بسم الله الرحمن الرحيم

SUDAN UNVERSITY OF SCIENCE AND TECHNOLOGY

COLLEGE OF GRADUATE STUDIES

Quantitative Structure - Activity Relationship Study of some N-substituted -2-Isonicotinoyl Hydrazine Carboxamide Derivatives as Anti Tuberculosis Agents

دراسة العلاقة بين البنية و الفعالية لبعض مشتقات أيزونيكوتينويل هيدرازين كربوكساميد كعوامل مضادة للسل

BY:

Shima Ibrahim Mohammed Ali

BSc.(Honours)Chemistry

A thesis submitted in partial fulfillment of the requirements
of Master degree in
Chemistry at Sudan University of Science and Technology.

Supervisor:

Prof.Dr. Ahmed Elsadig Mohammed Saeed

Chemistry Department

College of Science

November 2020

بسم الله الرحمن الرحيم

قال تعالي:

﴿ **وَيَسْأَلُونَكَ عَنِ الرُّوحِ ۖ قُلِ الرُّوحُ مِنْ أَمْرِ رَبِّي وَمَا أُوتِيتُم مِّنَ الْعِلْمِ إِلَّا قَلِيلًا** ﴾

صدق الله العظيم

سورة الإسراء (85)

Dedication

I dedicate this work with love and grateful to

my parents,
brother,
sisters
and
friends

Acknowledgement

I would like to say a special thanks to my supervisor, Prof.Dr. Ahmed Elsadig Mohammed Saeed , for guidance and overall insights in this field have made this an inspiring experience to me.

Thanks also to everyone who helped me so much. Furthermore, I would like to thank my family for supporting me during working on this thesis .Finally, special thanks to Rawan, Hanaa, Heba and Esraa for providing advice and supporting.

Abstract

Tuberculosis (TB) is a chronic disease caused by *Mycobacterium tuberculosis.* The appearance of multidrug-resistant strains of *Mycobacterium tuberculosis* (*M.tb*) has led to an urgent search for new and effective anti-TB drugs. Isoniazid remains the main and most effective component of all the multiple therapeutic regimens recommended by the World Health Organization.

In this work number of computer programs were used, such as ACD/Labs , MOE and SPSS. From previously published article , (21 compounds) of N-substituted -2-isonicotinoyl hydrazine carboxamide derivatives were divided into two groups, training set (15 compounds) and a test set(6 compounds) ,one of them was excluded). The biological activity (MIC) of data set was converted to logarithmic scale (pMIC) .The biological activity was correlated with selected descriptors were chosen according to correlation matrix to create a (QSAR) models. The models were used in study the relationship between structure and effectiveness by using the (MOE) program by (PLS) method. A number of equations were obtained in 2D ,3D and (2D+3D) dimensionals. The validity of the equations was confirmed by using the internal and external validation in addition to multiple linear regression (MLR) statistical analysis method to increase the strength and robustness of the equations. The equations were used to predicted the biological activity of data set compounds.

The best equations in the (2D) dimensional with value of correlation coefficient R^2 = 0.87. Also based on internal validation by using the training set (LOO, leave-one-out) method was obtained value Q^2 =0.73 and the external investigation was conducted using the test set the value was R^2_{pred}= 0.70 and the equation is

pMIC=8.00423-0.24582chiov-0.00617slog_VSA9-0.35165kier2 ,

And also the equation with a correlation coefficient R^2=0.84, Q^2=0.70, R^2_{pred}=0.71

pMIC=6.07370+0.10173kierA3-0.70720chi1v-0.21145logp(o/w)

The best equation in the (3D) dimensional has the highest correlation coefficient R^2 =0.96 values of Q^2= 0.9 and R^2_{pred}= 0.70 according to the equation

pMIC=10.25002-0.01536ASA-1.15755E_oop-2.58249npr1.

In addition to in (2D+3D) dimensional

pMIC=7.5644-1.1662E_oop-0.0103ASA-0.2475logp(o/w)

with correlation coefficient R^2=0.93 ,Q^2=0.89 and R^2_{pred}=0.73

The Fifty eight new compounds were designed , the four equations were applied on the new designed compounds (58) of N-substituted -2-isonicotinoyl hydrazine carboxamide to predict the biological activity values (pMIC) for them and compared to biological activity of Isoniazid (INH) pMIC = -0.3 . The compounds (XXIX, XXX, XXXII,XLI, XLIX) were showed biological activity values close to the biological activity value of Isoniazid (INH).

The molecular docking was applied on the compounds of the data set and the 58 new designed compounds with (4TRO) protein. 4TRO protein was obtained from the Protein Data Bank (PDB) ,the designed compounds that were showed more interaction with 4TRO (VII, VIII XLVII, XXXVII, XL, LI,LII).

الخلاصة

السل (TB) هو مرض مزمن ينتج عن .Mycobacterium tuberculosis أدى ظهور سلالات مقاومة للأدوية المتعددة من المتفطرة السلية (Mtb) إلى البحث العاجل عن أدوية جديدة وفعالة لمكافحة السل. .يظل أيزونيازيد هو المكون الرئيسي والأكثر فعالية لجميع الأنظمة العلاجية المتعددة التي أوصت بها منظمة الصحة العالمية.

في هذه الدراسة تم استخدام عدد من برامج الكمبيوتر مثل (ACD/Labs ,SPSS,MOE من ورقه منشوره مسبقا ، مجموعة بياناتها (21 مركب) -2- N-substituted isonicotinoyl hydrazine carboxamide قسم إلي مجموعتين ، مجموعة تدريب (15 مركب) ومجموعة اختبار (6 مركبات ، تم استبعاد أحدها). تم تحويل النشاط البيولوجي (MIC) لمجموعة البيانات إلى مقياس لوغاريثمي (pMIC) ، وتم ربط النشاط البيولوجي بمعاملات محددة وفقًا لمصفوفة الارتباط لإنشاء نماذج (QSAR) .استخدمت النماذج في دراسة العلاقة بين البنية والفعالية باستخدام برنامج (MOE) بطريقة (PLS) . تم الحصول على عدد من المعادلات بأبعاد ثنائية وثلاثية الأبعاد و(ثنائية + ثلاثية) الأبعاد. تم التأكد من صحة المعادلات باستخدام طريقة التحقيق الداخلي والداخلي و طريقه التحليل الإحصائي للانحدار الخطي المتعدد (MLR) لزيادة قوة ومتانة المعادلات. تم استخدام المعادلات للتنبؤ بالنشاط البيولوجي لمركبات مجموعة البيانات.
أفضل المعادلات في البعد (2D dimensionl) ذات قيمه معامل ارتباط (R^2)=0.87 وبنائاً علي قيم التحقق الداخلي باستخدام مجموعة التدريب بطريقه (LOO, leave-one-out) وجدت قيمة Q^2=0.73 وأجرى التحقيق الخارجي باستخدام مجموعة الإختبار وحسبت قيمه
R^2_{pred} = 0.70 والمعادلة هي

pMIC=8.00423-0.24582chiov-0.00617slog_VSA9-0.35165kier2 ,

و أيضا المعادلة ذات معامل الارتباط R^2=0.84 ، Q^2 =0.70 R^2_{pred}= 0.71

PMIC=6.07370+0.10173kierA3-0.70720chi1v-0.21145logp(o/w)

وأفضل معادله في البعد (3D dimensionl) ذات اعلي معامل ارتباط (R^2)= 0.96 وقيم
Q^2 =0.91 و R^2_{pred}=0.70 وفقا للمعادلة

pMIC=10.25002-0.01536ASA-1.15755E_oop-2.58249npr1

بالإضافة إلى المعادلة في الأبعاد(2D+3D)

pMIC=7.5644-1.1662E_oop-0.0103ASA-0.2475logp(o/w)

مع معامل الارتباط R^2 = 0.93 ، Q^2 =0.73 ، R^2_{pred} = 0.89

تم تطبيق الأربع معادلات على المركبات المصممة الجديدة (58) من -N المستبدلة -2-isonicotinoyl hydrazine carboxamide للتنبؤ بقيم النشاط البيولوجي (pMIC) لها ومقارنتها بالنشاط البيولوجي لـ أيزونيازيد (INH) .pMIC = -0.3. أظهرت المركبات(XXX XXXII، XLI، XXLIX) قيم نشاط قريبة من قيمة النشاط البيولوجي للإيزونيازيد (INH) تم تطبيق الالتحام الجزيئي على مركبات مجموعة البيانات والمركبات المصممة الجديدة ببروتين4 تم الحصول على بروتين 4TRO من بنك بيانات البروتين (PDB) ، وهذه هي المركبات المصممة التي أظهرت تفاعلًا أكبر مع 4TRO (VII, VIII, XLVII, XXXVII, XL, LI,LII).

Content

LIST OF TABLES

LIST OF FIGURE

ABBREVIATIONS

Abbreviation	ABBREVIATION MEANING
ADMET	Absorption, distribution, metabolism, excretion and toxicity
ANN	Artificial neural network
ASA	Negative accessible surface area
Asp	Aspartic acid
Å	Bond length
BCUT-SMR-1	Molarre fractivity Bcut(1/3)
balabanJ	balaban averaged distance sum connectivity
CADD	Computer aided drug design
chi1v	Atomic valence connectivity index (order1)
chiov	Atomic valence connectivity index (order0)
dens	Mass density
E-oop	Out-of -plane Energy
E-tor	Torsion energy
F-value	Fischers value
GA-MLR	Genetic algorithms with multiple linear regression
Gln	Glutamine
Gly	Glycine
Kier2	Second Kaapa shape index
KierA3	Third Kaapa shape index
lle	Isoleucine
IR	Infra- red
IUPAC	International union of pure and applied chemistry
LMO	Leave more out
Log p(o/w)	Log octanol /water.
LOO	Leave one out
Lys	Lysine
Me	Methyl
MD	Molecular dynamics
MIC	Minimal inhibitory concentration
ML	Machine learning
MLR	Multiple linear regression
MOE	Molecular operating environment
NMR	Nuclear magnetic resonance
npr1	Normalized PMI ratio (pmi1/pmi3)
pMIC	log of the MIC
RMSE	Root mean square error

Phe	Phenylalanine
QSAR	Quantitative structure activity relationship
QSAAR	Quantitative structure activity activity relationship
QSPR	Quantitative structure property relationship
QSTR	Quantitative structure toxicity relationship
Q^2	Cross validation regression coefficient
R	correlation coefficient
RMSE	Root mean square Error
R^2	square of correlation coefficient
S	Free binding energy
SAR	Structure activity relationship
SEE	Standard error of estimation
Ser	Serine
SPSS	Statistical package for social science
Std-dim3	Standard dimension 3
Slogp-VSA9	Bin 9 Slog(0,40,10)
TB	Tuberculosis
U.V	Ultra Violet
Val	Valine
Vol	Van der waals volume

Chapter One

Introduction

1.1.Computational Chemistry:

Computational Chemistry is a science with applications in most areas of chemistry, biochemistry and material sciences.

It employed tools that have been developed during years to model a wide variety of chemical processes, ranging from very accurate studies of small molecules in the gas phase to complex simulations of macromolecular systems, crystals and solutions. Different theoretical methods are used in different types of applications. Small molecules are treated using sophisticated and very accurate models of the Schrodinger equation. Larger molecules are treated with more approximate methods, like density functional theory. Effects of a solvent or a crystal environment can be modeled using reaction-field Hamiltonians and model potentials. Macromolecular systems are treated at an even simpler level, where only a ‘‘reactive center’’ is studied by Quantum Chemistry while the surroundings are modeled by classical mechanics Monte-Carlo (MC) or molecular dynamics (MD) .(Jensen,2007)

Computational Chemistry also called molecular modeling , the two terms mean about the same thing is a set of techniques for investigating chemical problems on a computer. (Lewars, 2011)

As one of the most active fields, the development of Computational Chemistry can not only analyze the experimental data but also predict ideal model complexes to inspire synthetic chemists. (Luo and Zheng, 2021)

With the exponential increase in data to be analysed, obtained through the introduction of automated whole genome and protein sequencing techniques, the field of bioinformatics rapidly emerged with the pioneering laborious mapping and comparison of protein and gene sequences in molecular biology, *via* an intense phase, which to a large extent can be viewed as ‘database mining’ and the development of efficient computer based algorithms, into a science of its own, reached a high level of maturity and sophistication. Tools in bioinformatics are used with great success in structural biology, Computational Chemistry, genetics, molecular biology, the pharmaceutical industry, pharmacology and more.(Genheden *et al*.,2017)

1.2.Molecular modeling:

Molecular modeling (MM) is one of the fastest growing fields in science. It may vary from building and visualizing simple molecules in three

dimensions (3D) to performing complex computer simulations on large proteins and nanostructures. MM is a collection of computer-based techniques for driving, representing and manipulating the structures and reactions of molecules, and the properties are dependent on these 3D structures. The techniques in MM cover several issues among them Computational Chemistry, drug design, computational biology, nanostructures, and material science.

It is found that the issue drug design is an interactive topic in MM and contribute to drug discovery both in academia and in industry. Computer-aided and structure-based drug design relies on knowledge of the 3D structure of the biological target. Drug design is an iterative process that begins when a compound is identified to display an interesting biological profile and ends when its activity profile and the chemical synthesis are optimized. MM permeates all aspects of drug design. Scientists have used computer models of new chemical entities to help define activity profiles, geometries, and relativities. Three stages of drug discovery can be achieved by MM: virtual screening, hit-to-lead optimization and lead optimization of other pharmaceutical properties while maintaining affinity. (Pimentel *et al*.,2013)

1.3 Molecular docking:

Pharmaceutical research has successfully incorporated a wealth of molecular modeling methods, within a variety of drug discovery programs, to study complex biological and chemical systems. The integration of computational and experimental strategies has been of great value in the identification and development of novel promising compounds. Broadly used in modern drug design, molecular docking methods explore the ligand conformations adopted within the binding or the sites of macromolecular targets. This approach also estimates the ligand-receptor binding free energy by evaluating critical phenomena involved in the intermolecular recognition process. Active site as a variety of docking algorithms are available, an understanding of the advantages and limitations of each method is of fundamental importance in the development of effective strategies and the generation of relevant results. (Ferreira *et al*.,2015)

One method for exploring the interactions between a ligand and a protein is to synthesize the ligand, co-crystallize it with the protein and then try to obtain an X-raystructure of the complex. Although both

synthesis and crystallography can sometimes be quite unpredictable and time-consuming, the method may be viable for small collections of ligands. If synthesis or crystallization fails, or if the aim is to screen many ligands for binding to the protein, computational molecular docking is often the method of first choice, and has become popular within both academia and industry. Furthermore, docking can be valuable when forming hypotheses regarding the way a ligand binds to the protein, or for modeling parts of the ligand whose structure or conformation when bound have not been successfully determined by crystallography. More than 60 docking programs are used, which 10 of them are roughly widely used . Docking requires a 3D structure of the protein as input. Typically, the software will generate 3D conformations of the ligands and optimize their interactions with the protein by computing the binding affinity scoring between the two. In most docking programs used, the ligand is treated as a flexible structure but the conformation of the protein is treated as being mostly rigid, and water molecules are typically not -considered at all. Obviously, both of these approximations constitute major simplifications of the real environment in which ligands and proteins interact. Still they are useful because of the immense amount of computation that would be necessary to accurately model the effects of water and protein flexibility – imagine the difficulty of modeling a lock and key that are constantly changing shape, in aqueous solution, and trying to measure the interactions between the two.

However, this simplifications are thought to be the two most important reasons why docking fails to correctly predict the affinity of a ligand for a protein, and the position of the ligand upon binding. (Andersson ,2010)

Molecular docking leads to discovery of therapeutic drugs through multiple ways that include:

I. Identification of potential target.

II. Screening of potent drugs as activators or inhibitors against certain diseases .

III. Designing of novel drugs by lead optimization

IV. Prediction of binding mode and nature of active site

V. Synthesis of chemical compounds with less time consumption.(Tripathi and Misra,2017)

1.4 Conformational of protein and ligand

The design of docking programs only considered ligand flexibility but was able to achieve certain degree of success in their results. Other studies took into account receptor flexibility and allowed adjustment of the protein conformation to accommodate more ligand motions at the binding site. However, modeling the receptor as flexible during docking increases the degree of freedom exponentially, which challenges the search ability of docking programs. Existing flexible-receptor protein–ligand docking programs can be grouped into three categories: (a) Soft-docking methods that modify the potential energy function to allow a closer approach between ligand and receptor such that small side-chain conformational changes is mimicked.

(b) Ensemble docking methods, that generate a set of receptor conformations as receptor candidates to simulate the protein conformational changes caused by binding the ligand into the pocket.

(c) Induced-fit docking methods , where both receptor and ligand conformations are dynamic during the docking process. (Wong *et al.,* 2021)

The most active area of theoretical research using molecular orbital theory has been in the prediction of the preferred conformation of molecules. Most molecules exist in multiple conformations. The preferred conformation of a molecule is a structural characteristic feature that arises as a response to the force of attraction and repulsion. The shape should be considered primarily in determining the interaction of the molecule with the receptor. The minimization energy is a function of bond angles, bond lengths, torsion angles and non-covalent interactions. By varying these parameters in a systematic way and calculating the total energy as a sum of orbital energies, can determine a minimum energy structure.(Nadendla,2004)

1.5 Quantitative Structure Activity Relationship (QSAR) :

Quantitative Structure Activity Relationship (QSAR) are mathematical models that seek to predict complicated physicochemical /biological properties of chemicals from their simpler experimental or calculated properties.

QSAR has emerged and has evolved trying to fulfill the medicinal chemist's need and desire to predict biological response. It found its way

into the practice of agro chemistry, pharmaceutical chemistry, and eventually most divisions of chemistry. (Muhammad *et al*.,2019)

A QSAR attempts to find consistent relationships between the variations in the values of molecular properties and the biological activity for a series of compounds so that these rules can be used to evaluate new chemical entities. A QSAR generally takes the form of a linear equation :

Biological activity = Constant $+ (C_1 . P_1) + (C_2 . P_2) + (C_3 . P_3)^{+} \ldots,$

where the parameters P_1 through P_n are computed for each molecule in the series and the coefficients C_1through C_n are calculated by fitting variations in the parameters and the biological activity. (Patel *et al*.,2014)

Computer-Aided Drug Design (CADD) involved widely employed computational approaches to discover and design new bioactive compounds. As examples of CADD techniques molecular docking, molecular dynamics (MD) simulations, pharmacophore modeling, similarity analysis, quantitative structure–activity relationship (QSAR) analysis, and machine learning (ML) techniques. (Serafim *et al*., 2021)

Advances in computing power have enabled development of software which allows simulation of the drug-receptor binding processes, a subset of computer-aided drug design (CADD) also referred to as virtual screening, with tremendous benefits to drug discovery efficiency (Kiriiri *et al*., 2020).

