inhibition of glucogon release, delay in gastric emptying and promotion of β-cell trophism in intestinal L-cells are advantageous to glucose homeostasis in both the animal models and human [4,5]. Studies revealed that GLP-1 levels are noticeably reduced in type 2 diabetics and exogenous infusion of it may lead to normal insulin response to glucose [6-8] and this fact is the basis for GLP-1 and its analogeus as novel treatments of type 2 diabetes. One such example of a GLP-1 analogue is exenatide [9,10]. Half-maximal effective concentration of 10 pM of the most potent incretin hormone, GLP-1 (7-36), is required to show its effects on pancreatic *β*-cells [11]. The biological functions of GLP-1 (7-36) are exerted through circulation and binding to the GLP-1 receptor that is highly expressed in pancreatic *β*-cells. After secretion GLP-1 (7-36) is rapidly degraded by DPP4 (EC 3.4.14.5) to afford inactive GLP-1 (9-36) under normal physiological conditions. The apparent half-life for GLP-1 (7-36) in this quick inactivation process is 60-90s. It is evinced that due to this natural degradation mechanism

* Corresponding author: Sharma Brij Kishore

less than 50% of released active GLP-1 (7-36) can reach circulation [12]. Thus it is apparent that a DPP4 inhibitor can prevent degradation of and lead to potentiation of GLP-1 and further improve glucose and insulin homeostasis [13,14].

DPP4, ubiquitously expressed throughout the body, is a nonclassical and sequence-specific serine protease. Membrane-bound DPP4 is highly expressed in the endothelium of the capillary bed in close proximity to intestinal L-cells where secretion of GLP-1 takes place. The other form which circulates in plasma is soluble form of DPP4 plays a little role in the cleavage of GLP-1 [15,16]. Vildagliptin [17], sitagliptin [18], saxagliptin [19] and alogliptin [20] are examples of small molecule DPP4 inhibitors which have demonstrated ability to lower blood glucose and HbA1c levels and to improve glucose tolerance in type 2 diabetic patients [21]. Several novel series of azolopyrimidine amines, containing an aromatic or heteroaromatic group on the azolo ring, as potent and selective DPP4 inhibitors were reported in view of medicinal chemistry efforts to discover novel scaffolds [22]. The substitution of aromatic or heteroaromatic group on the azolo ring in these compounds showed enhancement in the binding affinity to DPP4 but displayed high levels of the human ether-à-go-go related gene (hERG) and sodium channel inhibition.

As an attempt to minimize undesired hERG and sodium channel activities a novel series of imidazolopyrimidine amides as a highly potent and selective class of DPP4 inhibitors has been reported by Meng *et al*. [23]. The general structure of these analogues is shown in Figure 1 and structural variations are given in Table 1.

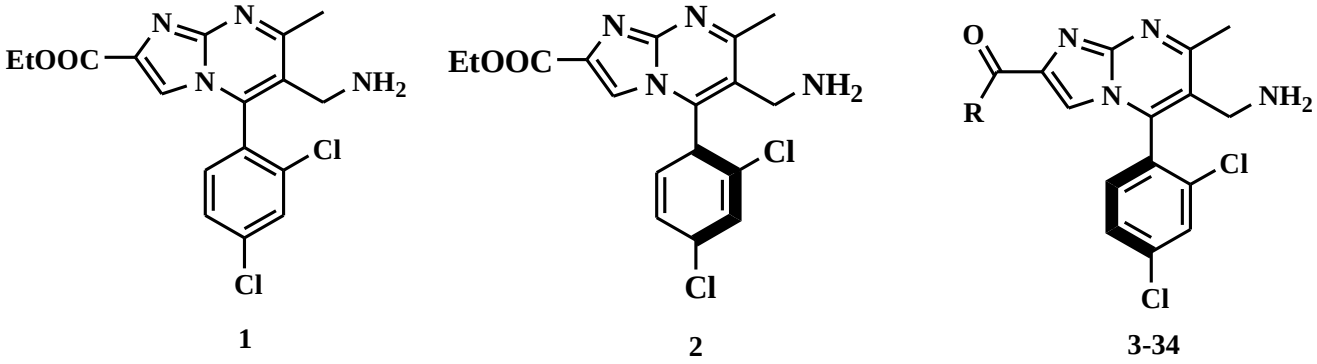


Figure 1 General structures of imidazolopyrimidine derivatives

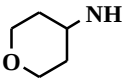
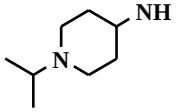
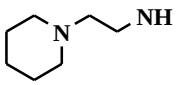
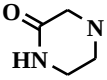
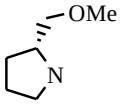
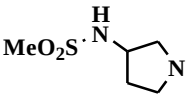
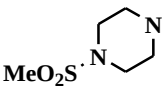
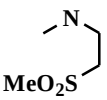
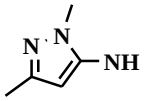
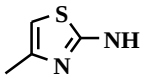
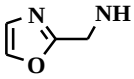
In the present communication a 2D-quantitative SAR (2D-QSAR) has been conducted to provide the rationale for drug-design and to explore the possible mechanism of the action. In the congeneric series, where a relative study is being carried out, the 2D-descriptors may play important role in deriving the significant correlations with biological activities of the compounds. The novelty and importance of a 2D-QSAR study is due to its simplicity for the calculations of different descriptors and their interpretation (in physical sense) to explain the inhibition actions of compounds at molecular level.

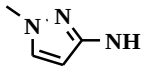
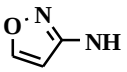
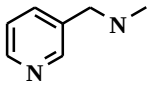
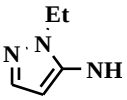
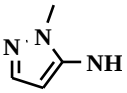
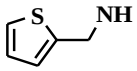
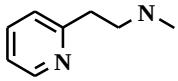
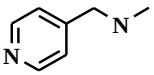
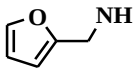
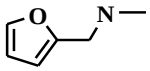
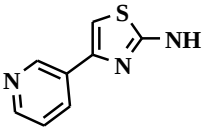
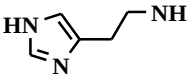
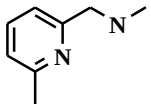
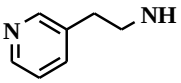
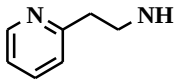
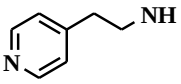
2. Material and methods

2.1. Data-set

For present work the imidazolopyrimidine amides (Table 1), along with their in vitro human DPP4 inhibition activity, have been taken from the literature [23]. The inhibition activity reported in terms of K_i is expressed as pK_i on a molar basis and considered as the dependent variable for the present quantitative analysis.