There are several Quantitative Structure-Activity Relationship (QSAR) methods to assist in the design of compounds for medicinal use. Owing to the different QSAR methodologies, deciding which QSAR method to use depends on the composition of system of interest and the desired results.(Esposito *et al*.,2004)

Based on chemometric methods, sometimes QSAR methods are also classified depending upon the type of correlation technique employed to establish a relationship between structural properties and biological activity. This includes linear methods including linear regression (LR), multiple linear regression (MLR), partial least squares(PLS), and principal component analysis/regression (PCA/PCR). Non-linear methods consist of artificial neural networks (ANN), k-nearest neighbors (kNN), and Bayesian neural nets .(Patel *et al*.,2014)

The QSAR models were developed by a combination of genetic algorithms with multiple linear regression (GA-MLR) methods to investigate the correlation between the activity and descriptors **.** (Setiawan *et al*.,2021)

The ability of accurate predictions of biological response (biological activity/property/toxicity) of a given chemical makes the quantitative structure-activity/property/toxicity relationship (QSAR/QSPR/QSTR) models unique among the *in silico* tools. In addition, experimental data of selected species can also be used as an independent variable along with other structural as well as physicochemical variables to predict the response for different species formulating quantitative activity–activity relationship (QAAR)/quantitative structure–activity–activity relationship (QSAAR) approach. Irrespective of the models type, the developed models quality, and reliability need to be checked through multiple classical stringent validation metrics. Among the validation metrics, error-based metrics are more significant as the basic idea of a good predictive model is to improve the predictions quality by lowering the predicted residuals for new query compounds.(Gajewicz-Skretna *et al*., 2021)

QSAR is the final result of computational processes that start with a suitable description of molecular structure and ends with some inference, hypothesis, and predictions on the behavior of molecules in environmental, physicochemical and biological system under analysis. The final outputs of QSAR computations are set of mathematical equations relating chemical structure to biological activity. Multivariate QSAR analysis employs all the molecular descriptors from various representations of a molecule (1D, 2D and 3D representation) dimensional to compute a model, in a search for the best descriptors valid for the property in analysis. The development of QSAR approaches drastically evolved and several multidimensional QSAR congeners(e.g., 4D-, 5D, and 6D-QSAR approaches) dimensional were introduced . (Minovski and Novič ,2017)

The success of any quantitative structure–activity relationship model depends on the accuracy of the input data, selection of appropriate descriptors. Validation is a crucial aspect of QSAR modeling. (Veerasamy *et*
al .,2011)

QSAR methodologies have the potential of decreasing substantially the time and effort required for the discovery of new medicines . The QSAR analysis employs statistical methods to derive quantitative mathematical relationship between chemical structure and biological activity. The process of QSAR modeling can be divided into three stages, development, model validation and application . (Muhammad *et al*.,2019)

1.5.1 Importance of validation of QSAR models:

Validation of QSAR is the most important parts in QSAR. The QSAR models can lead to false prediction of biological activity if the developed QSAR model is not validated. So validation of QSAR models, after model development, is most important part in QSAR studies. (Veerasamy *et al*., 2011)

1.5.1.1 Internal model validation:

The Leave-one-out (LOO), and Leave-More-out (LMO) cross-validation procedures are used for internal validation. (Setiawan *et al*., 2021)

A necessary condition for the validity of a regression model is that the multiple correlation coefficient R^2 is as close as possible to one and the standard error of the estimate s is small. However, this condition, which measures how well the model is able to mathematically reproduce the end point data of the training set, is an insufficient condition for model robustness and validity, as it do not express the ability of the model to make reliable predictions on new data. A necessary approach is to apply various cross-validations . Cross-validation refers to the use of one or more statistical techniques for internal validation in which different proportions of chemicals are omitted from the training set such as Leave-one-out (LOO), Leave-More-out (LMO) and bootstrapping and iteratively put in test set. QSAR is developed on the basis of the data of the remaining chemicals, and then used to make predictions for the chemicals that were omitted .This procedure is repeated a number of times, so that a number of statistics can be derived from the comparison of predicted data with the known data. Cross-validation techniques allow the assessment of the internal prediction power and robustness of the model. (Gramatica , 2013)

A cross-validated correlation coefficient R^2 (Q^2). Frequently, Q^2 is used as a criterion of both robustness and predictive ability of the model. Many authors consider high Q^2 (for instance, $Q^2 > 0.5$) as an indicator or even as the ultimate proof of the high predictive power of the QSAR model. (Veerasamy *et al* .,2011)

1.5.1.2 External model validation

Indeed, it is important to emphasize that the true predictive power of a

QSAR model can be established only through model validation procedure which consists of prediction of activities of compounds which were not included in model building, i.e., compounds in the test set. In contrast to the test set, compounds used for model building constitute the training set. In many QSAR studies multiple models are built and from the best models are selected، which are defined as those based on the prediction statistics for the test set. Thus, the test set is actually used to select models. This use of the test set for model selection practically negates the consideration of such routine as an adequate external model validation. In fact, it does not guarantee at all that models selected in this way will make accurate predictions if used for chemical database mining i.e. predicting activities of compounds in truly external database, to simulate the use of QSAR models for database mining, a so called external evaluation set is employed. It should consist of compounds with known activities that are not included in either training or test sets. External evaluation set can be selected randomly from the entire initial dataset. In general, the size of the external evaluation set should be about 15%–20% of the entire dataset. The remaining part of the dataset is called modeling set that can be divided into training and test sets . (Tropsha, 2010).

The use of internal versus external validation has been a matter of great debate . One group of QSAR workers supports internal validation, while the other group considers that internal validation is not a sufficient test for checking robustness of the models and external validation must be done. The major group of supporters of internal validation, are of the opinion that cross-validation is able to assess the model fit and to check whether the predictions will carry over to fresh data not used in the model fitting exercise. They have argued that when the sample size is small, holding a portion of it back for testing is wasteful and it is much better to use "computationally more burdensome" leave-one-out cross-validation.

The value of r^2_m(test) should be greater than 0.5 for an acceptable model. (Pratim Roy *et al*.,2009)

1.5.2 Fit of the Model:

Fit of the QSAR models can be determined by the method root-mean squared error (RMSE). These method are used to decide if the model possesses the predictive quality reflected in the R^2. The use of RMSE shows

the error between the mean of the experimental values and predicted activities.

For good predictive model the RMSE values should be <0.3. These method of error checking can also be used to aid in creating models and are especially useful in creating and validating models for nonlinear data sets, such as those created with Artificial Neural Network (ANN).

However, excellent values of R^2, RMSE are not sufficient indicators of model validity. Thus, alternative parameters must be provided to indicate the predictive ability of models. In principle, two reasonable approaches of validation can be envisaged one based on prediction and the other based on the fit of the predictor variables to rearranged response variables. (Veerasamy *et al* .,2011)

1.5.3.Applicability domain of QSAR:

The application of QSAR models depends on statistical significance and predictive ability of the models. The prediction of a modeled response using QSAR is valid only if the compound being predicted is within the applicability domain of the model. The applicability domain is a theoretical region of the chemical space, defined by the model descriptors and modeled response and thus by the nature of the training set molecules. It is possible to check whether a new chemical lies within applicability domain using the leverage approach. A compound will be considered outside the applicability domain when the leverage values is higher than the critical value of 3p/n, where p is the number of model variables plus 1 and n is the number of objects used to develop the model. (Muhammad *et al*.,2019)

1.6.Heterocyclic:

Usually heterocyclic compound, also called a heterocycle, as any of a class of organic compounds whose molecules contain one or more rings of atoms with at least one atom (the heteroatom) being an element other than carbon, most frequently oxygen, nitrogen, or sulfur . Although heterocyclic compounds may be inorganic, most contain within the ring structure at least one atom of carbon, and one or more elements such as sulfur, oxygen, or nitrogen. Since non-carbons are usually considered to have replaced carbon atoms, they are called heteroatoms. The structures may consist of either aromatic or non-aromatic rings. Heterocyclic derivatives, can be divided

into two broad area which are aromatic and non-aromatic. (Alvárez-Builla, and Barluenga, 2011)

The reasons for utlizing heterocycles are able to get involved in an extraordinarily wide range of reaction types. Depending on the pH of the medium, they may behave as acids or bases, forming anions or cations. Some interact readily with electrophilic reagents, others with nucleophiles, yet others with both. Some are easily oxidized, but resist reduction, while others can be readily hydrogenated but are stable toward the action of oxidizing agents. Many natural drugs such as papaverine, theobromine, quinine, emetine, theophylline, atropine, procaine, codeine, reserpine and morphine are heterocycles. Almost all the compounds we know as synthetic drugs such as diazepam, chlorpromazine, isoniazid, metronidazole, azidothymidine, barbiturates, antipyrine, captopril and methotrexate are also heterocycles. Some dyes (e.g. mauveine), luminophores, (e.g. acridine orange), pesticides (e.g. diazinon) and herbicides (e.g. paraquat) are also heterocyclic in nature. All these natural and synthetic heterocyclic compounds can and do participate in chemical reactions in the human body. (Dua *et al*., 2011)

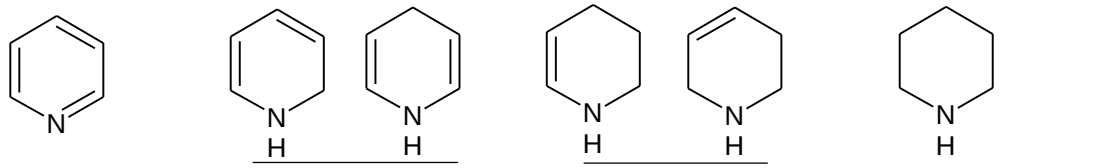


Figure (1.1) Simple heterocycle.

1.6.1 Naming simple monocyclic compounds:

The names are derived from the following four rules:

1. The heteroatom is given a name and is used as a prefix: N, aza-; O, oxa-; S, thia-; P, phospha-; As, arsa-; Si, sila-; Se, selena-, B, bora, and so on. The "a" ending is dropped if the next syllable starts with a vowel. Thus "aza-irine" is properly written "azirine."

2. Ring size is designated by stems that follow the prefix: 3-atoms,- ir-; 4-atoms, -et-; 5-atoms, -ol-; 6-atoms, -in-; 7-atoms, -ep-; 8-atoms, -oc-; 9-atoms, -on-; and so on.

3. If fully unsaturated, the name is concluded with a suffix for ring size: 3-atoms, -ene (except -ine- for N); 4-, 5-, and 6-atoms, -e; 7-, 8-, and 9- atoms, -ine.

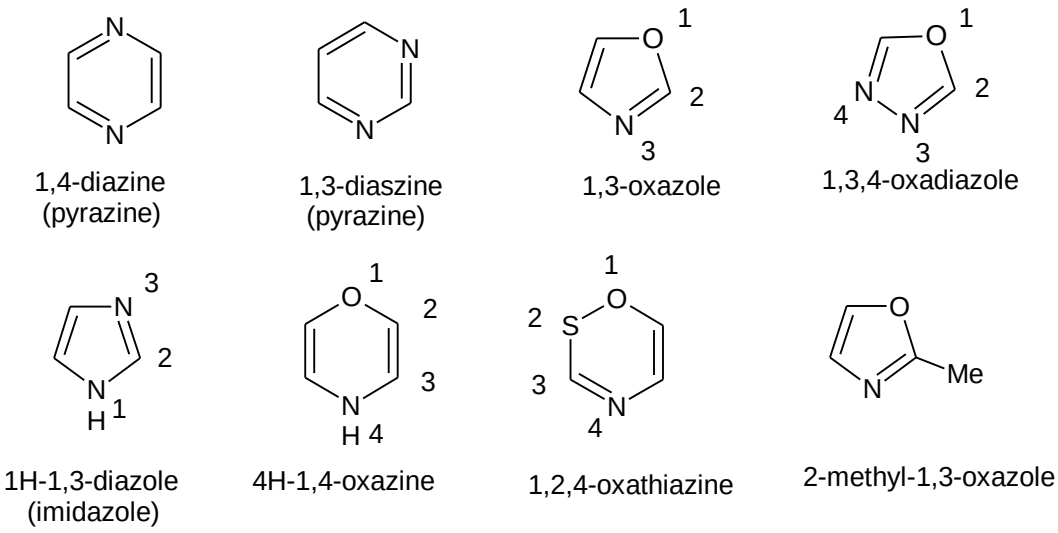
4. If fully saturated, the suffix is -ane for all ring sizes, except for N, which uses -idine for rings of 3-, 4-, or 5-atoms, and for 6- atoms, a prefix of hexahydro- is used. Also, the name oxane, not oxinane, is used for the 6-membered ring with O present. Other exceptions exist for P, As, and B rings, but they will not be given here. (Quin and Tyrell ,2010)

1.6.2.Naming the rings with more than one heteroatom:

It consider the common case where more than one heteroatom is present in the ring. The usual rules for stems to indicate prefixes for the various heteroatoms. They are listed in the following order of priorities, derived from the main groups of the Periodic System, and then within each group by increasing atomic number:

- Group I (O>S>Se>Te)>Group II (N>P>As)>Group III (Si>Ge)>. This listing can be simplified greatly by taking out the most commonly found heteroatoms in their order, which gives O>S>N>P. Each heteroatom is then given a number as found in the ring, with that of highest priority given position 1.
- A saturated heteroatom with an extra-hydrogen attached is given priority over an unsaturated form of the same atom, as in 1H-1,3-diazole
- The numbers are grouped together in front of the heteroatom listings (thus, 1,3-oxazole, not 1-oxa-3-azole).
- The heteroatom prefixes follow the numbers in the priorities given previously.).
- Punctuation is important; in the examples to follow, a comma separates the numbers and a dash separates the numbers from the heteroatom prefixes.
- A slight modification is used when two vowels adjoin; one is deleted, as in the listing for “oxaaza,” which becomes simply “oxaza.”

As for monohetero systems, substituents on the ring are listed alphabetically with a ring atom number for each (not grouped together). (Quin and Tyrell ,2010)



Figure(1.2) Some naming of multi heteroatom systems

1.6.3.Classification of heterocyclic compounds:

The Heterocycles as pyridine, thiophene, pyrrole, and furan being fused to benzene rings led to the development of quinoline, benzothiophene, indole, and benzofuran, respectively. However, fusion of two benzene rings developed a third large class of compounds, including acridine, dibenzothiophene, carbazole, and dibenzofuran. . The most common heterocycles are those having five- or six-membered rings and containing heteroatoms of nitrogen (N), oxygen (O), or sulfur (S).(Mandour *et al*., 2020)

Based on these rules, conjugated heterocyclic compounds containing $(4n+2)\pi$ electrons are aromatic(hence referred to as hetero aromatics in order to be able to recognize their heterocyclic and aromatic nature while those containing $4n\pi$ electrons cannot be aromatic even though they may be cyclic, planar and conjugated and are said to be anti- aromatic as delocalization of their π-electrons will instead lead to destabilization. By definition an aromatic compound is a planar ring of atoms linked by alternate single and double bonds. Delocalization of the π electrons of aromatic systems is a major contribution to the stabilization of these molecules and yield properties that are characteristic of aromaticity such as diamagnetic ring current. The Huckel molecular orbital theory is often used to express the relationship between a molecular orbital description of the structure and aromaticity.

There are three criteria used for evaluating aromaticity and include:

a) Energy data indicating thermodynamic stabilization or destabilization.
b) Structural data that relate to bond lengths indicating delocalized structures.
c) Electronic properties which are; energy levels, electron distribution, and polarizability.

These include the response of electrons to a magnetic field. Magnetic susceptibility measurements or NMR spectroscopy (in which aromatic compounds exhibit a diamagnetic ring Current) could be used as important experimental tools for assessing or observing aromaticity . pyridine, pyrrole and thiophene are all aromatic.(Dze *et al*.,2020)

1.7.Isoniazid:

Isoniazid, also named as isonicotinylhydrazide (INH), is a hydrazide compound derived from isonicotinic acid compound (pyridine-4-carboxylic acid), and is one of the first-line drugs for tuberculosis. INH enables KatG to form the INH-nicotinamide adenine dinucleotide (NAD) adduct. This adduct inhibits the acyl carrier protein (ACP) and InhA reductase, resulting in the synthesis of type II fatty acids, which in turn will synthesize mycolic acid, causing cell death.(Ruswanto *et al*.,2019).

1.7.1.Isoniazid Synthesis:

Isoniazid is prepared through the reaction of 4-cyanopyridine and hydrazine hydrate in an aqueous alkaline medium at 100C under reflux for 7 hours with subsequently crystallization in ethanol, thereby leading to the desired compound with 62% of yield . (Fernandes *et al*., 2017)

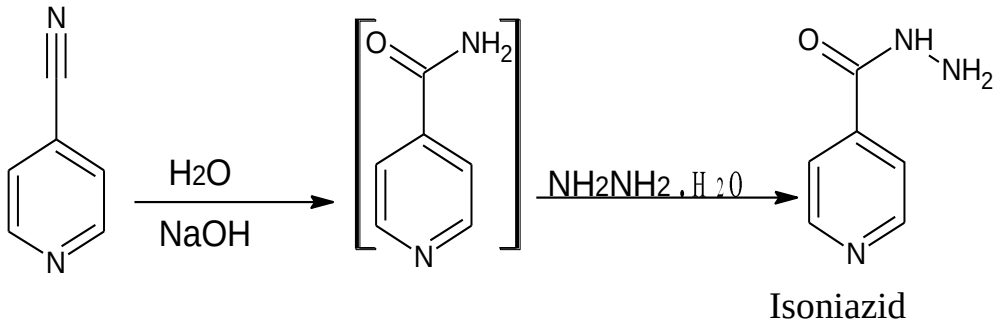


Figure (1.3) Synthesis of isoniazid

1.7.2.Biological properties:

Oral, intravenous and intramuscular are routes of administration for isoniazid, but the most common is the oral route. Isoniazid is promptly absorbed after oral administration and reaches the peak of seric concentrations in 0.5–2 hours within 100% of bioavailability in the most of the cases .Co-administration of isoniazid with food significantly reduces its bioavailability. Moreover, there is evidence that suggest that HIV-positive patients may exhibit poor absorption of isoniazid.

Isoniazid is widely distributed throughout body fluids and tissues, with a volume of distribution of approximately 61% of body weight and its plasma protein binding is very low. Acetylation is the main metabolic transformation that occurs with isoniazid. Specifically, this metabolic process is strongly influenced by genetic aspects. The predominant liver metabolism justify the high risk of hepatotoxicity, especially for those TB-patients using an association with rifampicin. The elimination half-life of isoniazid and its metabolites is 0.5 to 4 hours and their main elimination route is through the kidney, with 75% to 96% of the drug and metabolites excreted in urine within 24 hours .

Upon absorption, isoniazid is metabolized *invivo* and leads to the formation of several metabolites, namely acetylisoniazid, isonicotinic acid, isonicotinamide, monoacetylhydrazine and diacetylhydrazine . Therefore, it is important the simultaneous detection of these metabolites through the analytical methods. (Fernandes *et al*., 2017)

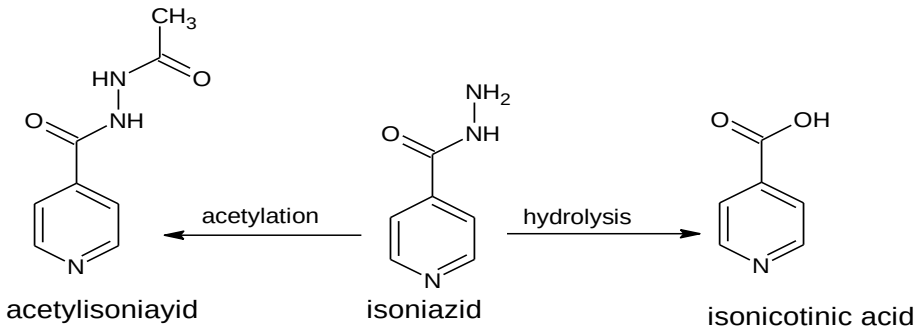


Figure (1.4) Principal metabolites of isoniazid detected by analytical methods.

metabolites of isoniazid, especially acetylisoniazid and isonicotinic acid, which are the acetylated and hydrolyzed metabolites respectively in Figure(1.4**)**

1.8 Aims and objectives of the study:

- Development QSAR(Quantitative Structure-Activity Relationship) model by using ACD/Labs, MOE, SPSS(software program) and data set to predict the biological activity of some designed derivatives compounds in (2D,3D and 2D+3D) dimensionals.
- Applying the molecular docking study on data set and designed derivatives with 4TRO receptor.
- Selecting some of designed derivative compounds and data set , which have more interaction with the receptor and higher biological activity to discuss .

Chapter Two

Materials and Methods

2. Materials and Methods

2.1 Materials

2.1.1 MOE 2009.10:

Molecular operating environment (Version 10) is a drug discovery software platform that integrates visualization, modeling and simulations, as well as methodology development, in one package. MOE scientific applications are used by biologists, medicinal chemistry and computational chemists in pharmaceutical, biotechnology and academic research.
Main applications area in MOE include structure-based design, fragment–based design, pharmacophore discovery, medicinal chemistry applications , biologics applications, protein and antibody modeling, molecular modeling and simulations ,chemo informatics and QSAR.

2.1.2 ACD/ labs program:

ACD/Labs (Version 14.01) is a two-dimensional structure drawing program primarily designed for organic molecules. Although the drawing mode is essentially a two-dimensional drawing routine, it is also possible to rotate the molecule in three dimensions. The program automatically keeps track of the number of hydrogens bonded to each atom.

2.1.3 SPSS program:

SPSS (Version 16.020) is a software package used for interactive ,or batched, statistical and statistical analysis. Can be used to perform data entry and to create tables and graphs. SPSS is allowing ordinary researchers to do their own statistical analysis and handling complex data manipulations and analysis.

2.1.4 Data set:

Twenty one(21) compounds of N-substituted-2-Isonicotinoylhyrazinecarboxmides derivative were taken from the literatures which are given in table 1(Rychtarčíková *et al*., 2014).The minimum inhibitory concentration (MIC) values were converted to the logarithmic scale (pMIC values)to ensure normal distribution in statistical analyses, and these were used as the dependent variable in the QSAR analyses.

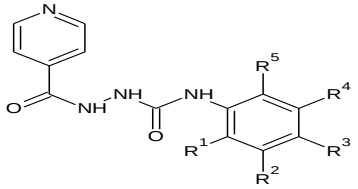
2.2. Methods:

2.2.1 QSAR data set analysis:

For the development of the model, the 21compounds N-substituted-2-Isonicotinoylhyrazinecarboxmides derivative gathered from (Rychtarčíková *et al*., 2014) used in the study were divided into training and test set .The training set is used in model construction and in internal validation while the test set for external validation.

The(21)compounds were divided randomly into the training set (15 compounds) and the test set (6 compounds) were showed in table (2-1) ,(2-2) respectively and one compound was excluded from the test set to make the result of the validation more accurate . All compounds were sketched by ACD/Labs, some descriptors were calculated by using MOE2009 software ,the descriptors should be correlated to each other according to the matrix relationship must be less than 50. The MIC value of the data set is converted to a logarithmic scale with the aim that the range of MIC values between compounds does not differ significantly, and the distribution of MIC values is getting better . The training set compounds were analyzed statistically to produce the QSAR models by using partial least square method (PLS). The values of three physicochemical parameters were selected as independent variables while biological activity (PMIC) as dependent variable to generated the models in(2D and 3D) (2D,3D and 2D+3D) dimensional ,The expected biological activity values were calculated from the derived equation in (2D,3D and 2D+3D) for the data set showed also in table (2-1)(2-2) for training set and test set respectively .

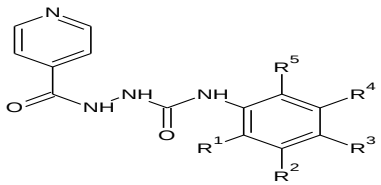
Table (2-1) Training set compounds of N-substituted-2-Isonicotinoyl hydrazine carboxmides derivative , ,observed MIC,observed pMIC ,Predicted pMIC in(2D ,3D and (2D+3D)) , residuals and descriptors in (2D and3D) dimensional



Compound	3a	3b	3c	3d	3i	3j	3k	3l
R1	-H	-H	-H	-H	-H	-H	-Br	-F
R2	-H	-H	-H	-H	-H	-H	-H	-H
R3	-Methyl	-iso Propyl	-tret -butyl	-octyl	-OMe	$-CF_3$	-Br	-Br
R4	-H	-H	-H	-H	-H	-H	-H	-H
R5	-H	-H	-H	-H	-H	-H	-H	-H
Observed MIC	62.5	16	8	1	32	62.5	8	62.5
pMIC	1.796	1.204	0.903	0.000	1.505	1.796	0.903	1.796
Predicted pMIC Model 1 (2D)	2.0295	1.1235	0.840	0.6868	1.8336	1.639	0.917	1.499
Predicted pMIC Model 5 (2D)	1.932	1.1381	0.870	0.6674	1.9500	1.6367	0.7922	1.4248
Predicted pMIC Model 6 (3D)	1.815	1.181	0.900	0.593	1.623	1.680	0.783	1.701
pMIC Predicted Model 11 (2D+3D)	1.8156	1.1817	0.900	0.5932	1.7018	1.6235	0.7838	1.6235
Residual model 1	-0.2335	0.0805	0.062	-0.0848	-0.3285	0.157	-0.014	0.297
Residual model 5	-0.136	0.0659	0.032	-0.6674	-0.4449	0.1593	0.1108	0.3712
Residual model 6	0.0196	0.0222	0.002	0.0087	0.0941	0.1159	0.1192	-0.1185
Residual model 11	-0.0466	0.0223	0.003	-0.593	-0.1968	0.1725	0.1192	0.1725
Slog P_VSA9 (2D)	59.944	93.270	126.5	59.944	31.862	71.499	118.52	87.533
Kier2(2D)	8.444	9.333	8.909	10.78	9.209	8.909	8.585	8.585
chi0v(2D)	10.88	12.45	13.38	13.00	11.29	11.51	13.73	12.14
LogP(o/w) (2D)	1.2200	2.0650	2.421	2.5790	0.8780	1.8568	2.5530	1.9080
KierA3 (2D)	3.4764	3.9753	4.212	4.7075	3.7118	4.0890	4.6288	3.9938
chi1v (2D)	5.9900	6.9334	7.240	7.5507	6.1024	6.3070	7.3707	6.5777
E_oop (3D)	0.0485	0.0417	0.040	0.0422	0.1421	0.0534	0.3904	0.0452
ASA (3D)	510.44	567.58	586.4	536.48	534.98	536.70	551.89	551.89
npr1 (3D)	0.207	0.116	0.103	0.207	0.098	0.106	0.207	0.098

Compound	3m	3o	3q	3r	3s	3t	3u
R1	-Br	-H	-H	-H	-H	-H	-Cl
R2	-H	-Cl	CF_3	CF_3	-H	-Cl	-H
R3	-F	-H	-Br	-F	-Br	-Cl	-Cl
R4	-H	-H	-H	-H	-H	-H	-H
R5	-H	-H	-H	-H	-H	-H	-Cl
Observed MIC	16	62.5	8	32	62.5	32	4
pMIC	1.204	1.796	0.903	1.505	1.796	1.505	0.602
Predicted pMIC Model 1 (2D)	1.4990	1.7450	0.8289	1.4101	1.4990	1.4099	0.8538
Predicted pMIC Model 5 (2D)	1.4248	1.7386	0.8848	1.5231	1.4248	1.430	0.9776
Predicted pMIC Model 6 (3D)	1.101	1.906	1.011	1.512	1.788	1.4588	0.7571
pMIC Predicted Model 11 (2D+3D)	1.1018	1.9060	1.0113	1.9080	1.7881	1.4588	0.7572
Residual model 1	-0.295	0.051	0.0741	0.095	0.297	0.0952	0.2518
Residual model 5	-0.2208	0.0574	0.0182	-0.018	0.3712	0.0751	-0.3756
Residual model 6	0.1021	-0.1100	-0.1083	-0.0077	0.0078	0.0461	-0.1551
Residual model 11	0.1022	-0.11	-0.1083	-0.403	0.0079	0.0462	-0.1552
Slog P_VSA9 (2D)	87.533	80.727	117.45	86.459	87.533	104.91	144.06
Kier2(2D)	8.585	8.585	9.087	9.087	8.585	8.585	8.740
chi0v(2D)	12.147	11.317	13.403	11.817	12.147	12.072	13.129
LogP(o/w) (2D)	1.9080	1.7020	2.6898	2.0448	1.9080	2.2170	2.7680
KierA3 (2D)	3.9938	3.7651	4.6728	4.0971	3.9938	4.3874	4.4900
chi1v (2D)	6.5777	6.1626	7.2056	6.4126	6.5777	6.5346	7.0242
E_oop (3D)	0.4618	0.0389	0.0640	0.0627	0.0372	0.0461	0.5060
ASA (3D)	518.86	506.69	561.37	543.47	525.00	527.51	548.58
npr1 (3D)	0.248	0.1995	0.2094	0.1223	0.1369	0.2455	0.1857

Table (2-2) Test set compounds of N-substituted-2-Isonicotinoyl hydrazine carboxmides derivative , observed MIC, observed pMIC ,Predicted PMIC in (2D , 3D and (2D+3D)), residuals ,descriptors in (2D and 3D) dimensional



Compound	3e	3f	3h	3n	3p
R1	-H	-H	-H	-H	-H
R2	-H	-H	-H	-Cl	$-CF_3$
R3	-n-butyl	-pentyl	-hexyl	-Cl	-Cl
R4	-H	-H	-H	-H	-H
R5	-H	-H	-H	-H	-H
Observed MIC	4	32	32	125	32
pMIC	0.602	1.505	1.505	2.0978	1.505
Predicted pMIC Model 1 (2D)	0.2304	-0.2299	-1.1608	1.4100	1.0750
Predicted pMIC Model 5 (2D)	0.2757	-0.1169	-0.8937	1.4159	1.1983
Predicted pMIC Model 6 (3D)	1.1334	-0.440	-1.375	-1.577	1.315
pMIC Predicted Model 11 (2D+3D)	0.2377	-0.2631	-1.1273	1.5465	1.1924
Residual model 1	0.3716	1.735	2.6659	0.6878	0.4301
Residual model 5	0.3263	1.622	2.3988	0.6819	0.3068
Residual model 6	0.4686	1.9451	2.8801	3.6748	0.1901
Residual model 11	0.3643	1.7681	2.6323	0.5513	0.3126
Slog P_VSA9 (2D)	11.583	12.398	14.057	8.585	9.087
Kier2(2D)	59.94	59.94	59.94	104.91	110.64
chi0v(2D)	13.71	14.41	15.83	12.07	12.57
LogP(o/w) (2D)	3.0210	3.4630	4.3470	2.1410	2.4838
KierA3 (2D)	5.2514	5.7868	6.9405	4.1331	4.4446
Chi1v (2D)	8.0507	8.8807	9.5507	6.5406	6.7906
E_oop (3D)	327.37	345.87	382.25	270.37	281.87
ASA (3D)	636.14	674.48	737.13	530.52	555.09
npr1 (3D)	0.1138	0.1088	0.0972	0.1856	0.1344

2.2.2 The validation of QSAR models:

N-substituted-2-Isonicotinoyl hydrazine carboxamides derivative compounds were modeled by using two-dimensional and three dimensional, the parameters were calculated by using MOE 2009 software**.**

The QSAR model was generated by the PLS method based on electronic parameters this method refers to the dependent variable y (biological activity) with a number of independent variables x (electronic descriptors) using linear regression, the QSAR internally validating by calculating R^2 ,Q^2, as well as externally by calculating R^2_{pred} to give models robustness all that was aided by MOE software. Moreover, MLR linear regression method estimates the values of the regression coefficients by applying some statistical parameters such as R^2, standard estimation of error (SEE), F-ratio between the variance of prediction and observation activity and p-value all using SPSS software.

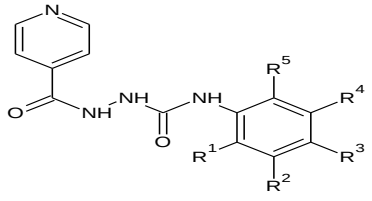
2.2.3 The molecular modeling design :

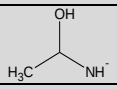
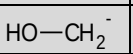
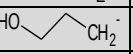
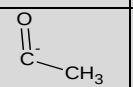
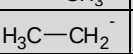
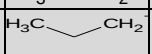
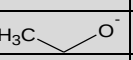
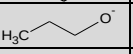
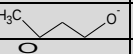
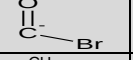
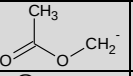
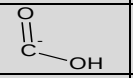
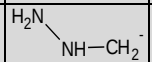
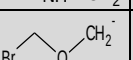
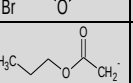
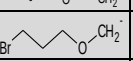
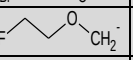
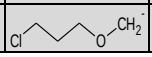
58 new compounds of N-substituted-2-Isonicotinoyl hydrazine carboxamides derivatives were modeled using ACD/Labs in table(2-3) .The selected parameters in the development of(QSAR) modeling were calculated for the 58 compounds.

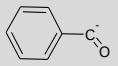
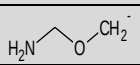
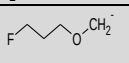
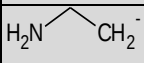
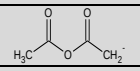
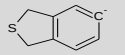
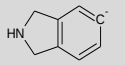
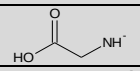
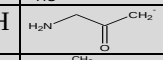
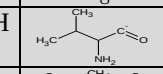
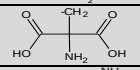
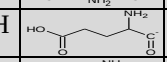
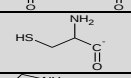
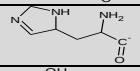
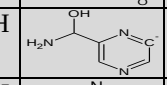
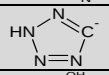
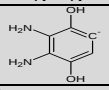
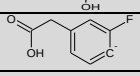
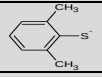
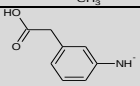
2.2.4 The application of QSAR models:

The obtained models of QSAR were applied on the new designed compounds predicted the biological activity of the new compounds ,2D-QASR , 3D-QSAR, or (2D+3D QSAR) models can be used to calculate the biological activity after validation of the models internally and externally.

Table(2-3) The designed new N-substituted-2-Isonicotinoyl hydrazine carboxmides derivatives and descriptors



No	Comp-ound	R_1	R_2	R_3	R_4	R_5	chi0v 2D	Kier2 2D	SlogP_VSA9 2D	Logp (o/w) 2D	khierA3 2D	chi1v 2D	ASA 3D	E-oop 3D	npr1 3D
1	I	-H	-H	-Cl	H	-H	12.29	10.09	62.701	1.514	3.804	1.857	515.1	0.0385	0.174
2	II	-H	-H		-H	-H	12.29	10.09	62.701	0.414	4.574	6.864	571.5	0.1111	0.175
3	III	-Br	-H	-H	-H	-H	11.84	8.444	72.573	1.718	3.800	6.478	526.7	0.0179	0.117
4	IV	-NH_2	-H	-H	-H	-H	10.45	8.444	29.375	0.248	3.152	5.784	512.5	0.0421	0.193
5	V	-H	-H	-OH	-H	-H	10.32	8.444	31.862	0.614	3.476	5.713	507.4	0.0175	0.058
6	VI	-H	-H	-Br	-H	-H	11.84	8.444	72.573	1.720	4.047	6.472	530.5	0.1884	0.114
7	VII	-OH	-H	-OH	-H	-H	10.69	8.585	37.105	0.341	3.496	5.853	519.5	0.2712	0.067
8	VIII	-OH	-H	-OH	-H	-OH	11.06	8.740	42.349	0.068	3.557	5.994	514.5	0.0213	0.083
9	IX	-H	-OH	-H	-H	-H	10.32	8.444	31.862	0.651	3.476	5.713	505.5	0.0168	0.066
10	X	-Cl	-H	-Cl	-H	-H	12.22	8.585	102.159	2.141	4.133	1.415	538.2	0.3740	0.074
11	XI	-H	-H		-H	-H	11.24	9.209	23.862	0.400	3.711	2.015	530.6	0.0211	0.055
12	XII	-H	-H		-H	-H	12.65	10.78	23.862	0.930	4.707	7.159	594.6	0.0187	0.071
13	XIII	-H	-Br	-H	-Cl	-H	13.11	8.585	108.96	2.423	4.645	6.949	554.0	0.0146	0.107
14	XIV	-H	-NH_2	-H	-NH_2	-H	11.16	8.585	23.862	-0.311	3.478	5.978	523.4	0.2073	0.176
15	XV	-H	-NH_2	-H	-H	-H	10.66	8.444	23.862	0.287	3.359	5.778	516.1	0.0194	0.066
16	XVI	-H	-H	-NH_2	-H	-H	10.66	8.444	23.862	0.250	3.359	5.778	507.3	0.1267	0.187
17	XVII	-H	-NO_2	-H	-Cl	-H	12.40	9.474	63.010	1.560	4.294	6.556	550.9	0.0899	0.210
18	XVIII	-H	-H		-H	-H	11.99	9.333	57.188	0.775	3.767	6.444	546.2	0.2056	0.120
19	XIX	-H	-H		-H	-H	11.79	9.209	57.188	1.695	3.711	6.550	547.2	0.1987	0.127
20	XX	-H	-H		-H	-H	12.50	9.988	57.188	2.137	4.212	7.050	579.2	0.0196	0.062
21	XXI	-H	-H		-H	-H	12.20	9.988	57.188	1.219	4.212	6.689	553.7	0.1939	0.162
22	XXII	-H	-H		-H	-H	12.91	10.78	57.188	1.833	4.707	1.080	599.9	0.0423	0.160
23	XXIII	-H	-H		-H	-H	13.61	11.58	57.188	2.275	5.251	0.688	617.1	0.0178	0.055
24	XXIV	-H	-H		-H	-H	12.96	9.333	69.816	1.705	4.322	6.926	555.2	0.0774	0.116
25	XXV	-H	-H		-H	-H	13.11	10.87	57.188	0.989	5.005	7.040	588.0	0.2001	0.133
26	XXVI	-H	-H		-H	-H	11.44	9.333	23.862	0.599	3.767	6.167	532.1	0.2090	0.127
27	XXVII	-H	-H		-H	-H	11.69	9.988	23.862	-0.177	4.212	6.485	549.6	0.0703	0.197
28	XXVIII	-H	-H		-H	-H	13.87	10.78	69.816	1.882	5.349	7.809	607.3	0.1770	0.159
29	XXIX	-H	-H		-H	-H	14.52	12.45	57.188	1.906	5.790	8.101	637.7	0.0603	0.171
30	XXX	-H	-H		-H	-H	15.29	12.39	69.816	2.214	6.493	8.809	683.1	0.0839	0.071
31	XXXI	-H	-H		-H	-H	12.99	11.58	38.822	1.217	5.202	7.188	604.4	0.2168	0.172
32	XXXII	-H	-H		-H	-H	14.46	12.38	63.010	2.008	6.193	8.222	660.9	0.0216	0.057

No	Comp-ound	R_1	R_2	R_3	R_4	R_5	chi0v 2D	Kier2 2D	SlogP_VSA9 2D	Logp (o/w) 2D	khierA3 2D	Chi1v 2D	ASA 3D	E-oop 3D	npr1 3D
33	XXXIII	-H	-H		-H	-H	14.38	11.25	23.862	2.437	4.106	8.104	634.1	0.2074	0.110
34	XXXIV	-H	-H		-H	-H	11.37	9.333	38.822	1.060	3.728	6.133	522.6	0.0926	0.172
35	XXXV	-H	-H		-H	-H	12.48	10.78	23.862	0.206	4.707	6.829	583.8	0.2016	0.130
36	XXXVI	-H	-H		-H	-H	13.70	12.39	38.822	1.569	5.735	7.688	638.2	0.0248	0.052
37	XXXVII	-H	-H		-H	-H	12.08	9.988	23.862	0.266	4.212	6.751	570.7	0.2216	0.124
38	XXXVIII	-H	-H		-H	-H	13.47	10.36	23.862	2.882	3.800	7.650	610.6	0.0205	0.045
39	XXXIX	-H	-H	$-NO_2$	-H	-H	11.35	9.333	23.862	0.857	3.740	6.078	525.5	0.0528	0.169
40	XL	-H	-H		-H	-H	14.02	11.75	57.188	0.928	5.541	7.513	644.5	0.0291	0.052
41	XLI	-H	-H		-H	-H	15.96	10.34	90.548	3.463	3.996	9.917	662.4	0.1870	0.100
42	XLII	-H	-H		-H	-H	15.23	10.34	60.741	2.518	3.737	8.892	669.4	0.0242	0.039
43	XLIII	-H	-H		-H	-H	12.65	10.87	23.862	0.074	4.870	6.864	578.3	0.2763	0.164
44	XLIV	-H	-H		-H	-H	13.10	10.87	23.862	-0.307	4.971	7.326	603.7	0.0172	0.062
45	XLV	-H	-H		-H	-H	14.58	11.40	23.862	1.167	4.763	8.054	612.1	0.1890	0.173
46	XLVI	-H	-H		-H	-H	14.72	11.11	90.514	-0.658	5.754	0.892	638.5	0.0210	0.073
47	XLVII	-H	-H		-H	-H	14.72	11.11	90.5142	-0.111	5.768	8.255	530.3	0.5533	0.424
48	XLVIII	-H	-H		-H	-H	14.20	10.98	23.860	0.305	4.856	7.923	647.5	0.0291	0.070
49	XLIX	-H	-H		-H	-H	15.46	12.14	23.860	-0.410	5.084	8.932	677.7	0.3037	0.203
50	L	-H	-H		-H	-H	14.74	11.40	23.862	-0.706	4.490	8.161	652.3	0.0385	0.049
51	LI	-H	-H		-H	-H	12.43	9.629	23.862	0.437	8.084	1.492	563.9	0.0166	0.044
52	LII	-H	-H		-H	-H	15.21	11.03	39.862	0.015	3.927	8.341	667.0	0.1187	0.054
53	LIII	-H	-H		-H	-H	15.76	12.29	38.822	2.835	4.993	8.801	685.3	0.1056	0.116
54	LIV	-CN	-H	$-NH_2$	-H	-SH	12.74	9.474	23.862	0.469	3.696	6.756	572.7	0.5082	0.123
55	LV	-CN	-H	$-NH_2$	-H		11.32	9.333	29.375	-0.055	3.342	6.169	553.0	0.0122	0.090
56	LVI	-F	-H	-F	-H	-F	11.01	8.740	68.742	1.451	3.4472	5.890	506.4	0.0152	0.083
57	LVII	-H	-Br		-H	-H	18.43	11.57	168.687	3.760	4.853	9.458	698.2	0.0905	0.206
58	LVIII	-H	-Cl		-Cl	-H	18.07	13.18	104.916	2.979	6.189	10.18	713.7	0.5401	0.094

2.2.5 Docking:

2.2.5.1 Protein preparation:

The energy minimized and docking studies by using AutoDock Tools MOE (2009.10) software crystal structure of (4TRO) protein receptor of *mycobacterium tuberculosis* from PDB (Protein data bank) the charges were added to the protein molecule followed by addition of polar hydrogen atoms and saved the charged protein molecule after removed of unwanted molecules as water molecule and kept all unique ligands in the protein and save in pdb format. Enoyl-ACP reductase of Mycobacterium tuberculosis (PDB: 4TRO) implicated in the biosynthesis of mycolic acids, essential constituents of the mycobacterial cell wall. This enzyme is considered as a promising target for the discovery of novel antitubercular drugs , The protein 4tro are obtained from PDB(Protein Data Bank) was showed in figure (2.1) below in pdb form before prepare it.

The figure (2.2) were showed clear image to the receptor that used in docking ,and all the unique ligands are appeared in the three dimensional(3D) after prepared by using MOE program.

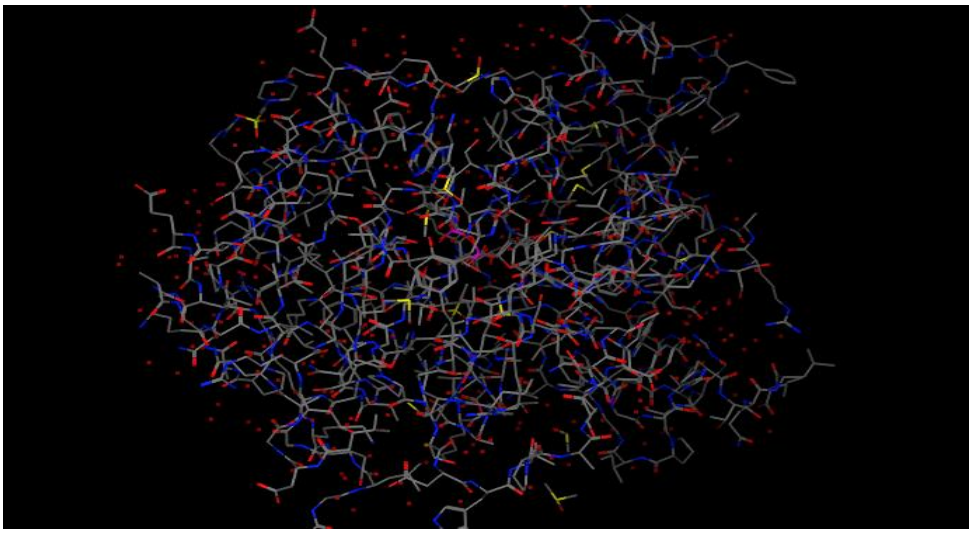


Figure (2.1) The receptor 4TRO in (PDB) form.