Table 1 Structural variations and DPP-4 binding affinities of imidazolopyrimidine amides (see fig. 1 for general structure)

Cpd.	R	$pK_i(M)$[a]					
		Obsd[b].	Calculated				
			Eq. (7)	Eq. (8)	Eq. (9)	Eq. (10)	PLS
1	racemic	9.15	8.98	8.94	9.01	9.06	8.99
2	chiral	7.93	8.08	8.03	7.93	8.10	7.83
3[c]	OEt (chiral)	9.40	8.98	8.94	9.01	9.06	8.99
4		8.62	8.81	8.80	8.77	8.64	8.94
5		8.54	8.45	8.46	8.43	8.68	8.75
6		8.66	8.47	8.68	8.50	8.32	8.50
7[c]		8.51	8.40	8.42	8.52	8.29	8.44
8		8.30	8.26	8.30	8.18	8.26	8.22
9		8.06	8.30	8.33	8.33	8.28	8.14
10		9.00	8.82	8.95	8.95	8.72	8.81
11		8.51	8.52	8.53	8.63	8.32	8.59
12		9.70	9.52	9.55	9.49	9.59	9.56
13		9.30	9.37	9.41	9.33	9.42	9.39
14		9.52	9.70	9.54	9.76	9.56	9.56
15		9.05	8.75	8.73	8.81	8.74	8.87
16		8.18	8.43	8.25	8.43	8.56	8.21
17[c]		8.70	8.54	8.49	8.91	8.65	8.71
18		9.22	9.22	9.19	9.01	8.86	9.10

19		8.74	8.70	8.68	8.88	8.72	8.82
20		8.42	8.73	8.91	8.59	8.58	8.59
21[c]		8.59	8.49	8.67	8.63	8.40	8.58
22		8.96	8.75	8.73	8.62	9.03	8.94
23		8.70	8.74	8.93	8.53	8.97	8.83
24		8.49	8.44	8.27	8.48	8.56	8.51
25[c]		9.15	9.18	8.98	9.04	8.78	8.85
26		8.59	8.46	8.65	8.61	8.47	8.60
27[c]		8.4	8.75	8.56	8.69	8.67	8.50
28		8.52	8.70	8.51	8.86	8.64	8.56
29		8.30	8.56	8.55	8.32	8.67	8.27
30		8.33	8.06	8.22	8.38	8.30	8.48
31		8.27	8.57	8.40	8.61	8.30	8.30
32		8.82	8.68	8.67	8.68	8.83	8.81
33		8.85	8.68	8.50	8.61	8.55	8.56
34[c]		8.89	8.68	8.67	8.68	8.76	8.77

[a]On molar basis; [b]Taken from reference [23]; [c]Compounds in test set.

For modeling purpose, the complete data-set was divided into training- and test-sets. The training-set was used to derive statistical significant models while the test-set, consisting nearly 20% of total compounds, was employed to validate such models. The selection of test-set compounds was made through SYSTAT [24] using the single linkage hierarchical cluster procedure involving the Euclidean distances of the binding activity, pK_i values. The test-set compounds were selected from the generated cluster tree in such a way to keep them at a maximum possible distance from each other. In SYSTAT, by default, the normalized Euclidean distances are computed to join the objects of cluster. The normalized distances are root mean-squared distances. The single linkage uses distance between two closest

members in clustering. It generates long clusters and provides scope to choose objects at intervals. Due to this reason, a single linkage clustering procedure was applied.

2.2. Molecular descriptors

The structures of the compounds (Table 1), under study, have been drawn in 2D ChemDraw [25] and were converted into 3D objects using the default conversion procedure implemented in the CS Chem3D Ultra. The generated 3D-structures of the compounds were subjected to energy minimization in the MOPAC module, using the AM1 procedure for closed shell systems, implemented in the CS Chem3D Ultra. This will ensure a well-defined conformer relationship across the compounds of the study. All these energy minimized structures of respective compounds have been ported to DRAGON software [26] for computing the descriptors corresponding to 0D-, 1D-, and 2D-classes. The combinatorial protocol in multiple linear regression (CP-MLR) [27] analysis and partial least-squares (PLS) [28-30] procedures have been used in the present work for developing QSAR models. A brief description of the computational procedure is given below.

2.3. Model development

The CP-MLR is a 'filter'-based variable selection procedure for model development in QSAR studies. Its procedural aspects and implementation are discussed in some of our publications [31-36]. The thrust of this procedure is in its embedded four 'filters'. They are briefly as follows: filter-1 seeds the variables by way of limiting inter-parameter correlations to predefined level (upper limit ≤ 0.79); filter-2 controls the variables entry to a regression equation through t-values of coefficients (threshold value ≥ 2.0); filter-3 provides comparability of equations with different number of variables in terms of square root of adjusted multiple correlation coefficient of regression equation, r-bar; filter-4 estimates the consistency of the equation in terms of cross-validated r^2 or q^2 with leave-one-out (LOO) cross-validation as default option (threshold value $0.3 \leq q^2 \leq 1.0$). All these filters make the variable selection process efficient and lead to a unique solution. In order to collect the descriptors with higher information content and explanatory power, the threshold of filter-3 was successively incremented with increasing number of descriptors (per equation) by considering the r-bar value of the preceding optimum model as the new threshold for next generation. Furthermore, in order to discover any chance correlations associated with the models recognized in CP-MLR, each cross-validated model has been put to a randomization test [37,38] by repeated randomization of the activity to ascertain the chance correlations, if any, associated with them. For this, every model has been subjected to 100 simulation runs with scrambled activity. The scrambled activity models with regression statistics better than or equal to that of the original activity model have been counted, to express the percent chance correlation of the model under scrutiny.

To support the findings, a partial least squares (PLS) analysis has been carried out on descriptors identified through CP-MLR. The study facilitates the development of a 'single window' structure-activity model and help to categorize the potentiality of identified descriptors in explaining the DPP4 inhibition activity profiles of the compounds. It also gives an opportunity to make a comparison of the relative significance among the descriptors. The fraction contributions obtainable from the normalized regression coefficients of the descriptors allow this comparison within the modeled activity.

2.4. Applicability domain

The utility of a QSAR model is based on its accurate prediction ability for new compounds. A model is valid only within its training domain and new compounds must be assessed as belonging to the domain before the model is applied. The applicability domain is assessed by the leverage values for each compound [39,40]. The Williams plot (the plot of standardized residuals versus leverage values, h) can then be used for an immediate and simple graphical detection of both the response outliers (Y outliers) and structurally influential chemicals (X outliers) in the model. In this plot, the applicability domain is established inside a squared area within ±x(s.d.) and a leverage threshold h^*. The threshold h^* is generally fixed at $3(k + 1)/n$ (n is the number of training-set compounds and k is the number of model parameters) whereas x = 2 or 3. Prediction must be considered unreliable for compounds with a high leverage value ($h > h^*$). On the other hand, when the leverage value of a compound is lower than the threshold value, the probability of accordance between predicted and observed values is as high as that for the training-set compounds.