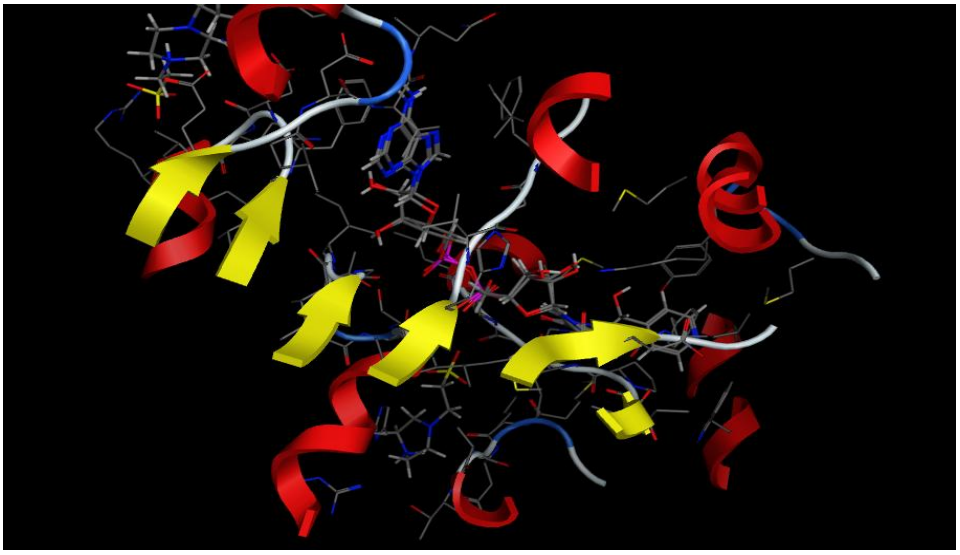


Figure (2.2) The receptor in moe form (3D)dimensional

2.2.5.2 Preparation of the ligands:

The Designed 58 new derivatives compounds were used as ligands and sketched by using ACD/Labs software .Table (2-3)

The energy of the new compounds were minimized by using MOE 2009.10 software and saved in (mdb) file .All the biological activity of ligands were predicted then were performed the molecular docking for them.

Chapter Three

Results and Discussion

3. Results and discussion:

3.1 Modeling Out Put:

3.1.1 Development of QSAR models :

The training set compounds were analyzed statistically to produce the QSAR model by using partial least square method (PLS) in MOE program, while the test set compounds were used to validate the QSAR model that generated by the training set.

The two groups of data were showed in tables (2-1),(2-2). QSAR model development depends on biological activity of training set. These two groups were correlated with different three descriptors in (2D, 3D and 2D+3D) dimensionals. The descriptors were selected according to the correlation matrix ,the relationship must be 50 or less than 50 to be used in the QSAR models .The selected descriptors in (2D) dimensional ,chi0v, slogp_VSA9, keir2, kierA3, chi1v, BCUT_SMR_1 ,SMR_VSA7,balabanJ and logp(o/w).In (3D) dimensional ASA, E_oop , npr1 , vol , std_dim , E_tor .

The best QSAR models were used to predict the biological activity of the new derivatives in (2D) two dimensional and in (3D)three dimensional. Noticeable that the biological activity (experimental and predicting) values are close to each other in the (2D) two dimensional and in (3D)three dimensional .table (3-1)(3-2) respectively

Log P(o/w) describe the molecular dimensions and hydrophobicity, respectively, for successful binding with the substrate. Log P(o/w) is the log value of the octanol/water partition coefficient.(Cotua *et al* .,2021)

chi1v, which is a topological descriptor related to molecular shape. (Bernal and Schmidt, 2019) Second kappa shape index: $(n\text{-}1)^2 / m^2$

KierA3: Third alpha modified shape index .(Sakagami et al., 2015).Is topological descriptors .The KierA3 descriptor describes shape of the molecules with third alpha shape index. It calculated by (s-1)(s-3)2 /p3 2 for odd n and (s-3)(s-2)2 /p3 2 for even n, where s = n+a. However, Kier and Hall kappa molecular shape indices compare the minimal and maximal molecular graphs and are intended to capture different aspects of molecular shape.(Moorthy *et al*., 2014)

Balaban J , which is a topological descriptor. (Roy and Ghosh,2010)

SMR: SMR is the molecular refractivity obtained from an atomic contribution model. Weight is the molecular. (Marighetti,2019).

slogp-VSA9,VSA-SMR1: SLOGPVSA9 measure the total surface area for logP atomic contributions . (Rosas-jimenez *et al*., 2020)

The nature and calculation of van der Waals surface area (VSA) descriptors is well. The VSA-type descriptors are based on the van der Waals surface area (VSA). The VSA for each atom in a molecule is obtained from the surface area of the atom, . . The calculation of the descriptors occurs in two steps: Firstly, the VSA for each atom of a molecule is calculated. The calculated atomic VSAs are successively combined with physicochemical properties as molar refractivity (MR), lipophilicity (logP). (Marighetti,2019)

BCUT-SMR1: For BCUT descriptors, the Burden matrix is defined as an adjacency matrix where the diagonal elements are selected atomic properties and the o-diagonal elements Bij , only for adjacent atoms, take the value of π-1/2, where π is the conventional bond order. Remaining elements take the arbitrary value of 0.001. The matrix eigen values are calculated and the smallest, second, third, and largest eigen values are used as descriptors The GCUT descriptors differ to BCUT descriptors only for the definition of the used matrix. For calculation of GCUT descriptors, the values of off-diagonal entries are calculated as d-1/2, where d is the graph distance between the two atoms. MOE uses as atomic properties for the matrix diagonal the PEOE partial charges, the atomic contribution to logP and the molar refractivity (both calculated using the Wildman and Crippen SlogP/SMR methods .(Marighetti,2019)

std-dim3,dens,ASA, Vol: These descriptors are correlated with the molecular conformation and are useful to describe dimensional parameters (like volume and surface) of molecules.

The van der Waals surface area is calculated as the sum of the atomic van der Waals surface areas not included in other atoms. Analogously, the van der Waals volume is calculated as the volume included in the van der Waals surface area.

The value of globularity indicates how is the molecule in the bulk extended.

Std_dim3 : Descriptor type is shape definition and Standard dimension .(Cotua *et al*.,2021)

ASA: The solvent-Accessible Surface Area (ASA) is one such descriptor which measures the exposure of a residue to solvent (water) in its folded state .(Hanson *et al*., 2019):

VOL: The values of molecular volume and name it Vw (w as Van der Waals).(Laffort ,2020)

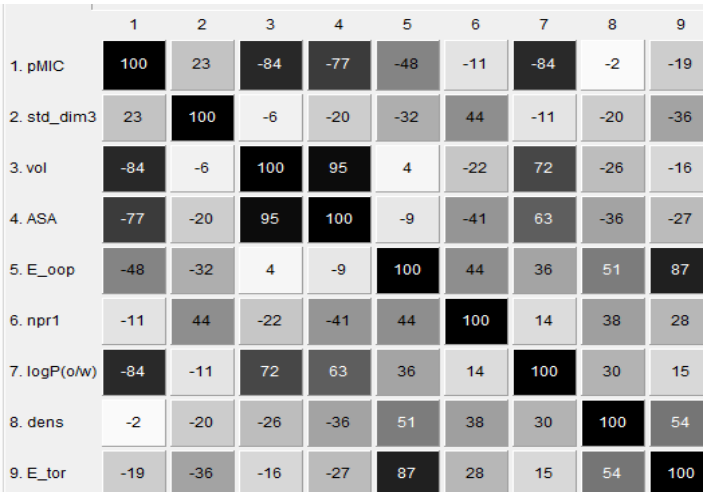
E_tor: Torsion (proper and improper) potential energy .(Sakagami *et al*., 2015)

A correlation matrix is a table showing correlation coefficients between variables. Each cell in the table shows the correlation between two variables . A correlation matrix help in selected the variables which used to build the QSAR model. If the relationship between the two variables is less than 50,it will be a better choice.

Figures (3-1) ,(3-2) elucidate the correlation coefficients between variables which used in build the QSAR models in 2D and 3D respectively .

	1	2	3	4	5	6	7	8	9	10
1. pMIC	100	-52	-89	-91	-45	-81	-61	-84	-59	-21
2. BCUT_SMR_1	-52	100	57	38	-26	52	72	54	63	61
3. chi0v	-89	57	100	96	25	86	74	88	65	30
4. chi1v	-91	38	96	100	49	86	57	87	53	9
5. Kier2	-45	-26	25	49	100	41	-31	25	-21	-40
6. KierA3	-81	52	86	86	41	100	58	90	39	34
7. SlogP_VSA9	-61	72	74	57	-31	58	100	80	74	64
8. logP(o/w)	-84	54	88	87	25	90	80	100	52	45
9. SMR_VSA7	-59	63	65	53	-21	39	74	52	100	15
10. balabanJ	-21	61	30	9	-40	34	64	45	15	100

Figure(3.1) Correlation matrix of training set compounds in (2D) two dimensional



	1	2	3	4	5	6	7	8	9
1. pMIC	100	23	-84	-77	-48	-11	-84	-2	-19
2. std_dim3	23	100	-6	-20	-32	44	-11	-20	-36
3. vol	-84	-6	100	95	4	-22	72	-26	-16
4. ASA	-77	-20	95	100	-9	-41	63	-36	-27
5. E_oop	-48	-32	4	-9	100	44	36	51	87
6. npr1	-11	44	-22	-41	44	100	14	38	28
7. logP(o/w)	-84	-11	72	63	36	14	100	30	15
8. dens	-2	-20	-26	-36	51	38	30	100	54
9. E_tor	-19	-36	-16	-27	87	28	15	54	100

Figure (3.2) Correlation matrix of training set compounds in (3D)three dimensional

The models of (QSAR) :

QSAR model of training data was reported in (Table 3-1) and (Table3-2) based on dependent variable correlated with different three descriptors (physicochemical parameters) independent, in (2D) and (3D) to get the equation of linear correlation . The best equations models of QSAR linear regression as follows below in table (3-3) in 2D dimensional and in 3D dimensional .For good predictive model the .RMSE values should be < 0.3, from model (1) (6), respectively= (0.155, 0.084) that indicate they are good models.

Table (3-1) Equations of QSAR model (in 2D dimensional) ,correlation coefficient(R^2) and root mean square error(RMSE)

No. Equation	**Equation**	**R2**	**RMSE**
Model 1	PMIC=8.04413-0.24582chi0v-0.00617slogp_VSA9-0.35165kier2	0.87098	0.15561
Model 2	PMIC=5.53099+0.11200kierA3-0.83579chi1v-2.05772BCUT_SMR_1	0.86930	0.15662
Model 3	PMIC=6.50454-0.13405kierA3-0.67353chi1v-0.00216SMR_VSA7	0.8585	0.16505
Model 4	PMIC=9.37683+0.11316khierA3-0.89950chi1v-1.51282balabanJ	0.85414	0.16546
Model 5	PMIC=6.07370+0.10173kierA3-0.70720chi1v-0.21145logp(o/w)	0.84600	0.17001

Table (3-2) Equations of QSAR model (in 3D and 2D+3D dimensional),correlation coefficient(R^2)and root mean square error(RMSE)

No. Equation	Equation	R2	RMSE
Model 6	PMIC=10.25002-0.01536ASA-1.15755E_oop-2.58249npr1	0.96017	0.08646
Model 7	PMIC=7.65042-0.02349vol-4.14115npr1+0.92373std_dim	0.94135	0.10492
Model 8	PMIC=7.75149-0.00340ASA-1.26587E_oop-0.01587vol	0.92318	0.12007
Model 9	PMIC=7.80302-0.01221ASA-2.34721E_oop+0.04596E_tor	0.92213	0.12089
Model10	PMIC=7.16302-0.02082vol+0.10694std_dim3-1.15447E_oop	0.92188	0.12108
Model 11	PMIC=7.5644-1.1662E_oop-0.0103ASA-0.2475logp(o/w)	0.9340	0.1109

R-squared (R^2) is a statistical measure of how close the data are to the fitted regression line.

3.1.2 Validation and statistical analysis of QSAR models:

The QSAR models were validated internally and externally to determine how good models in predict the anti-tuberculosis activity of the N-Substituted 2-Isonicotinoylhydrazinecarboxamides derivatives. MLR statistical validation for searching the best model was obtained by calculation for three equation in (2D) , three equation in (3D) and one equation in (2D+3D) . QSAR model acceptance criteria of Golbraikh and Tropsha's Model is chosen with higher values of R^2, R^2_{test} and Q^2_{cv} (Chtita *et al* ., 2021)

QSAR model is considered predictive if $R^2 > 0.6$. It consider the main R^2 of correlation coefficient the figure (3.3),(3.6),(3.9) and (3.12) explain that in (2D, 3D and (2D+3D)) = 0.87, 0.84 ,0.96 and 0.93 respectively . Internal validation **(**Q^2LOO (Leaving-One-Out)),LOO cross-validation coefficient , R^2 cross validation **)**or $Q^2 > 0.5$ value indicate excellent predictive quality of the developed QSAR models using as internal validation . A high value of this statistical characteristic of the Q^2 in figure (3.4) ,(3.7),(3.10) and (3.14) from (2D, 3D and (2D+3D)) = 0.73, 0.70, 0.91 and 0.89 respectively so could used to predict the biological

activity of new designed N-Substituted 2-Isonicotinoylhydrazinecarboxamides derivatives compounds .

Internal validation take place, the difference between Q^2 and R^2 should be less than 0.3 ,which means model is strong, so the value from Figure (3.6), (3.7) Q^2, R^2 in (3D) respectively 0.91, 0.96 the value of subtracting them equals 0.05 so it is considered good model. .

-Model (6) was correlated with different molecular descriptors ,(npr1) number of molecule ,(E-oop) out- of- plane energy and water accessible surface areas (SAS) consider to the matrix correlation coefficient .

External validation was used in order to determine the predictive ability of the developed models as judged by its application on test set ,figure (3.5) ,(3.8),(3.11) and (3.14) respectively showed that R^2_{pred} = 0.76,0.70 ,0.70 and 0.73 from 2D,3D and (2D +3D) ,it means the quality of judgment is good, from models (1),(5), (6) and (11)

The low values of RMSE indicate that the developed QSAR model is stable for predicting unknown compounds (Edache *et al*.,2015), as mentioned above in tables (3.1) and (3.2)

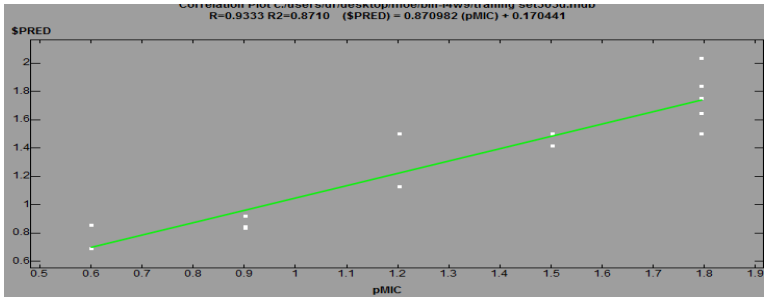


Figure (3.3) The predicted pMIC versus observed pMIC of Training set in (2D) dimensional (Model 1)

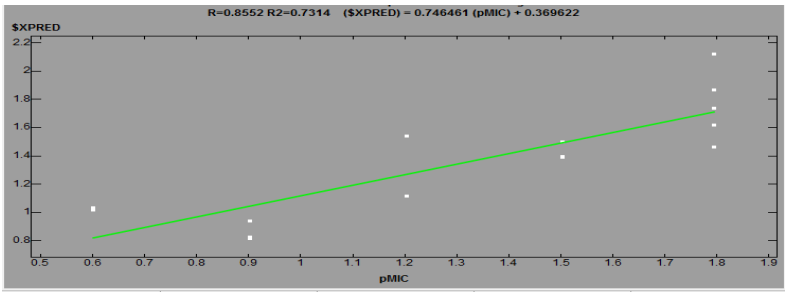


Figure (3.4) The predicted pMIC versus observed pMIC of training set cross validation in (2D)dimensional (Model 1)

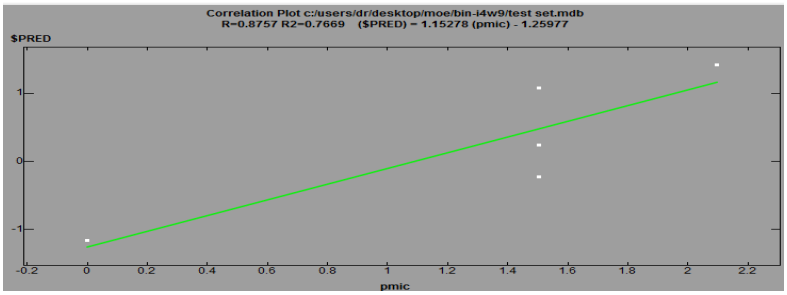


Figure (3.5) The predicted pMIC versus observed pMIC of test set compound in (2D)dimensional (Model 1)

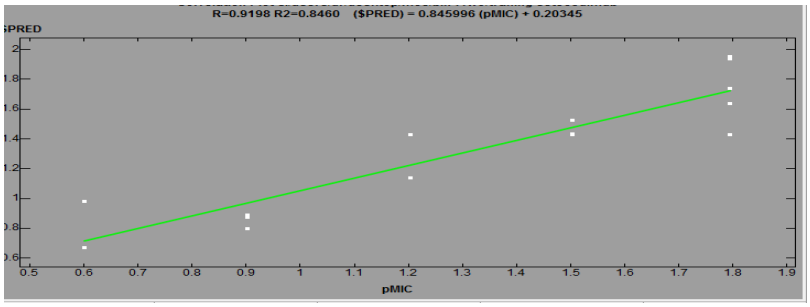


Figure (3.6) The predicted pMIC versus observed pMIC of training set compound in (2D)dimensional (Model 5)

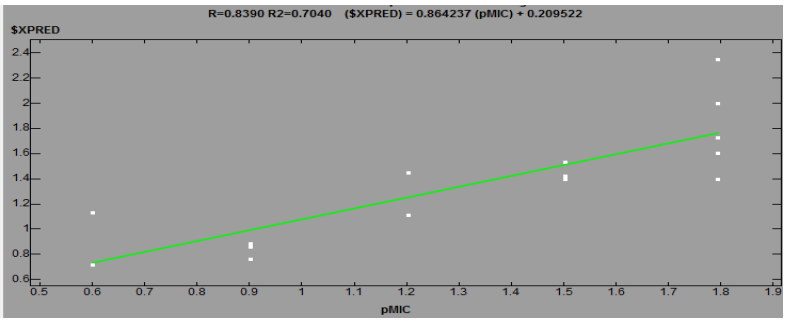


Figure (3.7) The predicted pMIC versus observed pMIC of training set cross validation in (2D)dimensional (Model 5)

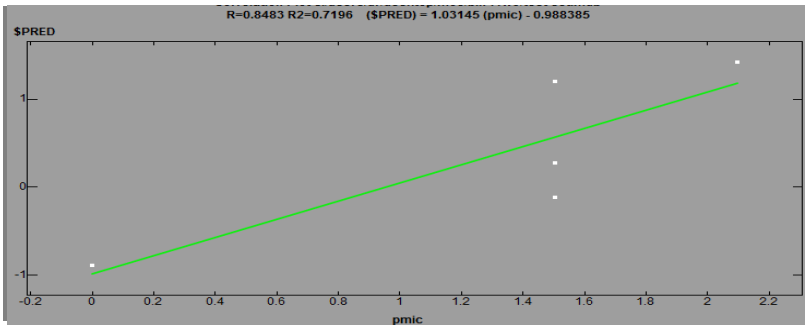


Figure (3.8) The predicted pMIC versus observed pMIC of test Set compound in(2D)dimensional (Model 5)

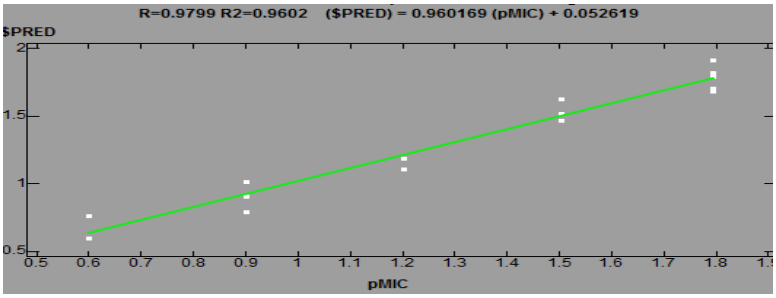


Figure (3.9) The predicted pMIC versus observed pMIC of Training set in (3D)dimensional (Model 6)

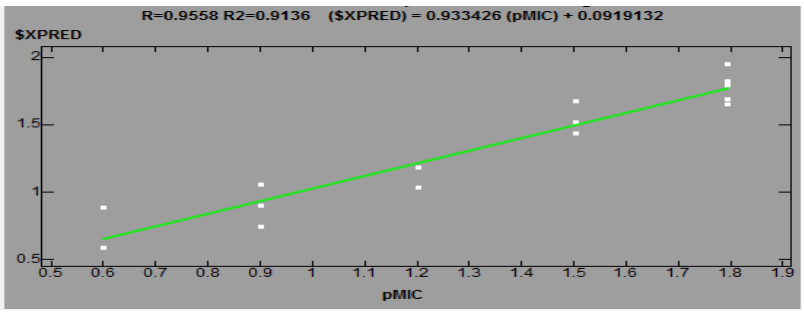


Figure (3.10) The predicted pMIC versus observed pMIC values of training set cross validation in (3D)dimensional (Model 6)