3. Results and discussion

3.1. QSAR results

For the compounds in Table 1, a total number of 479 descriptors belonging to 0D- to 2D- classes of DRAGON have been computed and were subjected to CP-MLR analysis. All the 34 compounds of data set were further divided into training-set and test-set. Seven compounds (nearly 20% of total population) have been selected for test-set through SYSTAT. The identified test-set was then used for external validation of models derived from remaining twenty seven compounds in the training-set. The squared correlation coefficient between the observed and predicted values of compounds from test-set, r^2_{Test}, was calculated to explain the fraction of explained variance in the test-set which is not part of regression/model derivation. It is a measure of goodness of the derived model equation. A high r^2_{Test} value is always good. But considering the stringency of test-set procedures, often r^2_{Test} values in the range of 0.5 to 0.6 are regarded as logical models. Following the strategy to explore only predictive models, CP-MLR resulted one model in three descriptors, five models in four descriptors and sixteen models in five descriptors at upper limit of filter-1. The highest significant of them, in statistical sense, are given through Equations (1)-(10):

pKi= 7.583 + 1.556(0.241)Ms+0.716(0.215)BELe5+ 0.729 (0.241)MATS2p

n = 27, r = 0.812, s = 0.271, F = 14.904, q^2_{LOO} = 0.515, q^2_{L50} = 0.504, r^2_{Test}= 0.125 (1)

pKi= 7.267 +1.503(0.229)Ms + 0.300(0.121)PJI2 + 0.985(0.256)BELv3

+ 0.738(0.223)MATS2p

n = 27, r = 0.844, s = 0.254, F = 13.731, q^2_{LOO} = 0.565, q^2_{L50} = 0.516, r^2_{Test}= 0.632 (2)

pKi = 8.043 – 0.677(0.253)Mv + 2.308(0.381)Ms – 0.880(0.255)nDB

+ 0.635(0.248)BELv3

n = 27, r = 0.836, s = 0.260, F = 12.856, q^2_{LOO} = 0.524, q^2_{L50} = 0.537, r^2_{Test}= 0.527 (3)

pKi = 7.935 + 2.678(0.448)Ms – 0.806(0.249)nDB – 0.619 (0.246)IC1

+ 0.575(0.199)BELe5

n = 27, r = 0.832, s = 0.263, F = 12.380, q^2_{LOO} = 0.507, q^2_{L50} = 0.594, r^2_{Test}= 0.525 (4)

pKi = 8.031 + 1.239(0.235)Ms + 0.315(0.127)PJI2 + 0.991(0.272)BELv3

– 0.679(0.232)GATS2v

n = 27, r = 0.831, s = 0.264, F = 12.295, q^2_{LOO} = 0.501, q^2_{L50} = 0.513, r^2_{Test}= 0.501 (5)

pKi = 8.080 + 1.325(0.227)MAXDN – 0.556(0.219)BELm8 + 0.706(0.345)BELv3

+ 0.957(0.256)MATS2p

n = 27, r = 0.821, s = 0.270, F = 11.449, q^2_{LOO} = 0.503, q^2_{L50} = 0.525, r^2_{Test}= 0.513 (6)

pKi = 7.360 + 1.914(0.288)Ms – 0.744(0.185)nDB + 0.214(0.101)PJI2

+ 1.110(0.215)BELv3 + 0.765(0.192)JGI2

n = 27, r = 0.901, s = 0.210, F = 18.170, q^2_{LOO} = 0.651, q^2_{L50} = 0.600, r^2_{Test}= 0.548 (7)

pKi = 7.644 + 1.753(0.277)Ms – 0.774(0.188)nDB + 1.059(0.211)BELv3

+ 0.729(0.196)JGI2 – 0.357(0.171)C-028

n = 27, r = 0.900, s = 0.211, F = 18.042, q^2_{LOO} = 0.667, q^2_{L50} = 0.690, r^2_{Test}= 0.579 (8)

pKi = 6.629 + 1.573(0.194)Ms + 0.331(0.102)PJI2 + 0.798(0.249)BEHv1

+ 0.705(0.232)BELv3 + 0.603(0.191)MATS2p

n = 27, r = 0.898, s = 0.213, F = 17.656, q^2_{LOO} = 0.644, q^2_{L50} = 0.696, r^2_{Test}= 0.616 (9)

pKi = 8.310 + 1.327(0.208)Ms + 0.265(0.107)PJI2 + 0.669(0.172)BELe5

– 0.723(0.208)MATS6e – 0.771(0.199)nCrHR

n = 27, r = 0.883, s = 0.227, F = 14.972, q^2_{LOO} = 0.569, q^2_{L50} = 0.550, r^2_{Test}= 0.511 (10)

where n and F represent respectively the number of data points and the F-ratio between the variances of calculated and observed activities. The data within the parentheses are the standard errors associated with regression coefficients. In all above equations, the F-values remained significant at 99% level. The indices q^2_{LOO} and q^2_{L50} (> 0.5) have accounted for their internal robustness. For all above models except equation (1) the r^2_{Test} values, obtained greater than 0.5, specified that the selected test-set is fully accountable for their external validation. The descriptors, in all above models, have been scaled between the intervals 0 to 1 [41] to ensure that a descriptor will not dominate simply because it has larger or smaller pre-scaled value compared to the other descriptors. In this way, the scaled descriptors would have equal potential to influence the QSAR models. The signs of the regression coefficients have indicated the direction of influence of explanatory variables in above models. The positive regression coefficient associated to a descriptor will augment the activity profile of a compound while the negative coefficient will cause detrimental effect to it.

Though Equations (1)-(10) emerged as significant predictive models but Equations (7)-(10) remained statistically more efficient. The later four models, involving five descriptors in each, could estimate up to 81.22 percent of variance in observed activity of the compounds. In fact, a total number of sixteen such models, sharing 19 descriptors among them, have been obtained through CP-MLR and the most significant four of them have been documented through Equation (7)-(10). The shared 19 descriptors along with their brief description, average regression coefficients and total incidences are given in Table 2.

Besides listed descriptors in Table 2, the other identified descriptors Mv is from constitutional and MAXDN is from topological class. The Mv represents mean atomic Van der Waals volume (scaled on carbon atom) (Equation 3) and MAXDN is maximal electrotopological negative variation (Equation 6). The further discussion is, however, based on the highest significant Equations (7)-(10). The derived statistical parameters of these four models have shown that these models are significant. These models were, therefore, used to calculate the activity profiles of all the compounds and are included in Table 1 for the sake of comparison with observed ones. A close agreement between them has been observed. Additionally, the graphical display, showing the variation of observed versus calculated activities is given in Figure 2 to ensure the goodness of fit for each of these four models.