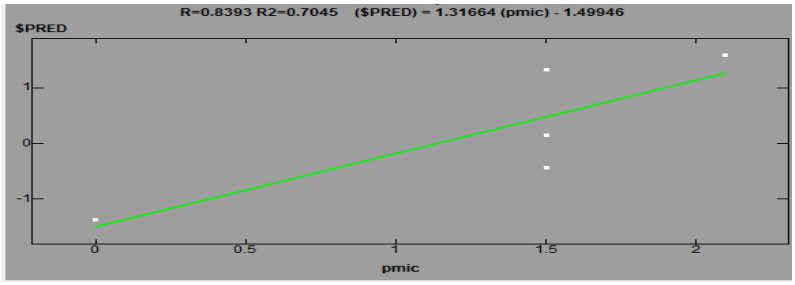
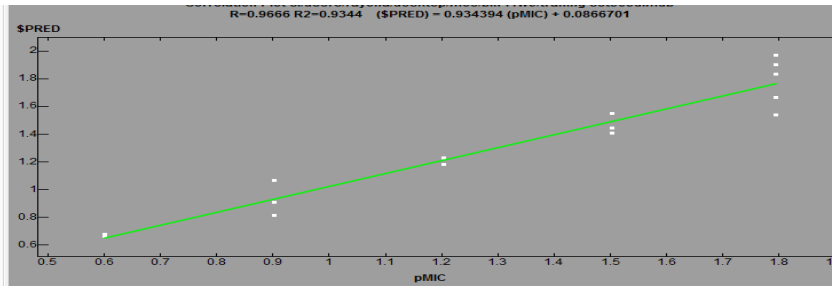


Figure (3.11) The predicted pMIC versus observed pMIC of test Set compound in(3D)dimensional (Model 6)



Figure(3.12) The predicted pMIC versus observed pMIC of Training set in (2D+3D)dimensional (Model 11)

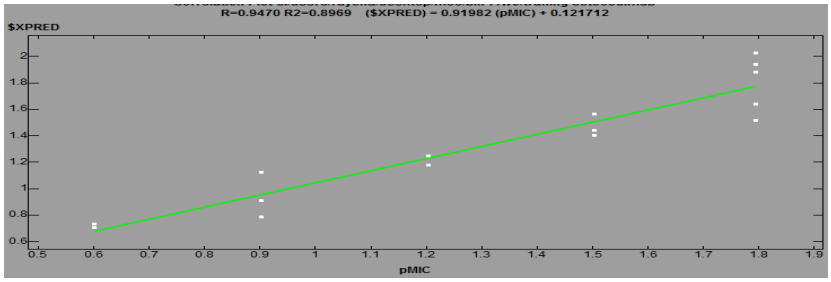


Figure (3.13) The predicted pMIC versus observed pMIC of training set cross validation in (2D+3D)dimensional (Model 11)

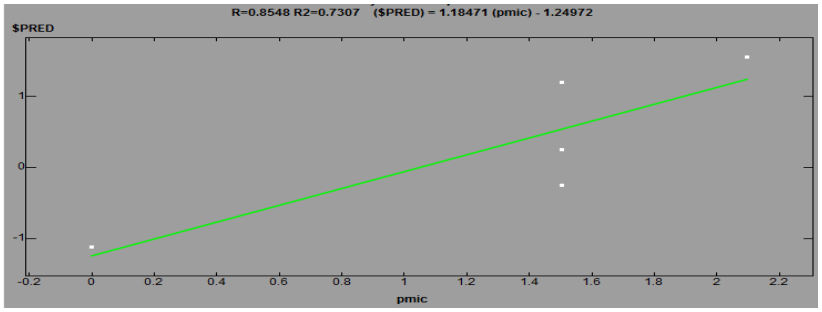


Figure (3.14) The predicted pMIC versus observed pMIC of test Set compound in(2D+3D)dimensional (Model 11)

Multiple linear regressions (MLR) method used as statistical technique in QSAR analysis by using SPSS program . MLR methodis applied for modeling linear Relationship between a dependent variable Y(biological activity)and independent variables X(3 descriptors)to providing the statistical parameters as standard error of estimate (SEE) ,P value andFisher statistic (F) ,that give the model more emphasis . The standard error of estimate (SEE < 0.3) represents an absolute measure of prediction accuracy. Fischer's value (F) or the Fisher ratio, reflects the ratio of the variance explained by the model and the variance due to the error in the regression. High values of the F-test indicate that the model is statistically significant the model (6) record best value of F=88 with R^2 =0.96.The p value is the statistical confidence level for evidence for the null hypothesis, which should not exceed more than 0.05 ($p < 0.05$).(Hajalsiddig *et al*.,2020)
Table (3.3) clarify that obviously ,take equation (6) as example p =0.00 , F=88.390, SEE= 0.100 they are fair enough values as results.

Table (3-3) The statistical parameters for equation1,2,3 from (2D) , 6,7,8 from(3D)and 11 from (2D+3D)

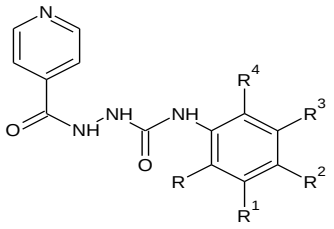
No. Equation	Training Set	Test set	R^2	Q^2	R^2 predicted	RMSE	SEE	F-value	P-value
1	15	5	0.8709	0.7314	0.7045	0.1556	0.1817	24.753	0.00
2	15	5	0.8693	0.7379	0.7669	0.1566	0.1828	24.388	0.00
5	15	5	0.8460	0.7014	0.7196	0.1700	0.1980	20.165	0.00
6	15	5	0.9601	0.9136	0.7027	0.0864	0.1009	88.390	0.00
7	15	5	0.9413	0.8687	0.7533	0.1049	0.1217	59.597	0.00
8	15	5	0.9231	0.8528	0.7438	0.1200	0.1409	43.570	0.00
11	15	5	0.9340	0.8969	0.7307	0.1109	0.1296	52.222	0.00

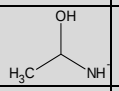
- R^2,R^2 predicted ,RMSE, Q^2, were calculated by using MOE 2009.10 software program.
- SEE,F,P were calculated by using SPSS software program.

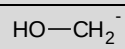
3.1.3 Application of QSAR models in 2D and 3D :

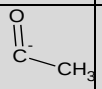
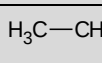
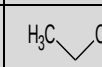
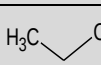
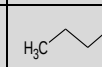
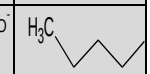
The most important result of (QSAR) models was predicted the biological activity of designed compounds by using QSAR models. The biological activity were calculated in the (2D and 3D) dimensional using MOE program . It is noticeable that the values of calculated biological activity in(2D and 3D) dimensional are closed to each other . The best models (model 1 in 2D ,model 6 in 3D)was applied on the new designed compounds (58 compounds)on table(3-4) . Based on the models was used, The best results of biological activity for compounds (XXIX,XXX,XXXII,XXXIII,XLI,XLII,XLVI,XLVII,LII,LIII,LVII,LVIII) comparison to recorded value by isoniazid (-0.3), (the MIC value of isoniazid converted to pMIC to facilitate comparison) all results were recorded in table (3-4).

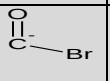
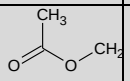
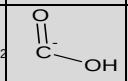
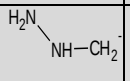
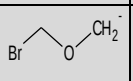
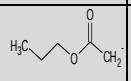
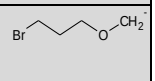
Table(3-4) The new designed N-Substituted 2-Isonicotinoyl hydrazine carboxamides derivatives compounds Predicated biological activity in(2D,3D and 2D+3D) and the descriptors Values (2D,3D dimensionals)

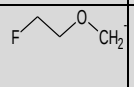
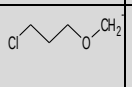
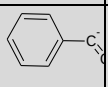
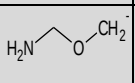
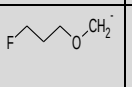
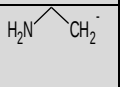


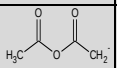
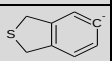
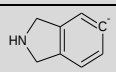
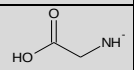
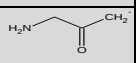
No	1	2	3	4	5	6	7	8
Compound	I	II	III	IV	V	VI	VII	VIII
R1	-H	-H	-Br	$-NH_2$	-H	-H	-OH	-OH
R2	-H	-H	-H	-H	-H	-H	-H	-H
R3	-Cl		-H	-H	-OH	-Br	-OH	-OH
R4	-H	-H	-H	-H	-H	-H	-H	-H
R5	-H	-H	-H	-H	-H	-H	-H	-OH
Predicted pMIC Model 1 (2D)	1.085	1.085	1.714	2.322	2.338	1.714	2.166	1.988
Predicted pMIC Model 5 (2D)	1.8571	1.5971	1.5158	2.2510	2.2568	1.5449	2.2174	2.1821
Predicted PMIC Model 6 (3D)	1.8413	0.8884	1.8342	2.0870	2.2819	1.5870	1.7785	2.1052
PMIC Predicted Model 11 (2D+3D)	1.8280	1.4283	1.3014	2.2986	2.2189	1.6574	2.2046	2.2045
Slog P_VSA9 (2D)	62.701	62.701	72.573	29.375	31.862	72.573	37.105	2.349
Kier2(2D)	10.095	10.095	8.444	8.444	8.444	8.444	8.585	8.740
chi0v(2D)	12.29	12.29	11.84	10.4599	10.32	11.84	10.69	11.06
LogP(o/w) (2D)	1.5140	0.4140	1.7180	0.2480	0.6140	1.720	0.3410	0.0680
KierA3 (2D)	3.8044	4.5747	3.8001	3.1524	3.4764	3.0472	3.4966	3.5572
chi1v (2D)	1.8571	6.8643	6.4780	5.7847	5.7136	6.4720	5.8539	5.9942
E_oop (3D)	0.0385	0.1111	0.0179	0.0766	0.0175	1.1884	0.2712	0.0213
ASA (3D)	515.1	571.5	526.7	498.2461	507.4	530.5	519.5	514.5
npr1 (3D)	1.8571	0.175	0.117	1.0656	0.058	0.114	0.067	0.083

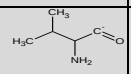
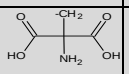
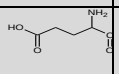
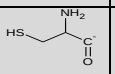
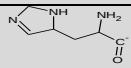
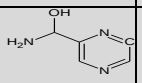
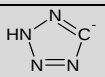
No	9	10	11	12	13	14	15	16
Compound	IX	X	XI	XII	XIII	XIV	XV	XVI
R1	-H	-Cl	-H	-H	-H	-H	-H	-H
R2	-OH	-H	-H	-H	-Br	$-NH_2$	$-NH_2$	-H
R3	-H	-Cl			-H	-H	-H	$-NH_2$
R4	-H	-H	-H	-H	-Cl	$-NH_2$	-H	-H
R5	-H	-H	-H	-H	-H	-H	-H	-H
Predicted pMIC Model 1 (2D)	2.338	1.388	1.894	0.994	1.130	2.132	2.305	2.305
Predicted pMIC Model 5 (2D)	2.249	1.4159	2.0105	1.2925	1.1191	2.2657	2.2680	2.2758
Predicted pMIC Model 6 (3D)	2.2933	1.356	1.9310	0.9087	1.4449	1.5113	2.1256	1.8261
PMIC Predicted Model 11 (2D+3D)	2.2897	1.0643	1.9679	1.1954	1.2993	2.2474	2.3060	2.2503
Slog P_VSA9 (2D)	31.862	102.15	23.862	23.862	108.96	23.862	23.862	23.862
Kier2(2D)	8.444	8.585	9.209	10.78	8.585	8.585	8.444	8.444
chi0v(2D)	10.32	12.22	11.24	12.65	13.11	11.16	10.66	10.66
LogP(o/w) (2D)	0.6510	2.1410	0.4000	0.9300	2.4230	-0.3110	0.2870	0.2500
KierA3 (2D)	3.4764	4.1331	3.7118	4.7075	4.6455	3.4788	3.3590	3.3590
chi1v (2D)	5.7136	1.4159	2.0125	7.1598	6.9496	5.9781	5.7787	6.5564
E_oop (3D)	0.0168	0.374	0.0211	0.0187	0.0146	0.2073	0.0194	0.1267
ASA (3D)	505.5	538.2	530.6	594.6	554.0	523.4	516.1	507.3
npr1 (3D)	0.066	0.074	0.055	0.071	0.107	0.176	0.066	0.187

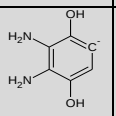
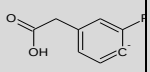
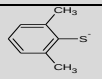
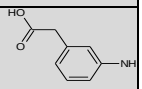
No	17	18	19	20	21	22	23
Compound	XVII	XVIII	XIX	XX	XXI	XXII	XXIII
R1	-H	-H	-H	-H	-H	-H	-H
R2	$-NO_2$	-H	-H	-H	-H	-H	-H
R3	-H						
R4	-Cl	-H	-H	-H	-H	-H	-H
R5	-H	-H	-H	-H	-H	-H	-H
Predicted pMIC Model 1 (2D)	1.272	1.459	1.552	1.105	1.178	0.726	0.269
Predicted pMIC Model 5 (2D)	1.5440	1.7358	1.4602	1.0641	1.5134	1.0803	0.6885
Predicted pMIC Model 6 (3D)	1.1373	1.3091	1.2827	1.1684	1.0993	0.5714	0.6044
PMIC Predicted Model 11 (2D+3D)	1.5210	1.6484	1.4900	1.1413	1.3737	0.8841	0.4545
Slog P_VSA9 (2D)	63.010	57.188	57.188	57.188	57.188	57.188	57.188
Kier2(2D)	9.474	9.333	9.209	9.988	9.988	10.78	11.583
chi0v(2D)	12.40	11.99	11.79	12.50	12.20	12.91	13.61
LogP(o/w) (2D)	1.5600	0.7750	1.6950	2.1370	1.2190	1.8330	2.2750
KierA3 (2D)	4.2946	3.7679	3.7118	4.2126	4.2126	4.7075	5.2514
chi1v (2D)	6.5564	6.4442	6.5507	7.0507	6.6899	7.1899	7.6899
E_oop (3D)	0.0899	0.2056	0.1987	0.0196	0.1939	0.0423	0.0178
ASA (3D)	550.9	546.2	547.2	579.2	553.7	599.9	617.1
npr1 (3D)	0.210	0.120	0.127	0.062	0.162	0.160	0.055

No	24	25	26	27	28	29	30
Compound	XXIV	XXV	XXVI	XXVII	XXVIII	XXIX	XXX
R1	-H	-H	-H	-H	-H	-H	-H
R2	-H	-H	-H	-H	-H	-H	-H
R3							
R4	-H	-H	-H	-H	-H	-H	-H
R5	-H	-H	-H	-H	-H	-H	-H
Predicted pMIC Model 1 (2D)	1.144	0.644	1.801	1.510	0.411	-0.2602	-0.505
Predicted pMIC Model 5 (2D)	1.2547	1.3948	1.9685	1.9529	0.6969	0.5306	0.0359
Predicted pMIC Model 6 (3D)	1.3291	0.6395	1.5022	1.2157	0.3022	-0.0596	-0.5249
PMIC Predicted Model 11 (2D+3D)	1.3494	1.1676	1.8467	1.7612	0.7624	0.2784	-0.0617
Slog P_VSA9 (2D)	69.816	57.188	23.862	23.862	69.816	57.188	69.816
Kier2(2D)	9.333	10.871	9.333	9.9881	10.78	12.45	12.398
chi0v(2D)	12.96	13.11	11.44	11.69	13.87	14.52	15.29
LogP(o/w) (2D)	1.7050	0.9890	0.5990	-0.1770	1.8820	1.9060	2.2140
KierA3 (2D)	4.3224	5.0058	3.7679	4.2126	5.3492	5.7904	6.4932
chi1v (2D)	6.9261	7.0405	6.1678	6.4858	7.8097	8.1012	8.8097
E_oop (3D)	0.0774	0.2001	0.2090	0.0703	0.1770	0.0603	0.0839
ASA (3D)	555.2	588.0	532.1	549.6	607.3	637.7	683.1
npr1 (3D)	0.116	0.133	0.127	0.197	0.159	0.171	0.071

No	31	32	33	34	35	36	37
Compound	XXXI	XXXII	XXXIII	XXXIV	XXXV	XXXVI	XXXVII
R1	-H	-H	-H	-H	-H	-H	-H
R2	-H	-H	-H	-H	-H	-H	-H
R3							
R4	-H	-H	-H	-H	-H	-H	-H
R5	-H	-H	-H	-H	-H	-H	-H
Predicted pMIC Model 1 (2D)	0.536	0.259	0.403	1.726	1.036	0.075	1.414
Predicted pMIC Model 5 (2D)	1.2811	0.4640	0.2444	1.8914	1.6794	0.8883	1.6711
Predicted pMIC Model 6 (3D)	0.2690	-0.0772	-0.0172	1.6667	0.7100	0.2824	0.9033
PMIC Predicted Model 11 (2D+3D)	1.0135	0.2293	0.3335	1.7912	1.4075	0.5007	1.6331
Slog P_VSA9 (2D)	38.822	63.010	23.862	38.822	23.862	38.822	23.862
Kier2(2D)	11.583	12.398	11.253	9.333	10.78	12.398	9.988
chi0v(2D)	12.99	14.46	14.38	11.37	12.48	13.70	12.08
LogP(o/w) (2D)	1.2170	2.0080	2.4370	1.0600	0.2060	1.5690	0.2660
KierA3 (2D)	5.2026	6.1934	4.1066	3.7281	4.7075	5.7356	4.2126
chi1v (2D)	7.1882	8.2227	8.1048	6.1331	6.8292	7.6882	6.7518
E_oop (3D)	0.2168	0.0216	0.2074	0.0926	0.2016	0.0248	0.2216
ASA (3D)	604.4	660.9	634.1	522.6	583.8	638.2	570.7
npr1 (3D)	0.172	0.057	0.110	0.172	0.130	0.052	0.124

No	38	39	40	41	42	43	44
Compound	XXXVIII	XXXIX	XL	XLI	XLII	XLIII	XLIV
R1	-H	-H	-H	-H	-H	-H	-H
R2	-H	-H	-H	-H	-H	-H	-H
R3		NO_2					
R4	-H	-H	-H	-H	-H	-H	-H
R5	-H	-H	-H	-H	-H	-H	-H
Predicted pMIC Model 1 (2D)	0.939	1.824	0.110	0.076	0.285	0.963	0.852
Predicted pMIC Model 5 (2D)	0.4403	1.9741	1.1276	-1.2655	-0.3671	1.6916	1.4630
Predicted pMIC Model 6 (3D)	0.7283	1.6773	0.1792	-0.4028	-0.1650	0.6215	0.7931
PMIC Predicted Model 11 (2D+3D)	0.5985	1.8591	0.7612	-0.2787	0.0138	1.5488	1.3689
Slog P_VSA9 (2D)	23.862	23.862	57.188	90.548	60.741	23.862	23.862
Kier2(2D)	10.364	9.333	11.755	10.346	10.346	10.871	10.871
chi0v(2D)	13.47	11.35	14.02	15.96	15.23	12.65	13.10
LogP(o/w) (2D)	2.8820	0.8570	0.9280	3.4630	2.5180	0.0740	-0.3070
KierA3 (2D)	3.8005	3.7408	5.5416	3.9964	3.7377	4.8701	4.9714
chi1v (2D)	7.6507	6.0788	7.5136	9.9172	8.8923	6.8749	7.3265
E_oop (3D)	0.0205	0.0528	0.0291	0.1870	0.0242	0.2763	0.0172
ASA (3D)	610.6	525.5	644.5	662.4	669.4	578.3	603.7
npr1 (3D)	0.045	0.169	0.052	0.100	0.039	0.164	0.062

No	45	46	47	48	49	50	51
Compound	XLV	XLVI	XLVII	XLVIII	XLIX	L	LI
R1	-H	-H	-H	-H	-H	-H	-H
R2	-H	-H	-H	-H	-H	-H	-H
R3							
R4	-H	-H	-H	-H	-H	-H	-H
R5	-H	-H	-H	-H	-H	-H	-H
Predicted pMIC Model 1 (2D)	0.300	-0.042	-0.042	0.543	-0.173	0.261	1.454
Predicted pMIC Model 5 (2D)	0.6155	0.8922	0.8455	0.9000	0.3605	0.9083	1.4922
Predicted pMIC Model 6 (3D)	0.1794	0.2259	0.3668	0.0862	-1.0397	0.0554	1.4527
PMIC Predicted Model 11 (2D+3D)	0.6721	0.8171	0.6946	1.0995	0.5725	1.0346	1.6038
Slog P_VSA9 (2D)	23.862	90.514	90.514	23.860	23.860	23.862	23.862
Kier2(2D)	11.407	11.111	11.111	10.982	12.142	11.407	9.629
chi0v(2D)	14.58	14.72	14.72	14.20	15.46	14.74	12.43
LogP(o/w) (2D)	1.167	-0.6580	-0.1110	0.3050	-0.4100	-0.7065	0.4370
KierA3 (2D)	4.7632	5.7543	5.7680	4.8562	5.0845	4.4904	3.0841
chi1v (2D)	8.0542	0.8922	8.2557	7.9231	8.9326	8.1612	1.4922
E_oop (3D)	0.0205	0.0528	0.5533	0.0291	0.3037	0.0385	0.0166
ASA (3D)	612.1	638.5	530.3	647.5	677.7	652.3	563.9
npr1 (3D)	0.173	0.073	0.424	0.070	0.203	0.049	0.044

No	52	53	54	55	56	57	58
Compound	LII	LIII	LIV	LV	LVI	LVII	LVIII
R1	-H	-H	-CN	-CN	-F	-H	-H
R2	-H	-H	-H	-H	-H	-Br	-Cl
R3			-NH	-NH	-F		
R4	-H	-H	-H	-H	-H	-H	-Cl
R5	-H	-H	-SH	-H	-F	-H	-H
Predicted pMIC Model 1 (2D)	0.175	-0.394	1.431	1.795	1.838	-1.597	-1.683
Predicted pMIC Model 5 (2D)	0.3727	-0.2424	1.5724	2.062	1.9519	-0,9169	-1.1310
Predicted Pmic Model 6 (3D)	-0.2759	-0.7012	0.7186	1.0860	2.2373	-1.1154	-1.5826
PMIC Predicted Model 11 (2D+3D)	0.1886	-0.1957	1.1747	1.7504	2.1111	-0.3365	-0.2366
Slog P_VSA9 (2D)	39.862	38.822	23.862	29.375	68.742	168.687	104.916
Kier2(2D)	11.039	12.296	9.474	9.333	8.740	11.571	13.185
chi0v(2D)	15.21	15.76	12.74	11.32	11.01	18.43	18.07
LogP(o/w) (2D)	9510.0	2.8350	0.4690	-0.0550	1.451	3.760	2.979
KierA3 (2D)	3.9272	4.9939	3.6969	3.3415	3.4472	4.8536	6.1892
chi1v (2D)	8.3419	8.8019	6.7565	6.1690	5.8903	9.4588	10.1872
E_oop (3D)	0.1187	0.1056	0.5082	0.0122	0.0152	0.0905	0.5401
ASA (3D)	667.0	685.3	572.7	553.0	506.4	698.2	713.7
npr1 (3D)	0.054	0.116	0.123	0.090	0.083	0.206	0.094

3.2 Docking study:

The docking process carried out by MOE program . The outcome docking presented (2D) forms in the appendix for data set from figure (4.1) to (4.21),isoniazid (4.22) and designed compounds (4.23) to (4.80) .

Molecular docking was used to predict the binding of ligands (data set ,isoniazid and the new designed derivatives)to the target 4TRO protein . Characteristics of the bindings were tabled in table (3-5) for the data set and isoniazid.Designed compounds were explained in table (3-6). From the result of docking ,the best compound have more interaction (VIII) from

designed compounds ,the best compounds have the lowest energy less than -22 and at least one interaction ,compounds (3f ,3c)from dat sata ,from the new designed derivatives are (XII, XXIII, XXIX, XXX, XXXIII, XL, XLI, XLII, XLVII, LII, LIII, LVII) .

The binding affinity for the data set ranging from (-14.8731 to -24.0150) kcal/mol while the binding affinity for the designed compounds rang from (-15. 7981to -28.1695) kcal/mol.

The docking score correlated to the potency of inhibitors ,the lowest docking score was shown most potent inhibitor ,while the high docking score was shown least potent inhibitor. The lowest docking score was performed by LIII = -28.1695 kcal/mol with one interaction Gly96

and the highest docking score was tabled by compound (XXVII) = -15.7981 kcal/mol with two interaction Val65, Phe41 .The isoniazid binding affinity values = -10.5989 kcal/mol

The new compounds were docked into active site of 4TRO were revealed the most residues (Phe41, Val65, Gly96, Gly14, Ser94) in the protein ligand interactions .

The isoniazid was showed three interactions hydrogen bond with the receptor (4TRO) .

Table (3-5): The Interaction between the data set with (4TRO) protein active sides receptor:

NO Compound	S Bond energy (kcal/mol)	Hydrogen bond length(Å)	Interaction of amino acid	Type of inter action
3a	-18.3771	2.19 -	Val65 Phe41	H-bond donor π -π bond
3b	-17.8684	-	-	No interaction
3c	-23.0513	2.12 2.09	Gly96 Ser94	H-bond acceptor H-bond acceptor
3d	-20.2616	-	Phe41	π -π bond
3e	-20.7632	-	Phe41	π -π bond
3f	-24.0150	2.13 -	Val65 Phe41	H-bond donor π -π bond
3g	-21.3515	-	Phe41	π -π bond
3h	-21.6033	2.18	Lys118 Lys118	H-bond acceptor π –cation bond
3i	-19.1373	2.23 -	Val65 Phe41	H-bond donor π -π bond
3j	-18. 4892	2.20 -	Val65 Phe41	H-bond donor π -π bond

NO Compound	S Bond energy (kcal/mol)	Hydrogen bond length(Å)	Interaction of amino acid	Type of inter action
3k	-19.3499	2.17 -	Val65 Phe41	H-bond donor π -π bond
3l	-18.3147	-	Phe41	π-π bond
3m	-20.4081	2.13 -	Val65 Phe41	H-bond donor π -π bond
3n	-18.9593	2.18 -	Val65 Phe41	H-bond donor π -π bond
3o	-19.0347	-	Phe41	π-π bond
3p	-17.4867	-	-	No interaction
3q	-14.8731	-	-	No interaction
3r	-20.5348	2.10 -	Val65 Phe41	H-bond donor π -π bond
3s	-18.1753	-	-	No interaction
3t	-17.9677	2.06 -	Val65 Phe41	H-bond donor π -π bond
3u	-18.5850	-	-	No interaction
INH	-10.5989	1.77 2.65 2.04	Ser94 Ser94 Gly14	H-bond acceptor H-bond acceptor H-bond acceptor

Table (3-6): The Interaction between the designing set with 4TRO protein active sides receptor

No	Compound	S Bond Energy (kcal/mol)	Hydrogen bond length(Å)	Interaction amino acid	Type of inter action
1	**I**	-18.8351	-	-	No interaction
2	**II**	-22.8545	2.07	Gly14	H-bond acceptor
3	**III**	-19.8033	2.15 -	Val65 Phe41	H-bond donor π -π bond
4	**IV**	-17.9148	2.93 2.26 -	Ser94 Val65 Phe41	H-bond acceptor H-bond donor π -π bond
5	**V**	-19.9600	-	Phe41	π -π bond
6	**VI**	-19.7667	2.02	Gly14	H-bond acceptor
7	**VII**	-19.8196	1.97	Gly96	H-bond acceptor
8	**VIII**	-20.1462	2.78 1.88 2.05 2.15 -	Gly14 Ser94 Gly96 IIe21 Phe41	H-bond acceptor H-bond acceptor H-bond acceptor H-bond donor π -π bond

No	Compound	S Bond Energy (kcal/mol)	Hydrogen bond length(Å)	Interaction amino acid	Type of inter action
9	**IX**	-21.9676	2.13 -	Val56 Phe41	H-bond donor π -π bond
10	**X**	-18.7306	-	-	No interaction
11	**XI**	-21.6312	-	-	No interaction
12	**XII**	-21.3501	2.17 2.06	Gly96 Gly192	H-bond acceptor H-bond acceptor
13	**XIII**	-18.3127	2.05 -	Val65 Phe41	H-bond donor π -π bond
14	**XIV**	-17.0935	2.14 -	Lys118 Phe97	H-bond acceptor π -π bond
15	**XV**	-20.6654	2.14 -	Val65 Phe41	H-bond donor π -π bond
16	**XVI**	-17.6141	2.07 -	Val65 Phe41	H-bond donor π -π bond
17	**XVII**	-20.6676	2.17 -	Val65 Phe41	H-bond donor π -π bond
18	**XVIII**	-22.2965	-	-	No interaction
19	**XIX**	-18.7245	-	-	No interaction
20	**XX**	-19.6942	2.03	Gly96	H-bond acceptor
21	**XXI**	-19.6494	-	Phe41	π -π bond
22	**XXII**	-20.1904	-	-	No interaction
23	**XXIII**	-23.4673	2.22	Ser94	H-bond acceptor
24	**XXIV**	-20.9059	-	-	No interaction
25	**XXV**	-20.3092	2.03	Gly14 Phe41	H-bond acceptor
26	**XXVI**	-18.5326	-	Phe41	π -π bond
27	**XXVII**	-15.7981	2.21 -	Val65 Phe41	H-bond donor π -π bond
28	**XXVIII**	-19.1809	-	-	No interaction
29	**XXIX**	-23.4358	2.10 2.05	IIe194 Ser94	H-bond donor H-bond acceptor
30	**XXX**	-25.0724	2.02 -	Gly14 Phe41	H-bond acceptor π -π bond
31	**XXXI**	-20.1683	2.02 -	Gly14 Phe41	H-bond acceptor π -π bond
32	**XXXII**	-22.4081	2.19	Gly14	H-bond acceptor
33	**XXXIII**	-24.2295	2.10 -	Val65 Phe41	H-bond donor π -π bond
34	**XXXIV**	-19.1542	2.09 2.75 -	Thr196 Asp64 Phe41	H-bond acceptor H-bond donor π -π bond

No	Compound	S Bond Energy (kcal/mol)	Hydrogen bond length(Å)	Interaction amino acid	Type of inter action
35	**XXXV**	-21.4784	2.12 -	Lys165 Phe149	H-bond acceptor π -π bond
36	**XXXVI**	-23.7257	1.97 -	Gly96 Phe41	H-bond acceptor π -π bond
37	**XXXVII**	-19.7007	2.13 2.19 2.14	Gly96 Gly14 Ser94	H-bond acceptor H-bond acceptor H-bond acceptor
38	**XXXVIII**	-21.4714	-	Phe41	π -π bond
39	**XXXIX**	-19.0195	2.11 -	Thr196 Phe41	H-bond acceptor π -π bond
40	**XL**	-23.0472	2.12 2.14 -	Gly96 Val65 Phe41	H-bond acceptor H-bond donor π -π bond
41	**XLI**	-26.0837	2.17 -	Val65 Phe41	H-bond donor π -π bond
42	**XLII**	-26.7429	2.09 -	Val65 Phe41	H-bond donor π -π bond
43	**XLIII**	-22.4260	2.00 -	Ser94 Phe41	H-bond acceptor π -π bond
44	**XLIV**	-21.8662	-	Phe41	π -π bond
45	**XLV**	-22.6569	3.22 -	Lys165 Phe41	H-bond acceptor π -π bond
46	**XLVI**	-22.2708	2.10	Ser94	H-bond acceptor
47	**XLVII**	-24.1621	2.11 1.88 -	Lys118 Asp42 Phe97	H-bond acceptor H-bo donor π -π bond
48	**XLVIII**	-22.4548	2.12	Gly96	H-bond acceptor
49	**XLIX**	-21.3022	-	Phe41	π -π bond
50	**L**	-25.1195	-	-	No interaction
51	**LI**	-19.5723	2.08 2.01 2.24	Gly96 Ser94 Gly14	H-bond acceptor H-bond acceptor H-bond acceptor
52	**LII**	-25.5646	2.29 2.04 2.27 1.93	Ser94 Ser94 IIe194 IIe194	H-bond acceptor H-bond acceptor H-bond donor H-bond donor
53	**LIII**	-28.1695	2.25	Gly96	H-bond acceptor

No	Compound	S Bond Energy (kcal/mol)	Hydrogen bond length(Å)	Interaction amino acid	Type of inter action
54	**LIV**	-20.4511	2.19 -	Gly14 Phe41	H-bond acceptor π -π bond
55	**LV**	-17.1932	2.08	Gly14	H-bond acceptor
56	**LVI**	-19.2803	-	Phe41	π -π bond
57	**LVII**	-25.9896	2.27	Ser94	H-bond acceptor
58	**LVIII**	-26.3585	-	-	No interaction

It was found that most of the interactions are between the receptor and nitrogen of pyridine ,ring of pyridine which existing in data set with Val65, Phe41of protein receptors to give hydrogen bond and π – π bond respectively . All data set showed lower energy than that tabled by isoniazid .

compounds (3d,3h) were selected according to the low pMIC ,low energy in their interaction with receptor.

The compound (3d) exhibited binding free energy -20.2616 Kcal/mol ,and (3d) forming one π interaction.

The π – π interactions between pyridine ring in compound (3d) with phe41 phenyl ring. Figure(3.15)

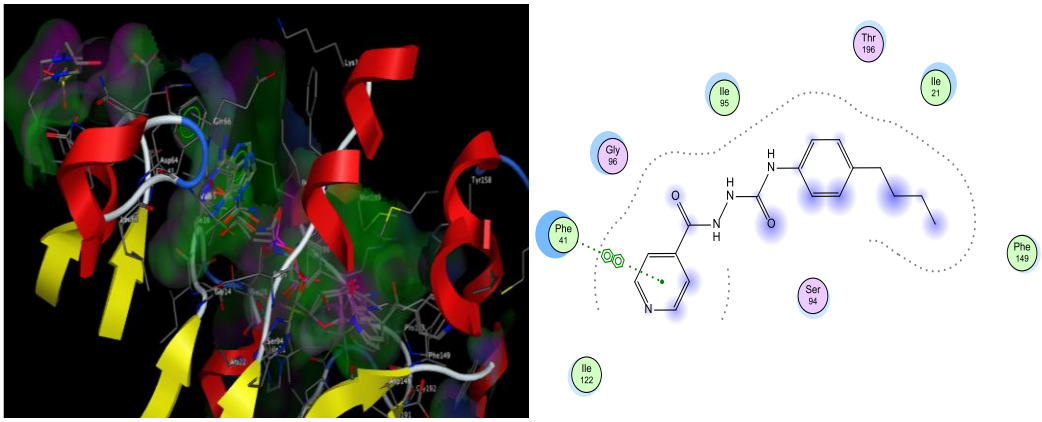


Figure (3.15) The interaction between the compound 3d and receptor in(2D -3D) dimensional

The distinguished of the two compounds (3d ,3h) , a hydrocarbon chain present in site (4) as substitutes

compound (3h)has a binding affinity of -20.2616 Kcal/mol ,and (3h) formed one H- bond and one π –cation bond interaction.

The H-bond between O oxygen of carbonyl in (3h) with N of amino group in Lys118 the bond length was 2.18 Å .

The π – cation interactions between pyridine ring in compound (3h) with Lys118 amino acid . Figure(3.16)

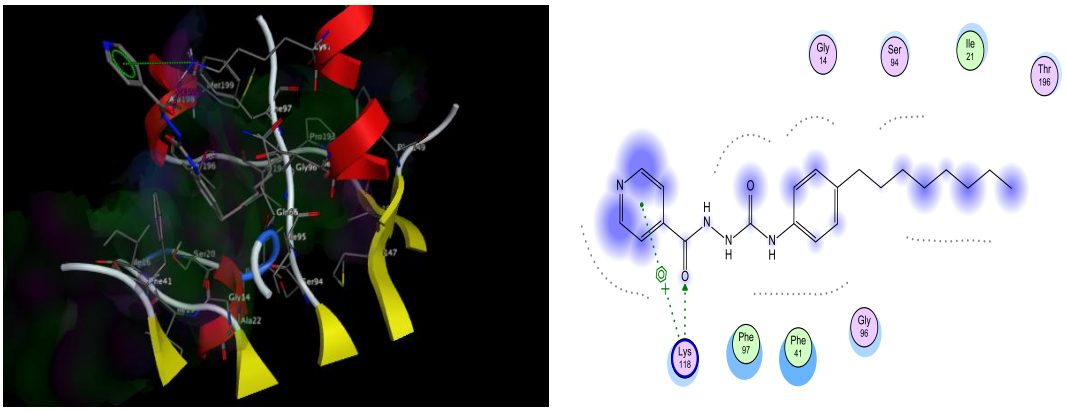
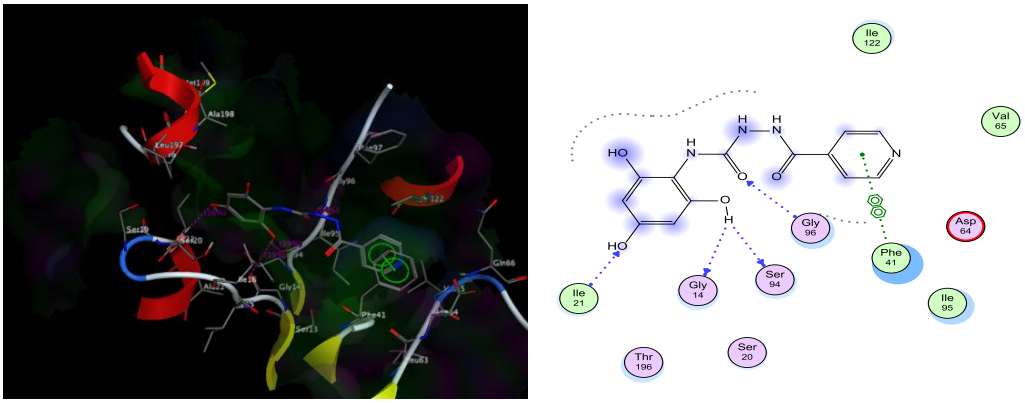


Figure (3.16) The interaction between the compound(3h)and receptor in (2D -3D) dimensional

Most of the designed compounds were showed more interactions with the amino acid residues (Phe41, Val65 ,Gly96, Gly14, Ser94) , and less interaction with (Asp64, Phe97, lle194, Lys118 ,Lys165, Thr196, lle21 , lle95).There are similarities between the compounds (XXX,XXXI,XXXII) in the interaction and energy, also compounds (XLI,XLII).

The compound (VIII) exhibited binding affinity of -20.1462 Kcal/mol and formed 4 H-bond interactions and one π Interaction.

H-bond between OH group in VIII compound and amino group in (lle21) , the bond length was 2.15 Å, H bond between H of hydroxyl group in VIII and O oxygen of carbonyl in (ser94) the bond length was 1.88 Å ,H bond between H hydrogen of hydroxyl in VIII compound and O oxygen of carbonyl in Gly14 the bond length 2.79 Å, H bond between O oxygen of carbonyl in compound VIII and amino group in Gly96 the length bond 2.05 Å ,finally the π – π interactions between pyridine ring and (Phy41) in phenyl ring : in Figure 3.17



Figure(3.17) The interaction between the compound VIII and receptor 4TRO in(2D -3D) dimensional

compound (VII)exhibited binding free energy -25.0724 Kcal/mol and forming H-bond interaction and one π Interaction.

The amino group in XXX compound interacted with O oxygen of carbonyl in Gly14*via*H bond, the bond length was 2.02 Å , also formed π – π interactions between pyridine ring in compound VII with phe41 phenyl ring . Figure (3.18)

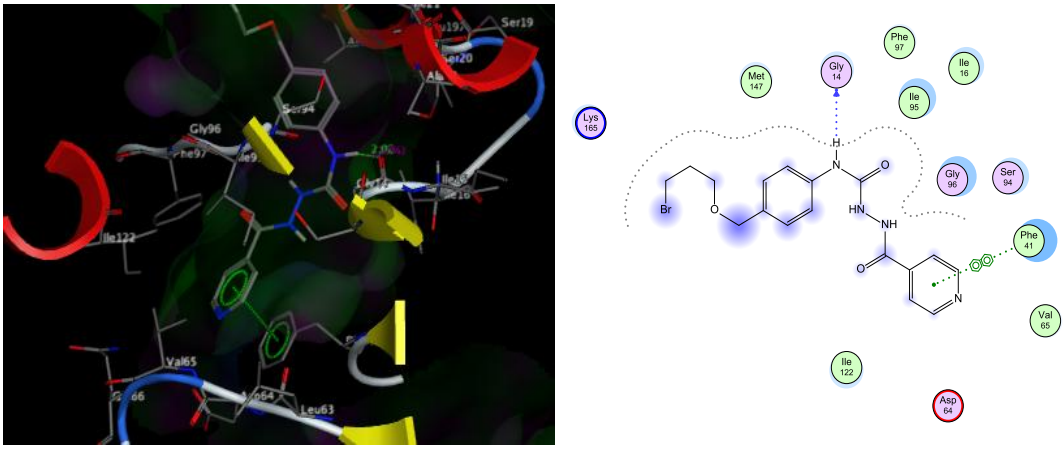


Figure (3.18) The interaction between the compound XXX and receptor 4TRO in(2D -3D) dimensional

Compounds (XXX,XXXI) are similar in contained halogen group ,and showed the same interacted with the receptor.