Descriptors Ms (mean electrotopological state) and nDB (number of double bonds in molecular structure) belong to constitutional class. From the sign of regression coefficients it is evident that higher value of mean electrotopological state (descriptor Ms) and lower number of double bonds (descriptor nDB) are helpful to augment the activity. The descriptor PJI2 participated in these models is topological class descriptor and represents 2D Petitijean shape index. The positive sign of regression coefficient of descriptor PJI2 suggest that a higher value of this descriptor is beneficiary to the DPP4 inhibition activity. The descriptors MATS2p (Moran autocorrelation of lag-2/weighted by atomic polarizabilities) and MATS6e (Moran autocorrelation of lag-6/weighted by atomic Sanderson electronegativities) are 2D autocorrelation descriptors. It is evinced from the models mentioned above the descriptor MATS2p contributed positively and descriptor MATS6e negatively to the activity. Thus a higher value of descriptor MATS2p and a lower value of descriptor MATS6e will be supportive to enhance the inhibition activity.

Table 2 Identified descriptors[a] along with their physical meaning, average regression coefficient and incidence[b], in modeling the DPP-4 binding affinity.

S. No.	Descriptor	Descriptor class	Physical meaning	Average regression coefficient (incidence)
1	Ms	Constitutional	Mean electrotopological state	1.835(15)
2	nDB	Constitutional	Number of double bonds	-0.685(11)
3	HNar	Topological	Narumi harmonic topological index	-1.304(2)
4	PJI2	Topological	2D Petitijean shape index	0.280(7)
5	IC1	Topological	information content index of neighborhood symmetry of 1-order	-0.585(1)
6	BELm8	BCUT	Lowest eigenvalue n.8 of Burden matrix/ weighted by atomic masses	-0.479(1)
7	BEHv1	BCUT	Highest eigenvalue n.1 of Burden matrix/ weighted by atomic van der Waals volumes	0.989(2)
8	BELv3	BCUT	lowest eigenvalue n.3 of Burden matrix/ weighted by atomic van der Waals volumes	0.935(12)
9	BELe5	BCUT	lowest eigenvalue n.5 of Burden matrix/ weighted by atomic Sanderson electronegativities	0.653(6)
10	JGI2	Galvez topological charge indices	Mean topological charge index of order 2	0.747(2)
11	MATS5e	2D autocorrelations	Moran autocorrelation of lag-5/ weighted by atomic Sanderson electronegativities	-0.597(1)
12	MATS6e	2D autocorrelations	Moran autocorrelation of lag-6/ weighted by atomic Sanderson electronegativities	-0.594(2)
13	MATS2p	2D autocorrelations	Moran autocorrelation of lag-2/ weighted by atomic polarizabilities	0.694(5)
14	GATS2v	2D autocorrelations	Geary autocorrelation of lag-2/ weighted by atomic van der Waals volumes	-0.794(1)
15	nCp	Functional	Number of total primary C (sp3)	0.274(1), -0.612(1)
16	nCrHR	Functional	Number of ring tertiary C(sp3)	-0.771(1)
17	nNR2	Functional	Number of tertiary aliphatic amines	-0.702(1)
18	C-028	Atom-centered fragments	R--CR--X	-0.464(6)
19	C-032	Atom-centered fragments	X--CX--X	-0.458(2)

[a]The descriptors are identified from the five parameter models, emerged from CP-MLR protocol with filter-1 as 0.79, filter-2 as 2.0, filter-3 as 0.813, and filter-4 as $0.3 \leq q2 \leq 1.0$ with a training set of 27 compounds. [b]The average regression coefficient of the descriptor corresponding to all models and the total number of its incidence. The arithmetic sign of the coefficient represents the actual sign of the regression coefficient in the models

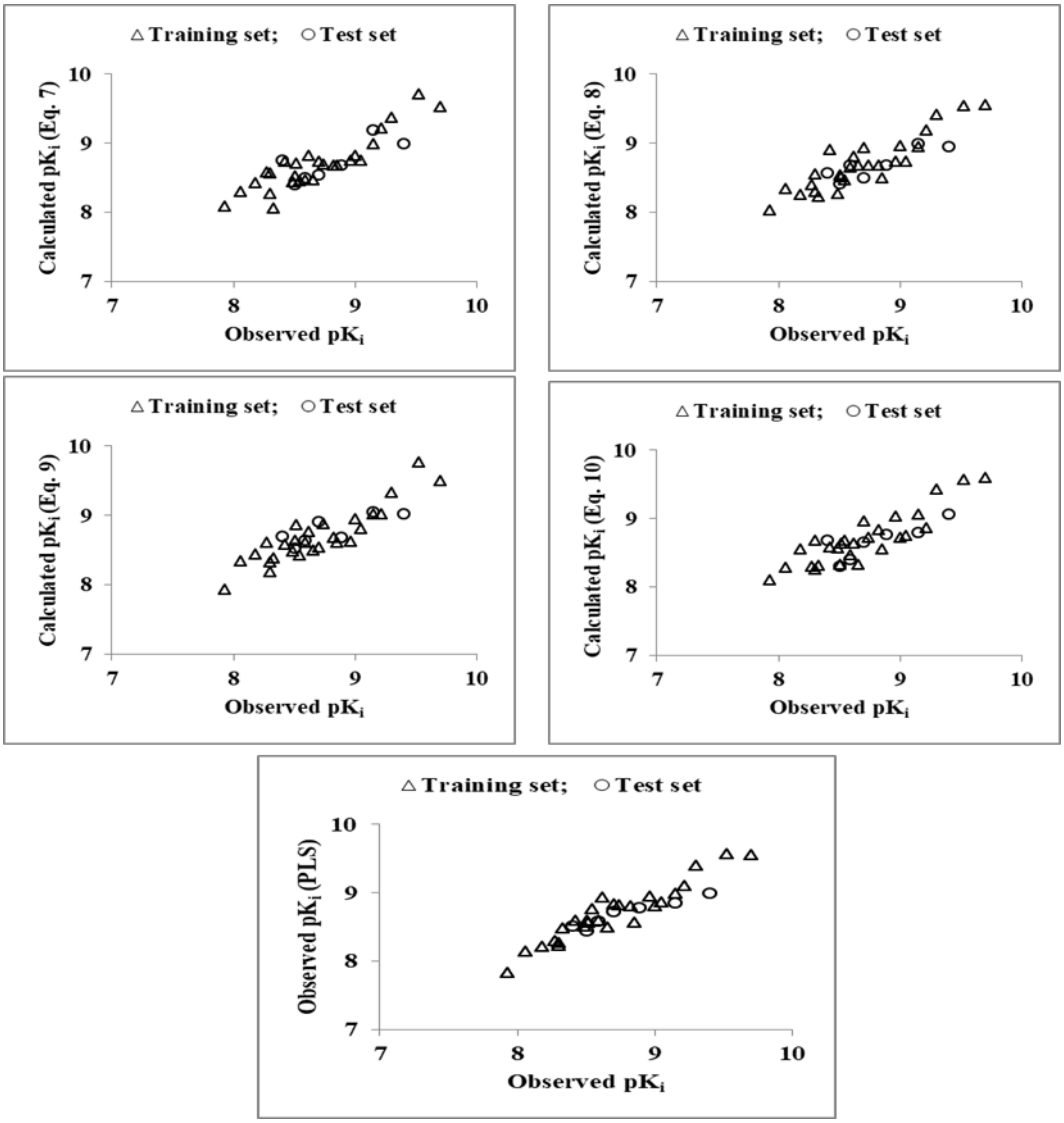


Figure 2 Plot of observed versus calculated pK$_i$ values for training- and test-set compounds.

The participated descriptors BELe5 (lowest eigenvalue n.5 of Burden matrix/weighted by atomic Sanderson electronegativities), BELv3 (lowest eigenvalue n.3 of Burden matrix/weighted by atomic van der Waals volumes) and BEHv1 (highest eigenvalue n.1 of Burden matrix/weighted by atomic van der Waals volumes) belong to BCUT class. All these descriptors contributed positively to the activity suggesting that higher value of these will augment the activity.

From Equations (7)-(10), it appeared that the descriptors nCrHR, a functional group accounting descriptor representing number of ring tertiary C(sp3) functionality in a structure and atom centered fragment accounting descriptor C-028 showing R—CR--X type fragment in a molecular structure make negative contribution to activity and JGI2, mean Galvez topological charge index of order 2 shown positive correlation to the activity. In this way absence of number of ring tertiary C(sp3) functionality along with R—CR--X type fragment in a molecular structure and higher value of mean Galvez topological charge index of order 2 would be advantageous in improving the DPP4 inhibition activity of a compound.

To corroborate the study further, a PLS analysis has also been carried out on 19descriptors identified through CP-MLR and results are given in Table 3. For this purpose, the descriptors have been autoscaled (zero mean and unit s.d.) to give each one of them equal weight in the analysis. In the PLS cross-validation, three components have been found to be the optimum for these 19 descriptors and they explained 89.7% variance in the activity (r^2 = 0.897).The MLR-like PLS coefficients of these 19 descriptors are given in Table 3.

Table 3 PLS and MLR-like PLS models from the descriptors of five parameter CP-MLR models for DPP-4 binding affinity.

A: PLS equation			
PLS components	PLS coefficient (s.e.)[a]		
Component-1	0.196(0.015)		
Component-2	-0.113(0.019)		
Component-3	0.078(0.023)		
Constant	8.693		
B: MLR-like PLS equation			
S. No.	Descriptor	MLR-like coefficient (f.c.)[b]	Order
1	Ms	0.311(0.109)	2
2	nDB	0.045(0.016)	17
3	HNar	-0.157(-0.55)	10
4	PJI2	0.184(0.064)	9
5	IC1	0.115(0.040)	13
6	BELm8	-0.083(-0.029)	14
7	BEHv1	0.318(0.111)	1
8	BELv3	0.215(0.075)	3
9	BELe5	0.210(0.073)	4
10	JGI2	0.130(0.045)	12
11	MATS5e	-0.077(-0.027)	15
12	MATS6e	-0.131(-0.045)	11
13	MATS6p	-0.020(-0.007)	18
14	GATS2v	-0.055(-0.019)	16
15	nCp	0.011(0.004)	19
16	nCrHR	-0.208(-0.072)	5
17	nNR2	-0.184(-0.064)	8
18	C-028	-0.187(-0.065)	7
19	C-032	-0.205(-0.071)	6
Constant		7.909	
C: PLS regression statistics		Values	
n		27	
r		0.947	
s		0.148	
F		67.415	
q^2_{LOO}		0.851	
q^2_{L50}		0.860	
r^2_{Test}		0.662	

[a]Regression coefficient of PLS factor and its standard error. [b]Coefficients of MLR-like PLS equation in terms of descriptors for their original values; f.c. is fraction contribution of regression coefficient, computed from the normalized regression coefficients obtained from the autoscaled (zero mean and unit s.d.) data.

The calculated activity values of training- and test-set compounds are in close agreement to that of the observed ones and are listed in Table 1. For the sake of comparison, the plot between observed and calculated activities (through PLS analysis) for the training- and test-set compounds is given in Figure 2. Figure 3 shows a plot of the fraction contribution of normalized regression coefficients of these descriptors to the activity (Table 3).

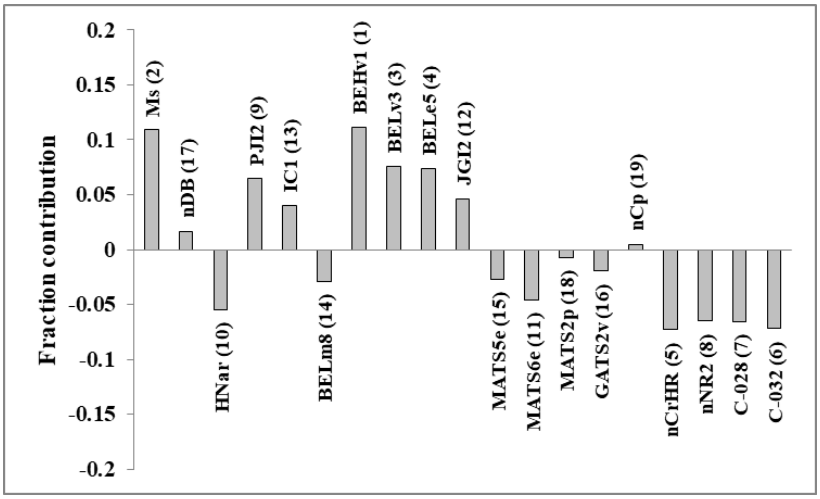


Figure 3 Plot of fraction contribution of MLR-like PLS coefficients (normalized) against 19 identified descriptors (Table 3) associated with DPP-4 binding affinity of the compounds.