The binding free energy was tabled for the compound (XXXIII) -24.2295 Kcal/mol and forming H-bond interactions and one π Interaction.
Two H-bond between N nitrogen of pyridine in XXXIII compound and amino group of Val65 the bond length was 2.10 Å , the π – π interaction between pyridine ring in compound XXXIII with phe41 phenyl ring . Figure(3.19)

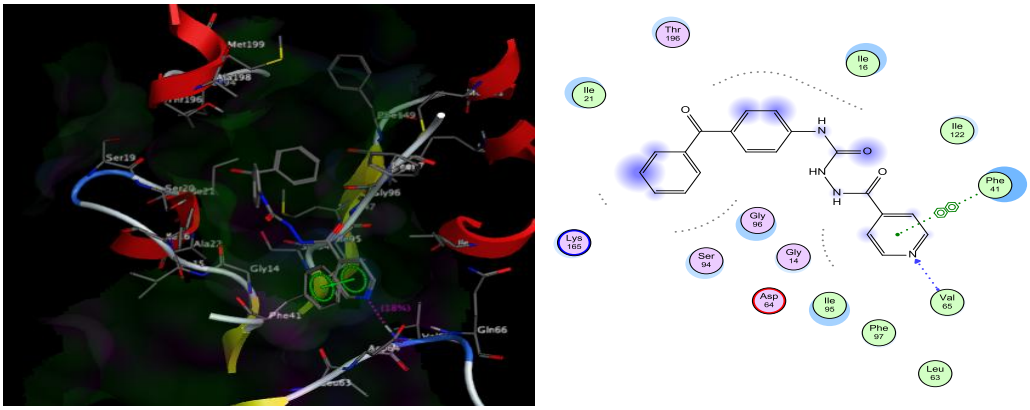


Figure (3.19) The interaction between the compound XXXIII and receptor 4TRO in(2D -3D) dimensional

The compound (XLII)formed H-bond interactions and one π Interaction with binding free energy -26.7429 Kcal/mol .

The H bond between N of pyridine ring group in XLII with backbone of amino group Val65 *via*hydrogen bond , the bond length was 2.09 Å and the π – π interactions between pyridine ring in compound XLII with phe41 phenyl group. Figure(3.20)

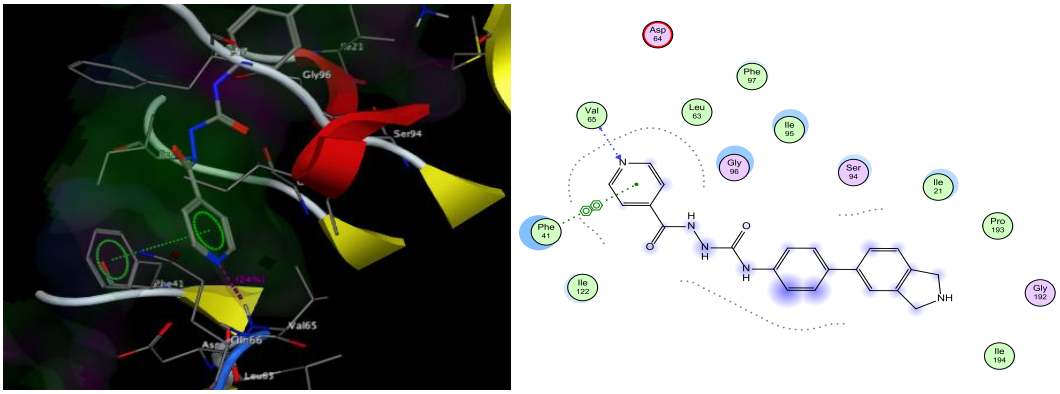


Figure (3.20) The interaction between the compound XLII and receptor 4TRO in(2D -3D) dimensional

Compounds (XLI, XLII) involve in consist heterocyclic (fused ring) at the site (4) of N-substituted -2-isonicotinoyl hydrazine carboxamide derivatives were showed interactions with (Val65,Phe41) residue.

The binding free energy tabled for the compound (XLVII) -24.1621 Kcal/mol and formed 2H-bond interactions and one π Interaction.
Two H-bond between H of hydroxyl group in XLVII compound and O oxygen of carboxylic in Asp42 , the bond length was 1.88 Å, H bond between O oxygen of carbonyl in XLVII and amino group in Lys118 the bond length was 2.11 Å and the π – π interactions between phenyl ring in compound XLVII with phe97 phenyl ring . Figure (3.21)

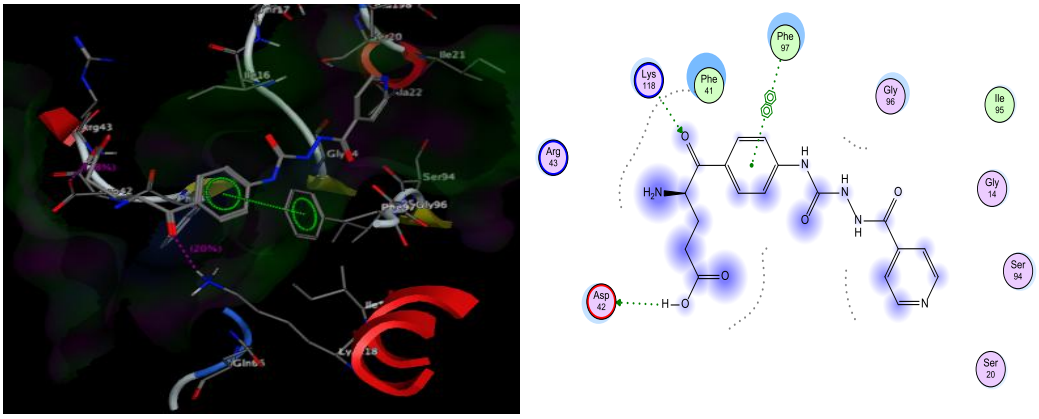


Figure (3.21) The interaction between the compound XLVII and receptor in(2D -3D) dimensional

Compound (XLVII) distinguished by the presence of two carbonyl group,one amino group and one hydroxyl group , carponyl and hydroxyl group that formed bonds with the receptor.

Compound (LII) exhibited with binding affinity -21.1446 Kcal/mol ,and showed the occurrence 2 H-bond interaction and one π Interaction.
The H-bond between amino group in LII compound and O oxygen of carbonyl in Ser94 , the bond length was 2.29 Å , The amino group in LII compound mediated hydrogen bond with O oxygen of carbonyl in Ser94 , the bond length was 2.04 Å . H bond formation occurred between hydroxyl

group in LII and amino group in IIe194 the bond length was 1.93 Å ,also formed H bond between hydroxyl group in LII and O oxygen of carbonyl in lle194 residue the bond length was 2.27 Å . Compound(LII) was showed three H bond interactions and was not showed any pi bond.figure (3.22)

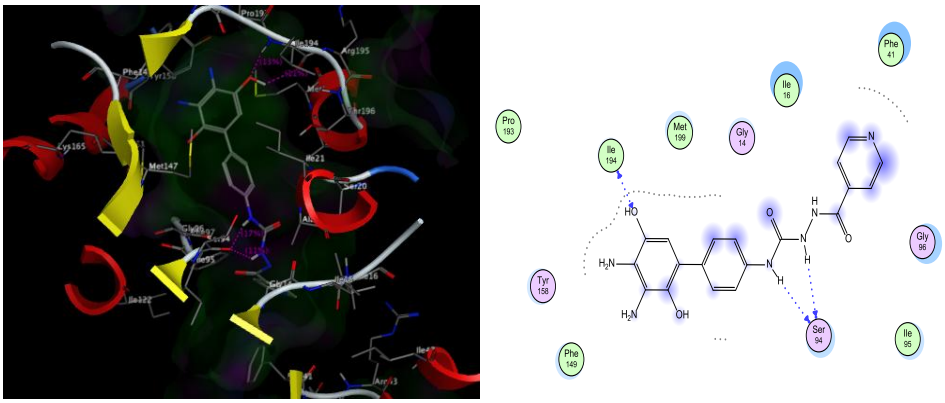


Figure (3.22) The interaction between the compound LII and receptor in(2D -3D) dimensional

Compound (LIII) formed one H-bond interaction with binding free energy -28.1695 Kcal/mol .

H hydrogen of amino group in LIII compound via H-bond interacted with O oxygen of carbonyl in Gly96, the bond length was 2.25 Å. Figure(3.23)

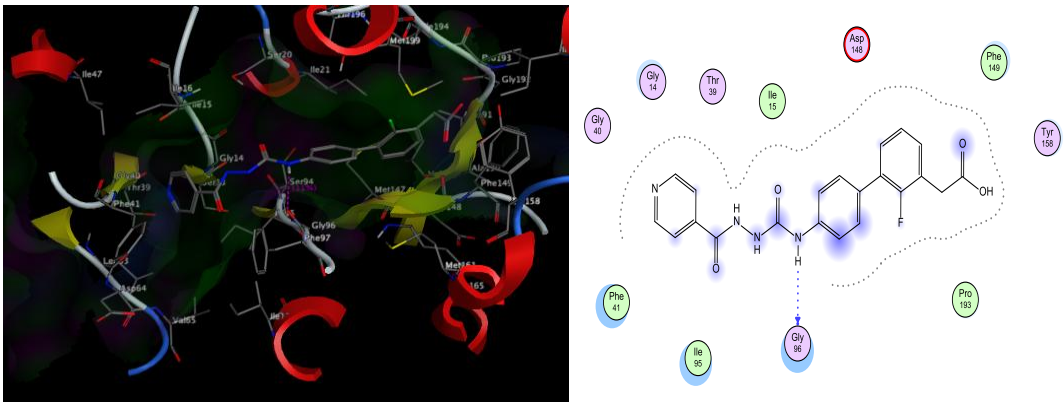


Figure (3.23) The interaction between the compound LIII and receptor in(2D -3D) dimensional

The binding free energy was revealed for the compound (LVII) -25.9896 Kcal/mol and which formed one H-bond interaction.

H-bond between amino group in LVII compound with O oxygen of carbonyl Ser94 residue , the bond length was 2.27 Å . Figure(3.24)

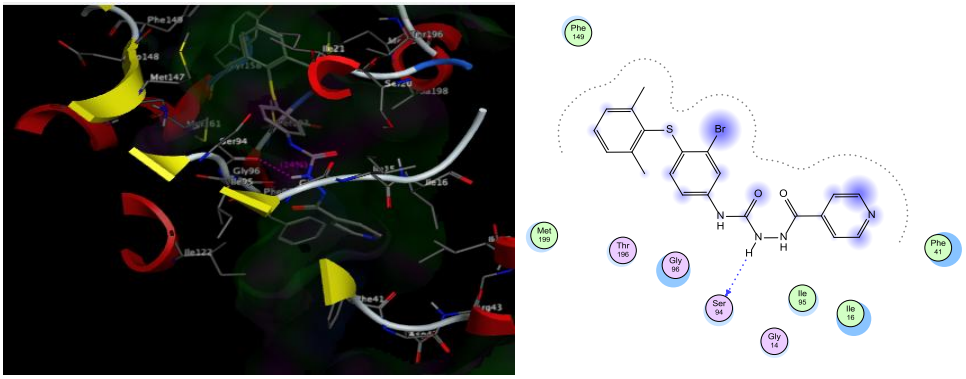


Figure (3.24) The interaction between the compound LVII and receptor in(2D -3D) dimensional

Chapter Four

Conclusion and recommendation

4. Conclusion and recommendation

- In the search for effective and selective antituberculosis agents, series of N-Substituted- 2-Isonicotinoylhydrazinecarboxamides derivatives were studied and modeled. The 3D-QSAR models presented here, are powerful enough to suggest improvement in N-substituted -2-isonicotinoyl hydrazine carboxamide derivatives
- 3D - QSAR approaches was superior on 2D-QSAR in modeling according to this research. By combining 3D-QSAR and molecular docking studies are good approach of design N-substituted -2-isonicotinoyl hydrazine carboxamide derivatives anti agents .
- The three modeled that applied in the designed new N-substituted -2-isonicotinoyl hydrazine carboxamide derivatives compounds, the compounds (XXIX, XXX, XXXII,XLI, XLIX) that revealed low pMIC when model (5) was used (that contain log p(o/w) reverse model 6 and model 1 in 3D and 2D respectively which showed high pMIC. Furthermore (LVII,LVIII) were revealed low pMIC when models (1,6) are used ,the rest of designed compounds showed similar values with used models 1,6 and 5 (which contain log p(o/w)).
- All the new designed N-substituted -2-isonicotinoyl hydrazine carboxamides derivatives sketched by ACD lab software , In addition were characterized and evaluated by MOE 2009.10 and SPSS software.
- The new designed compounds were indicated molecular docking or receptor- ligand interference , compounds (XII, XXIII, XXIX, XXX, XXXIII, XL, XLI, XLII, XLVII, LII, LIII, LVII) were showed interaction ,good biological activity and low energy. Some of this compound as (XXIX,XXXIII ,XLVII, LIII) contained carbonyl group ,and compound (XI, XLVII , LII,LIII) consiste of hydroxyl group .In addition to compounds(XLVII, LII,XLII) involve amino group ,also some of this compound contain of halogens or sulfur.
- Compound(VIII) was showed more interactions , low energy ; however, high pMIC .
- Applying modeling techniques in addition the tools that use in modeling as software programs, could be useful for chemists. considered the study computer aided methods as potential and complex tools that may serve as valuable partnership with wet lab experiments and may provide a rational aid to minimize the cost and time of research.

- Recommended reviewing and study the (VIII ,XII, XXIII, XXIX, XXX, XXXIII, XL, XLI, XLII, XLVII, LII, LIII, LVII) compounds in a broad and comprehensive manner according to their biological activity and docking study .
- Application of *silico* ADMET prediction for characteristic compounds, In the search for design and discovery of anti-tuberculosis drugs, ADMET (absorption, distribution, metabolism, excretion and toxicity) is a prerequisite. On account, the properties of the molecule play an important role in the initial clinical stage. It is necessary to predict the ADMET properties of compounds designed to ensure drug adaptability to the human body in advance, which include absorption in the human intestine, the blood-brain barrier and penetration into the central nervous system. Metabolism refers to the chemical and biological transformation of drugs in the body, the total chain, the toxicity level of drugs and molecules The admetSAR online web servers and ADME Tlab are used to predict the ADMET properties of the newly designed compounds.
- Some of QSAR models can be used in the two dimensional or the three dimensional to design more N-substituted -2-isonicotinoyl hydrazine carboxamide derivatives to search and design anti-tuberculosis drugs according to the results later.

References:

- Alvarez-Builla, J.; Vaquero, J.J. and Barluenga, J. eds.(2011). *Modern heterocyclic chemistry, volume* 2. John Wiley & Sons.

-Andersson, D. (2010). *Multivariate design of molecular docking experiments: An investigation of protein-ligand interactions,* Ph.D.Thesis , Umea University ,Sweden .

-Bernal, F. A. and Schmidt, T. J. (2019). A comprehensive QSAR study on antileishmanial and antitrypanosomal cinnamate ester analogues. *Molecules*, *24*(23), 4358

-Chtita, S.; Belhassan, A.; Bakhouch, M.; Taourati, A.I.; Aouidate, A.; Belaidi, S.; Moutaabbid, M.; Belaaouad, S.; Bouachrine, M. and Lakhlifi, T. (2021). QSAR study of unsymmetrical aromatic disulfides as potent avian SARS-CoV main protease inhibitors using quantum chemical descriptors and statistical methods. *Chemometrics and Intelligent Laboratory Systems*, *210*, (3)104266.

-Cotuá, V.J.; LLinás, S.H. and Cotes, O.S. (2021). Virtual Screening Based on QSAR and Molecular Docking of Possible Inhibitors Targeting Chagas CYP51. *Journal of Chemistry*, 2021,(6)1-15.

-Dua, R.; Shrivastava, S.; Sonwane, S.K. and Srivastava, S.K. (2011). Pharmacological significance of synthetic heterocycles scaffold: a review. *Advances in Biological Research*, *5*(3), 120-144.

-Dze, K. C. and Samad, F. (2020). Heterocycles, their Synthesis and Industrial Applications: A Review.*International Journal for Research in Applied Science & Engineering Technology*,*8*,(X),36-56

-Edache, E.I.; Uttu, A.J.; Oluwaseye, A.; Samuel, H. and Abduljelil, A. A. (2015). Semi-empirical based QSAR study of indole***β***-Diketo acid, Diketo acid and Carboxamide Derivativesas potent HIV-1 agent Using Quantum Chemical descriptors.*IOSR Journal of Applied Chemistry*,***70*****(11),4397-4403**

-Esposito, E. X.; Hopfinger, A. J and Madura, J. D. (2004). Methods for applying the quantitative structure-activity relationship paradigm.In *chemoinformatics*: *Methods in molecular biology* ,131–214,Springer.

-Fernandes, G.F.D.S.; Salgado, H.R.N. and Santos, J.L.D. (2017). Isoniazid: a review of characteristics, properties and analytical methods. *Critical reviews in analytical chemistry*,*47*(4), 298-308.

-Ferreira, L.G.; Dos Santos, R.N.; Oliva, G. and Andricopulo, A.D. (2015). Molecular docking and structure-based drug design strategies. *Molecules*, *20*(7),13384-13421.

-Gajewicz-Skretna, A.; Kar, S.; Piotrowska, M. and Leszczynski, J. (2021). The kernel-weighted local polynomial regression (KwLPR) approach: an efficient, novel tool for development of QSAR/QSAAR toxicity extrapolation models. *Journal of cheminformatics*, *13*(1), 1-20.

-Genheden, S.; Reymer, A.; Saenz-Méndez, P. and Eriksson, L.A. (2017). Computational chemistry and molecular modelling basics. *Computational Tools for Chemical Biology*,Royal ,1-38.

-Gramatica, P. (2013). On the development and validation of QSAR models. In *Computational toxicology*, 930 ,499-526.

-Hajalsiddig, T.T.; Osman, A.B.M. and Saeed, A.E.M. (2020). 2D-QSAR Modeling and Molecular Docking Studies on 1 H-Pyrazole-1-carbothioamide Derivatives as EGFR Kinase Inhibitors. *ACS omega* ,*5*(30),18662-18674.

-Hanson, J.; Paliwal, K.; Litfin, T.; Yang, Y. and Zhou, Y. (2019). Improving prediction of protein secondary structure, backbone angles, solvent accessibility and contact numbers by using predicted contact maps and an ensemble of recurrent and residual convolutional neural networks. *Bioinformatics*,*35*(14), 2403-2410

-Jensen, F. (2017). *Introduction to computational chemistry*. John wiley & sons.

- Kiriiri, G. K.; Njogu, P. M. and Mwangi, A. N. (2020). Exploring different approaches to improve the success of drug discovery and

development projects: a review. *Future Journal of Pharmaceutical Sciences*, *6*(1), 1-12.

-Laffort, (2020). PInterest of Splitting the Enthalpies of Vaporization in Four Distinct Parts Reflecting the Van der Waals and the Hydrogen Bonding Forces. *Open Journal of Physical Chemistry*, *10*(02), 117.

-Lewars, E. (2011). Computational chemistry. *Introduction to the theory and applications of molecular and quantum mechanics*, Springer.

-Luo, Q.C. and Zheng, Y.Z. (2021). Methods and Models of Theoretical Calculation for Single-Molecule Magnets. *Magnetochemistry*, *7*(8), 107.

-Mandour, A. A.; Nabil, N.; Zaazaa, H. E. and Abdelkawy, M. (2020). Review on analytical studies of some pharmaceutical compounds containing heterocyclic rings: brinzolamide, timolol maleate, flumethasone pivalate, and clioquinol. *Future Journal of Pharmaceutical Sciences*, *6*(1), 1-10.

-Marighetti, F. (2019). *Investigation of BCRP-inhibitors using QSAR and machine learning methods* , Ph.D. Thesis, University of und Landesbibliothek Bonn)., Bonn ,Germany .

-Minovski, N. and Novič, M. (2017). Integrated in Silico Methods for the Design and Optimization of Novel Drug Candidates: A Case Study on Fluoroquinolones–Mycobacterium tuberculosis DNA Gyrase Inhibitors. In *Oncology: Breakthroughs in Research and Practice* ,434-481,IGI Global.

-Moorthy, N. H. N.; Martins, S. A.; Sousa, S. F.; Ramos, M. J. and Fernandes, P. A. (2014). Classification study of solvation free energies of organic molecules using machine learning techniques. *RSC advances*, *4*(106), 61624-61630.

- Muhammad, U.; Uzairu, A. and Ebuka Arthur, D.(2018). Review on: quantitative structure activity relationship (QSAR) modeling. *J Anal Pharm Res*, *7*(2), 240-242.

-Nadendla, R.R. (2004). Molecular modeling: A powerful tool for drug design and molecular docking. *Resonance*, *9*(5), 51-60.

-Patel, H.M.; Noolvi, M.N.; Sharma, P.; Jaiswal, V.; Bansal, S.; Lohan, S.; Kumar, S.S.; Abbot, V.; Dhiman, S. and Bhardwaj, V. (2014). Quantitative

structure–activity relationship (QSAR) studies as strategic approach in drug discovery. *Medicinal chemistry research*, *23*(12), 4991-5007.

-Pimentel, A.S.; Guimarães, C.R. and Miller, Y. (2013) .Molecular modeling: advancements and applications *Journal of chemistry*,2013

-Pratim Roy, P.; Paul, S.; Mitra, I. and Roy, K. (2009). On two novel parameters for validation of predictive QSAR models. *Molecules*, *14*(5), 1660-1701.

-Quin, L.D. and Tyrell, J.A. (2010). *Fundamentals of heterocyclic chemistry: importance in nature and in the synthesis of pharmaceuticals*. John Wiley & Sons.

-Rosas-Jimenez, J.G.; Garcia-Revilla, M.A.; Madariaga-Mazon, A. and Martinez-Mayorga, K.A . (2020). Validated and interpretable predictive model of cruzain inhibitors. *BioRxiv*.