The PLS analysis in 19 identified descriptors revealed three components (Table 3) as optimum to explain the DPP4 inhibition activity. The top ten descriptors in decreasing order of significance are BEHv1, Ms, BELv3, BELe5, nCrHR, C-032, C-028, nNR2, PJI2 and HNar (Table 3, figure 3). Among these descriptors, BEHv1, Ms, BELv3, BELe5, nCrHR, C-028 and PJI2 are part of Equations discussed above and convey same inferences in PLS analysis. The negative contributions of atom centered fragment descriptor C-032 (X--CX--X type fragment), functional group count descriptor nNR2 (number of tertiary aliphatic amine functionality in a molecule) and topological descriptor HNar (Narumi harmonic topological index) advocated lower value of these are helpful in improving the activity profile. It is also observed that PLS model from the dataset devoid of 19 descriptors (Table 3) remained inferior in explaining the activity of the analogues.

3.2. Applicability domain

On analyzing the applicability domain (AD) in the Williams plot (Figure 3) of the model based on the whole dataset (Table 4), no any compound has been identified as an obvious 'outlier' for the DPP4 inhibitory activity if the limit of normal values for the Y outliers (response outliers) was set as 3×(standard deviation) units. One of the compound (**2**; Table 1) was found to have leverage (h) values greater than the threshold leverage (h*); suggesting it as chemically influential compound.

Table 4 Models derived for the whole data set (n = 34) for the DPP-4 binding affinity in descriptors identified through CP-MLR.

Model	r	s	F	q2LOO	Eq.
pKi = 7.366 + 1.980(0.276)Ms – 0.762(0.176)nDB + 0.231(0.092)PJI2 + 1.073(0.178)BELv3 + 0.731(0.181)JGI2	0.880	0.216	19.331	0.653	(7a)
pKi = 7.653 + 1.858(0.265)Ms – 0.844(0.176)nDB + 1.081(0.175)BELv3 + 0.694(0.180)JGI2 – 0.426(0.156)C-028	0.884	0.212	20.176	0.680	(8a)
pKi = 6.571 + 1.600(0.184)Ms + 0.356(0.092)PJI2 + 0.764(0.241)BEHv1 + 0.777(0.180)BELv3 + 0.698(0.171)MATS2p	0.884	0.213	20.026	0.156	(9a)
pKi = 8.402 + 1.269(0.194)Ms + 0.295(0.098)PJI2 + 0.685(0.159)BELe5 – 0.834(0.187)MATS6e – 0.792(0.185)nCrHR	0.865	0.228	16.727	0.623	(10a)

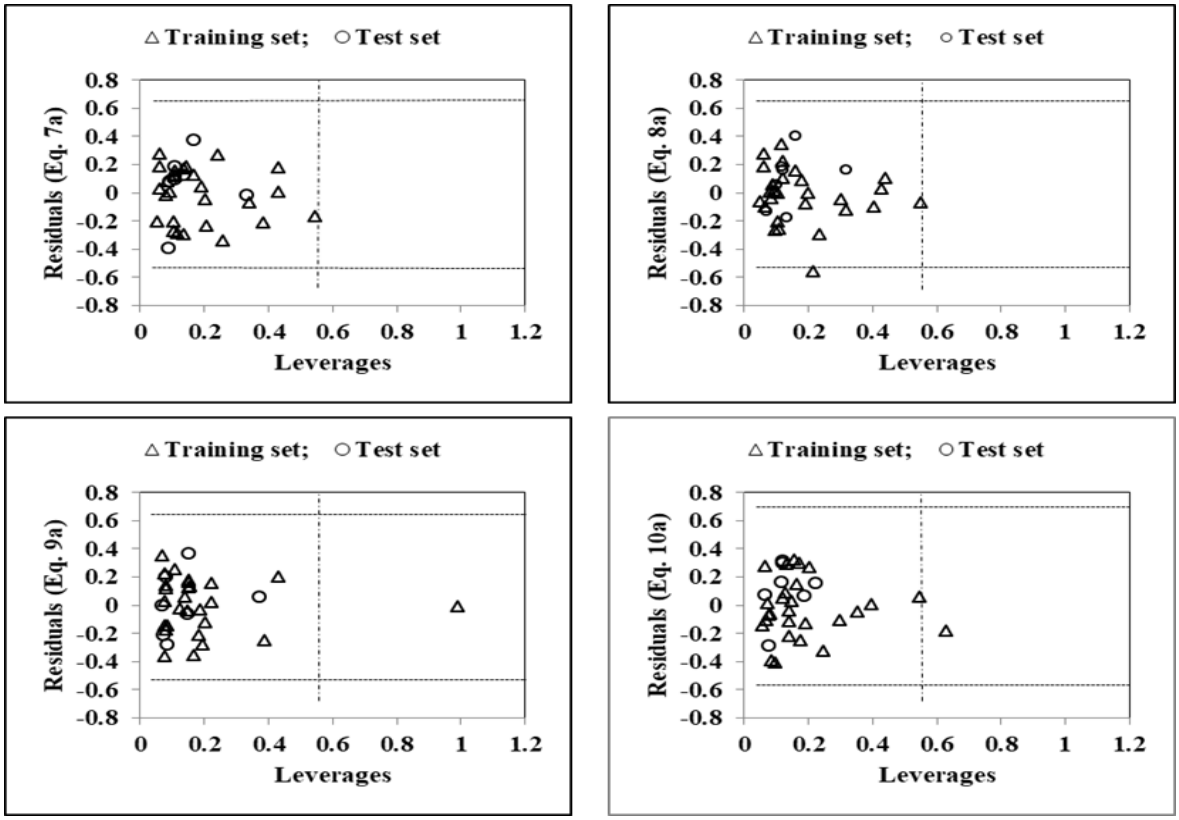


Figure 4 Williams plot for the training-set and test- set for inhibition activity of DPP4 for the compounds in Table 1. The horizontal dotted line refers to the residual limit (±3×standard deviation) and the vertical dotted line represents threshold leverage h* (= 0.529).

For both the training-set and test-set, the suggested model matches the high quality parameters with good fitting power and the capability of assessing external data. Furthermore, all of the compounds were within the applicability domain of the proposed model and were evaluated correctly.

4. Conclusion

The DPP4 inhibition activity of imidazolopyrimidine amides has been quantitatively analyzed in terms of chemometric descriptors. The statistically validated quantitative structure-activity relationship (QSAR) models provided rationales to explain the inhibition activity of these congeners. The descriptors identified through combinatorial protocol in multiple linear regression (CP-MLR) analysis have highlighted the role of mean electrotopological state (Ms), number of double bonds in molecular structure (nDB), 2D Petitijean shape index (PJI2), Moran autocorrelation of lag-2/ weighted by atomic polarizabilities (MATS2p), Moran autocorrelation of lag-6/weighted by atomic Sanderson electronegativities (MATS6e), lowest eigenvalue n.5 of Burden matrix/ weighted by atomic Sanderson electronegativities (BELe5), lowest eigenvalue n.3 of Burden matrix/ weighted by atomic van der Waals volumes (BELv3), highest eigenvalue n.1 of Burden matrix/ weighted by atomic van der Waals volumes (BEHv1). In addition to these 2nd order mean Galvez topological charge index (JGI2), number of ring tertiary C(sp3) (nCrHR) and R--CR--X type structural fragments (C-028) have also shown prevalence to model the inhibitory activity.

From statistically validated models, it appeared that the descriptors Ms, PJI2, JGI2, MATS2p, BELe5, BELv3 and BEHv1 make positive contribution to activity and their higher values are conducive in improving the DPP4 inhibition activity of a compound. On the other hand, the descriptors nDB, C-028, nCrHR and MATS6e render detrimental effect to activity. Therefore, absence or lower number of double bonds (nDB), R--CR--X type structural fragment (C-028), number of ring tertiary C(sp3) (nCrHR) and lower value of descriptor MATS6e would be advantageous. Such guidelines may be helpful in exploring more potential analogues of the series. The statistics emerged from the test sets have validated the identified significant models. PLS analysis has further confirmed the dominance of the CP-MLR identified descriptors. Applicability domain analysis revealed that the suggested models have acceptable predictability. All the compounds are within the applicability domain of the proposed models and were evaluated correctly.

Compliance with ethical standards

Acknowledgments

Authors are thankful to their institution for providing necessary facilities to complete this study.

Disclosure of conflict of interest

The authors declare no conflict of interest.

References

[1] Wideman RD and Kieffer TJ. (2009). Mining incretin hormone pathways for novel therapies. Trends in Endocrinology and Metabolism, 20, 280-286.

[2] Triplitt C, Wright A and Chiquette E. (2006). Incretin mimetics and dipeptidyl peptidase-IV inhibitors: potential new therapies for type 2 diabetes mellitus. Pharmacotherapy, 26, 360-374.

[3] Gallwitz B. (2005). Glucagon-like peptide-1-based therapies for the treatment of type 2 diabetes mellitus. Treatments in Endocrinology, 4, 361-370.

[4] Holst JJ. (1994). Glucagon-like peptide-1 (GLP-1) a newly discovered GI hormone. Gastroenterology, 107, 1048-1055.

[5] Drucker DJ. (1998). Glucagon-like peptides. Diabetes, 47, 159-169.

[6] Vilsboll T and Holst JJ. (2004). Incretins, insulin secretion and type 2 diabetes mellitus. Diabetologia, 47, 357-366.

[7] Elahi D, McAloon-Dyke M, Fukagawa NK, Meneilly GS, Sclater AL, Minaker KL, Habener JF and Andersen DK. (1994). The insulinotropic actions of glucose-dependent insulinotropic polypeptide (GIP) and glucagon-like peptide-1 (7-37) in normal and diabetic subjects. Regulatory Peptides, 51, 63-75.

[8] Nauck MA, Bartels E, Oerskov C, Ebert R and Creutzfedlt WJ. (1993). Additive insulinotropic effects of exogenous synthetic human gastric inhibitory polypeptide and glucagon-like peptide-1-(7-36) amide infused at near-physiological insulinotropic hormone and glucose concentrations. Clinical Endocrinology and Metabolism, 76, 912-917.

[9] Gentilella R, Bianchi C, Rossi A and Rotella CM. (2009). Exenatide: a review from pharmacology to clinical practice. Diabetes, Obesity and Metabolism, 11,544-556.

[10] Gonzalez C, Beruto V, Keller G, Santoro S and Di Girolamo G. (2006). Investigational treatments for type 2 diabetes mellitus: Exenatide and Liraglutide. Expert Opinion on Investigational Drugs, 15, 887-895.

[11] Holst JJ. (2002). Therapy of type 2 diabetes mellitus based on the actions of glucagon-like peptide-1. Diabetes/ Metabolism Research and Reviews, 18, 430-441.

[12] Hansen L, Deacon CF, Orskov C and Holst JJ. (1999). Glucagon-likepeptide-1-(7-36) amide is transformed to glucagon-like peptide-1-(9-36)amide by dipeptidyl peptidase IV in the capillaries supplying the L cells of the porcine intestine. Endocrinology, 140, 5356-5363.

[13] Nagakura T, Yasuda N, Yamazaki K, Ikuta H, Yoshikawa S, Asano O and Tanaka I. (2001). Improved glucose tolerance via enhanced glucose-dependent insulin secretion in dipeptidyl peptidase IV-deficient Fischer rats. Biochemical and Biophysical Research Communications, 284, 501-506.

[14] Marguet D, Baggio L, Kobayashi T, Bernard A, Pierres M, Nielsen PF, Ribel U, Watanabe T, Drucker DJ and Wagtmann N. (2000). Enhanced insulin secretion and improved glucose tolerance in mice lacking CD26. Proceedings of the National Academy of Sciences of the United States of America, 97, 6874-6879.

[15] Durinx C, Lambeir A, Bosmans E, Falmagne J, Berghmans R, Haemers A, Scharpe S and De Meester I. (2000). Molecular characterization of dipeptidyl peptidase activity in serum. Soluble CD26/dipeptidyl peptidase IV is responsible for the release of X-Pro dipeptides. European Journal of Biochemistry, 267, 5608-5613.

[16] Iwaki-Egawa S, Watanabe Y, Kikuya Y and Fujimoto Y. (1998). Dipeptidyl peptidase IV from human serum: purification, characterization, and N-terminal amino acid sequence. Journal of Biochemistry, 124, 428-433.

[17] Villhauer EB, Brinkman JA, Naderi GB, Burkey BF, Dunning BE, Prasad K, Mangold BL, Russell ME and Hughes TE. (2003). 1-[[(3-Hydroxy-1-adamantyl)amino]acetyl]-2-cyano-(S)-pyrrolidine: A potent, selective, and orally bio available dipeptidyl peptidase IV inhibitor with antihyperglycemic properties. Journal of Medicinal Chemistry, 46, 2774-2789.

[18] Kim D, Wang L, Beconi M, Eiermann GJ, Fisher MH, He H, Hickey GJ, Kowalchick JE, Leiting B, Lyons K, Marsilio F, McCann ME, Patel RA, Petrov A, Scapin G, Patel SB, Roy RS, Wu JK, Wyvratt MJ, Zhang BB, Zhu L, Thornberry NA and Weber AE. (2005). (2R)-4-Oxo-4-[3-(trifluoromethyl)-5,6-dihydro[1,2,4]triazolo[4,3-α]pyrazin-7(8H)-yl]-1-(2,4,5-trifluorophenyl)butan-2-amine: a potent, orally active dipeptidyl peptidase IV inhibitor for the treatment of type 2 diabetes. Journal of Medicinal Chemistry, 48, 141-151.

[19] Augeri DJ, Robl JA, Betebenner DA, Magnin DR, Khanna A, Robertson JG, Wang A, Simpkins LM, Taunk P, Huang Q, Han SP, Abboa-Offei B, Cap M, Xin L, Tao L, TozzoE, Welzel GE, Egan DM, Marcinkeviciene J, Chang SY, Biller SA, Kirby MS, Parker RA and Hamann LG. (2005). Discovery and preclinical profile of saxagliptin (BMS-477118): a highly potent, long-acting, orally active dipeptidyl peptidase IV inhibitor for the treatment of Type 2 diabetes. Journal of Medicinal Chemistry, 48, 5025-5037.

[20] Feng J, Zhang Z, Wallace MB, Stafford JA, Kaldor SW, Kassel DB, Navre M, Shi L, Skene RJ, Asakawa T, Takeuchi K, Xu R, Webb DR and Gwaltney SL. (2007). Discovery of alogliptin: a potent, selective, bio available, and efficacious inhibitor of dipeptidyl peptidase IV. Journal of Medicinal Chemistry, 50, 2297-2300.

[21] Havale SH and Pal M. (2009). Medicinal chemistry approaches to the inhibition of dipeptidyl peptidase-4 for the treatment of type 2 diabetes. Bioorganic and Medicinal Chemistry, 17, 1783-1802.

[22] Brigance RP, Meng W, Fura A, Harrity T, Wang A, Zahler R, Kirby MS and Hamman LG. (2010). Synthesis and SAR of azolopyrimidines as potent and selective dipeptidyl peptidase-4 (DPP4) inhibitors for type 2 diabetes. Bioorganic and Medicinal Chemistry Letters, 20, 4395-4398.

[23] Meng W, Brigance RP, Chao HJ, Fura A, Harrity T, Marcinkkeviciene J, O'Connor SP, Tamura JK, Xie D, Zhang Y, Klei HE, Kish K, Weigett CA, Turdi H, Wang A, Zahler R, Kirby MS and Hamann LG. (2010). Discovery of 6-(Aminomethyl)-5-(2,4-dichlorophenyl)-7-methylimidazo[1,2-α]pyrimidine-2-carboxamides as potent, selective dipeptidyl peptidase-4 (DPP4) inhibitors. Journal of Medicinal Chemistry,53, 5620-5628.

[24] SYSTAT, Version 7.0; SPSS Inc 444 North Michigan Avenue, Chicago IL, 60611.

[25] ChemDraw Ultra 6.0 and Chem3D Ultra, Cambridge Soft Corporation, Cambridge, USA.

[26] Todeschini R and Consonni V. Dragon software (version 1.11-2001), Milano, Italy.

[27] Prabhakar YS. (2003). A combinatorial approach to the variable selection in multiple linear regression: analysis of Selwood et al. Data Set-a case study. QSAR and Combinatorial Science, 22, 583-595.

[28] Wold S. (1978). Cross-validatory estimation of the number of components in factor and principal components models. Technometrics, 20, 397-405.

[29] Kettaneh N, Berglund A and Wold S. (2005). PCA and PLS with very large data sets. Computational Statistics and Data Analysis, 48, 69-85.

[30] Stahle L and Wold S. (1998). Multivariate data analysis and experimental design, in Biomedical Research, Progress in Medicinal Chemistry, Ellis GP and West WB, eds., Elsevier Science Publishers, BV, The Netherlands, 25, 291–338.

[31] Sharma S, Prabhakar YS, Singh P and Sharma BK. (2008). QSAR study about ATP-sensitive potassium channel activation of cromakalim analogues using CP-MLR approach. European Journal of Medicinal Chemistry, 43, 2354-2360.

[32] Sharma S, Sharma BK, Sharma SK, Singh P and Prabhakar YS. (2009). Topological descriptors in modeling the agonistic activity of human A_3 adenosine receptor ligands: The derivatives of 2-Chloro-N^6-substituted-4'-thioadenosine-5'-uronamide. European Journal of Medicinal Chemistry, 44, 1377-1382.

[33] Sharma BK, Pilania P, Singh P and Prabhakar YS. (2010). Combinatorial protocol in multiple linear regression/partial least-squares directed rationale for the caspase-3 inhibition activity of isoquinoline-1,3,4-trione derivatives. SAR and QSAR in Environmental Research, 21, 169-185.

[34] Sharma BK, Singh P, Sarbhai K and Prabhakar YS. (2010). A quantitative structure-activity relationship study on serotonin 5-HT_6 receptor ligands: Indolyl and piperidinyl sulphonamides. SAR and QSAR in Environmental Research, 21, 369-388.

[35] Sharma BK, Pilania P, Sarbhai K, Singh P and Prabhakar YS. (2010). Chemometric descriptors in modeling the carbonic anhydrase inhibition activity of sulfonamide and sulfamate derivatives. Molecular Diversity, 14, 371-384.

[36] Sharma BK, Singh P, Shekhawat M, Sarbhai K and Prabhakar YS. (2011). Modelling of serotonin reuptake inhibitory and histamine H_3 antagonistic activity of piperazine and diazepane amides: QSAR rationales for co-optimization of the activity profiles. SAR and QSAR in Environmental Research, 22, 365-383.

[37] So S-S and Karplus M. (1997). Three-dimensional quantitative structure activity relationships from molecular similarity matrices and genetic neural networks. 1. Method and validations. Journal of Medicinal Chemistry, 40, 4347-4359.

[38] Prabhakar YS, Solomon VR, Rawal RK, Gupta MK and Katti SB. (2004). CP-MLR/PLS directed structure–activity modeling of the HIV-1RT inhibitory activity of 2,3-diaryl-1,3-thiazolidin-4-ones. QSAR and Combinatorial Science, 23, 234-244.

[39] Gramatica P. (2007). Principles of QSAR models validation: internal and external. QSAR and Combinatorial Science, 26, 694-701.

[40] Eriksson L, Jaworska J, Worth AP, Cronin MTD, McDowell RM and Gramatica P. (2003). Methods for reliability and uncertainty assessment and for applicability evaluations of classification and regression-based QSARs. Environmental and Health Perspectives, 111, 1361-1375.

[41] Golbraikh A and Tropsha A. (2002). Beware of q^2! Journal of Molecular Graphics and Modeling, 20, 269-276.

How to cite this article

Parihar R and Sharma BK. (2020). QSAR rationales for the dipeptidyl peptidase-4 (DPP-4) inhibitors: The imidazolopyrimidine amides. GSC Biological and Pharmaceutical Sciences, 11(3), 130-144.