-Roy, K. and Ghosh, G. (2010). Exploring QSARs with Extended Topochemical Atom (ETA) indices for modeling chemical and drug toxicity. *Current pharmaceutical design*, *16*(24), 2625-2639.

-Ruswanto, R.; Mardianingrum, R.; Lestari, T.; Nofianti, T. and Rahayuningsih, N. (2019). Synthesis and molecular docking of isonicotinohydrazide derivatives as anti-tuberculosis candidates. *Malaysian Journal of Fundamental and Applied Sciences*, *15*(3), 367-371.

-Rychtarčíková, Z.; Krátký, M.; Gazvoda, M.; Komlóová, M.; Polanc, S.; Kočevar, M.;Stolaříková, J. and Vinšová, J. (2014). N-substituted 2-isonicotinoyl hydrazine carboxamides-new antimycobacterial active *molecules*, 19(4), 3851–3868.

-Sakagami, H.; Uesawa, Y.; Ishihara, M.; Kagaya, H.; Kanamoto, T.; Terakubo, S. and Sugita, Y. (2015). Quantitative structure–cytotoxicity relationship of oleoylamides. *Anticancer Research*, *35*(10), 5341-5351.

-Serafim, M. S.; Gertrudes, J. C.; Costa, D. M.; Oliveira, P. R.; Maltarollo, V. G.;Honorio, K. M. (2021). Knowing and combating the enemy: a brief review on SARS-CoV-2 and computational approaches applied to the discovery of drug candidates. *Bioscience Reports*, *41*(3), BSR20202616.

-Setiawan, E.; Karna Wijaya, M. M . (2021). QSAR modeling for predicting the antifungal activities of gemini imidazolium surfactants against Candida albicans using GA-MLR methods. *Journal of Applied Pharmaceutical Science*, *11*(04), 022-027.

-Tripathi, A.; Misra, K. (2017). Molecular docking: A structure-based drug designing approach. *JSM Chem*, *5*(2), 1042-1047.

-Tropsha, A.(2010). Best practices for QSAR model development, validation, and exploitation. *Molecular informatics*, *29*(7), 476-488.

-Veerasamy, R.; Rajak, H.; Jain, A.; Sivadasan, S.; Varghese, C. P. and Agrawal, R. K. (2011). Validation of QSAR models-strategies and importance. *Int. J. Drug Des. Discov*, *2*,(3)511-519.

-Wong, K.M.; Tai, H.K. and Siu, S.W. (2021). GWOVina: A grey wolf optimization approach to rigid and flexible receptor docking. *Chemical biology & drug design*, *97*(1), 97-110.

Appendixes

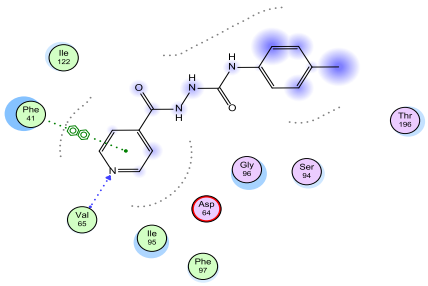


Figure (4.1) The interactions between the ligand (3a) and the 4TRO receptor active side

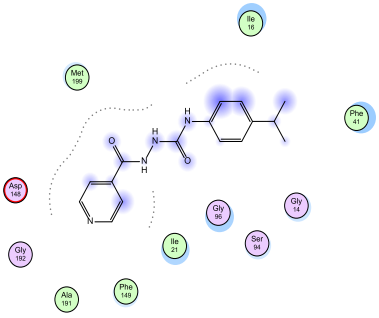


Figure (4.2) The interactions between the ligand (3b) and the 4TRO receptor active side

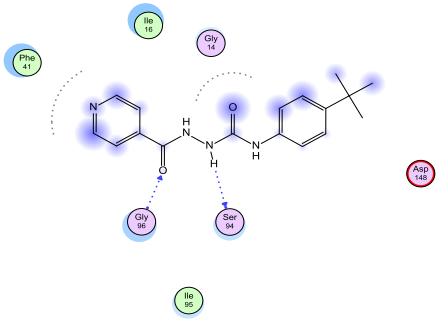


Figure (4.3) The interactions between the ligand (3c) and the 4TRO receptor active side

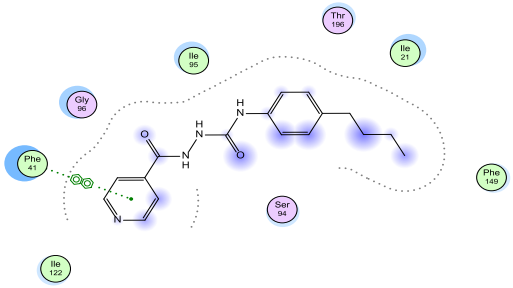


Figure (4.4) The interactions between the ligand (3d) and the 4TRO receptor active side

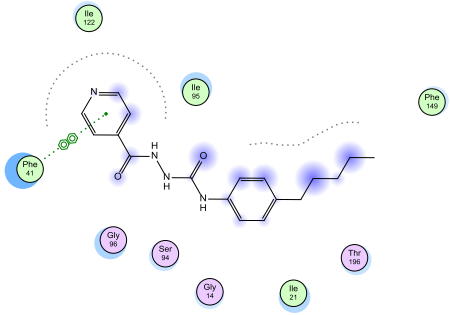


Figure (4.5) The interactions between the ligand (3e) and the 4TRO receptor active side

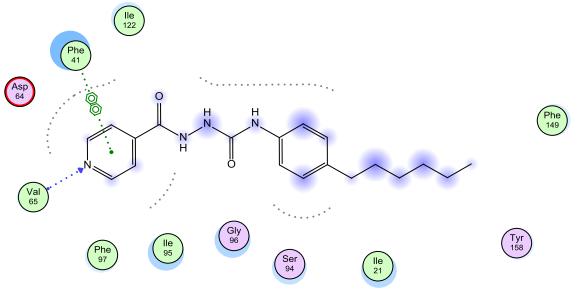


Figure (4.6) The interactions between the ligand (3f) and the 4TRO receptor active side

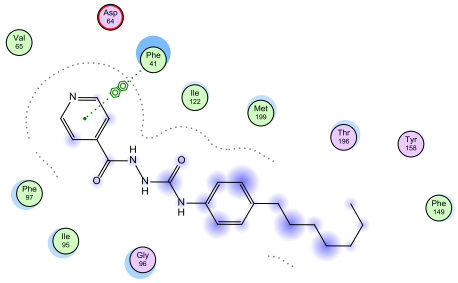


Figure (4.7) The interactions between the ligand (3g) and the 4TRO receptor active side

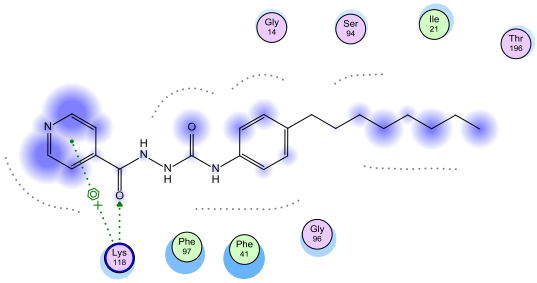


Figure (4.8) The interactions between the ligand (3h) and the 4TRO receptor active side

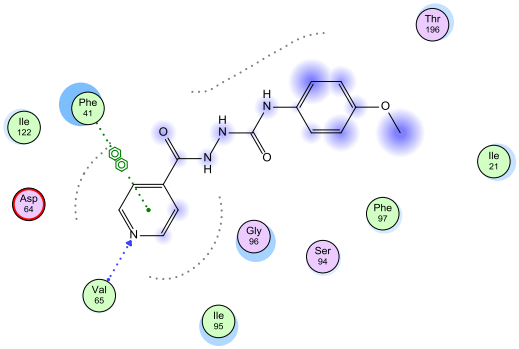


Figure (4.9) The interactions between the ligand (3i) and the 4TRO receptor active side

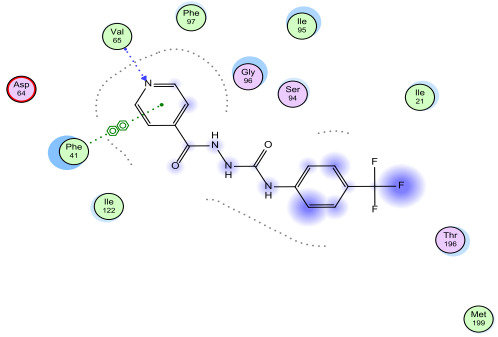


Figure (4.10) The interactions between the ligand (3j) and the 4TRO receptor active side

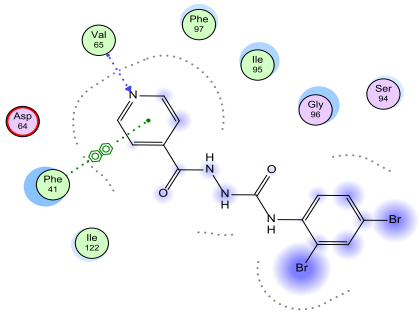


Figure (4.11) The interactions between the ligand (3k) and the 4TRO receptor active side

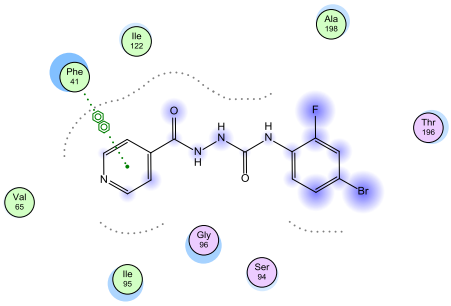


Figure (4.12) The interactions between the ligand (3l) and the 4TRO receptor active side

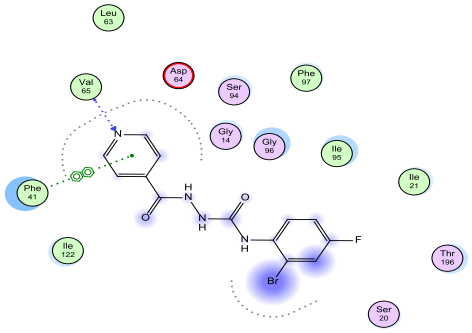


Figure (4.13) The interactions between the ligand (3m) and the 4TRO receptor active side

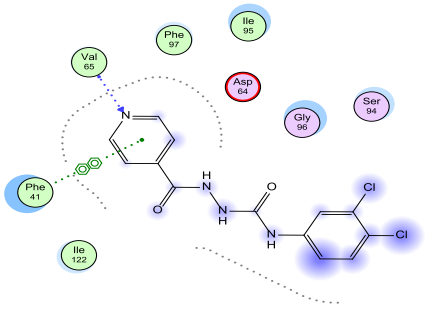


Figure (4.14) The interactions between the ligand (3n) and the 4TRO receptor active side

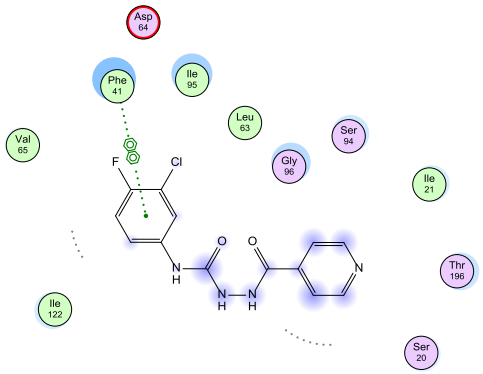


Figure (4.15) The interactions between the ligand (3o) and the 4TRO receptor active side

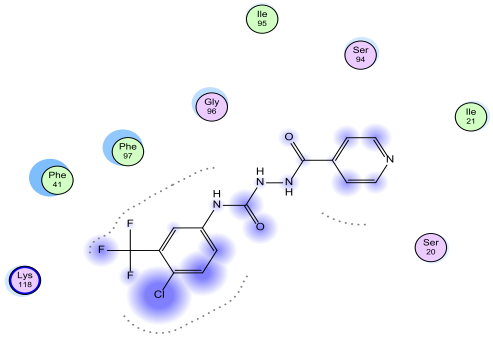


Figure (4.16) The interactions between the ligand (3p) and the 4TRO receptor active side

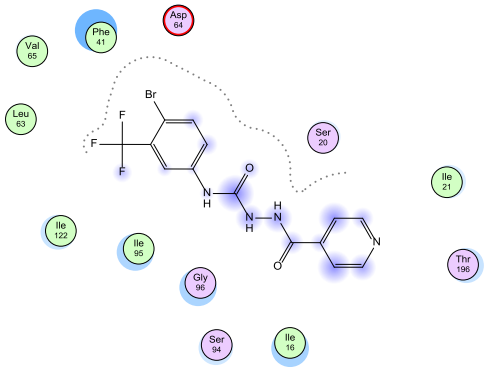


Figure (4.17) The interactions between the ligand (3q) and the 4TRO receptor active side

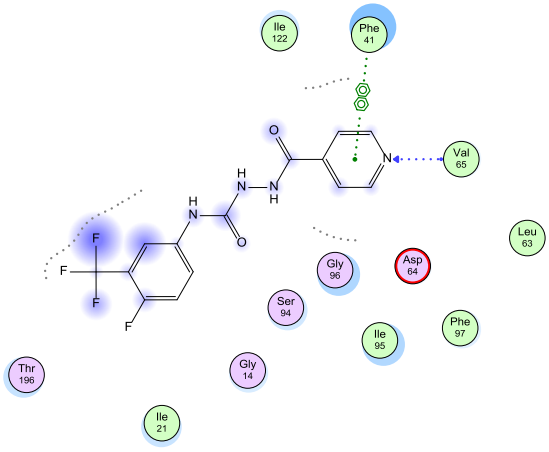


Figure (4.18) The interactions between the ligand (3r) and the 4TRO receptor active side

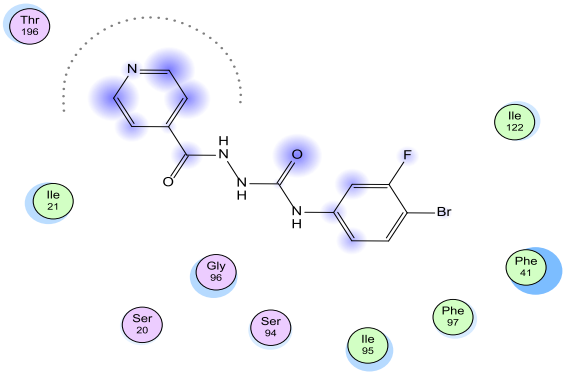


Figure (4.19) The interactions between the ligand (3s) and the 4TRO receptor active side

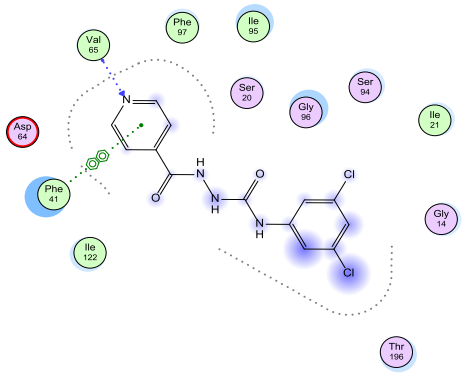


Figure (4.20) The interactions between the ligand (3t) and the 4TRO protein active side

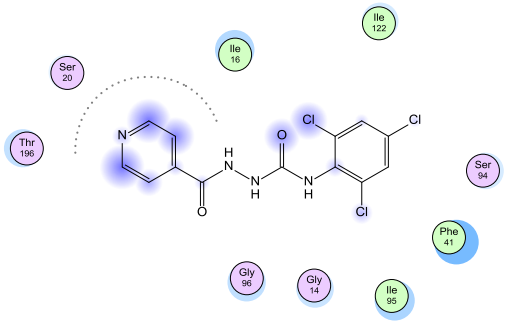


Figure (4.21) The interactions between the ligand (3u) and the 4TRO receptor active side

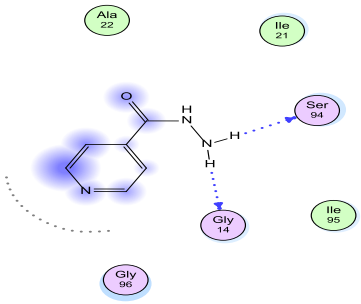


Figure (4.22) The interactions between the ligand (Isoniazid) and the 4TRO receptor active side

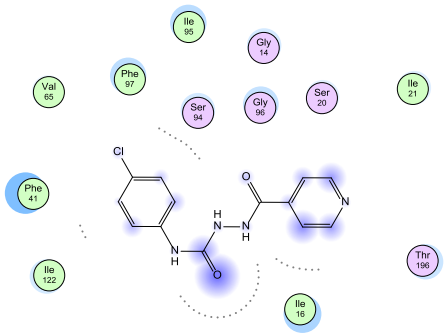


Figure (4.23) The interactions between the ligand (compound I) and the 4TRO protein active side

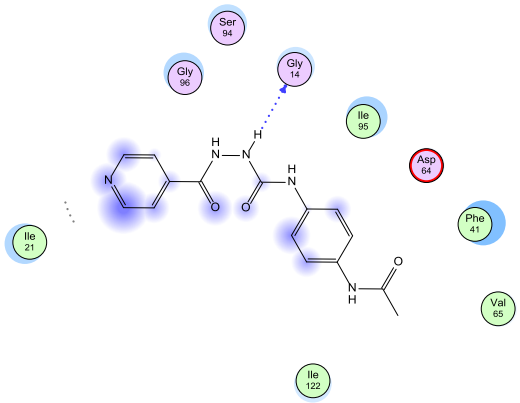


Figure (4.24) The interactions between the ligand (compound II) and the 4TRO receptor active side

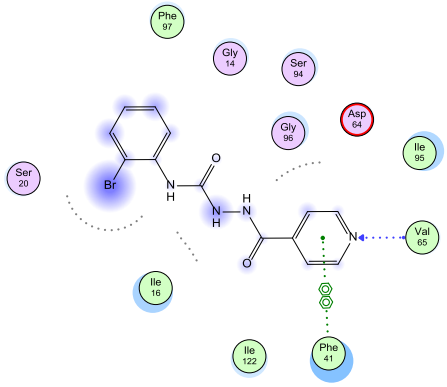


Figure (4.25) The interactions between the ligand (compound III) and the 4TRO receptor active side

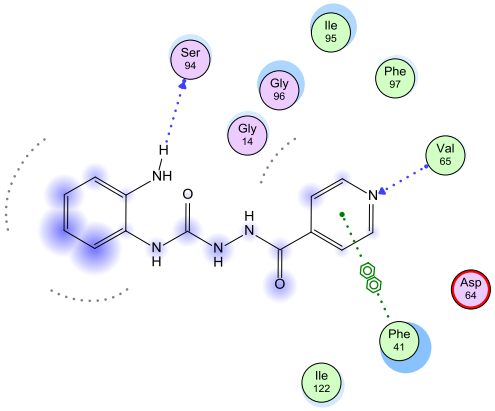


Figure (4.26) The interactions between the ligand (compound IV) and the 4TRO receptor active side

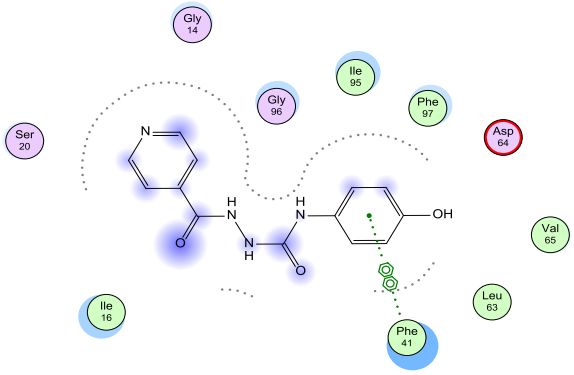


Figure (4.27) The interactions between the ligand (compound V)and the 4TRO receptor active side

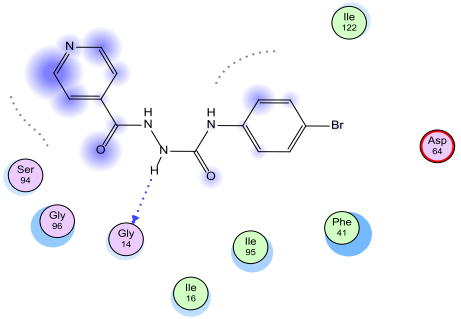


Figure (4.28) The interactions between the ligand (compound VI) and the 4TRO receptor active side

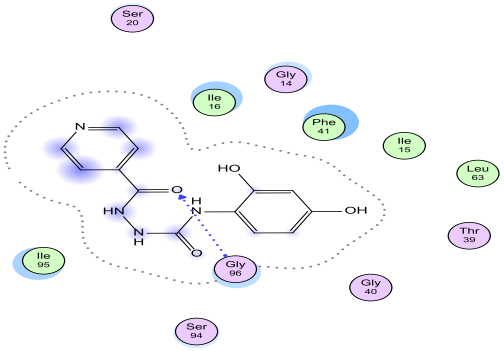


Figure (4.29) The interactions between the ligand (compound VII) and the 4TRO receptor active sid

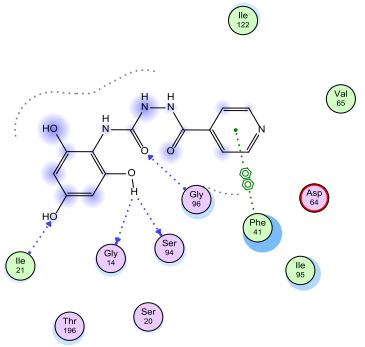


Figure (4.30) The interactions between the ligand (compound VIII) and the 4TRO receptor active side

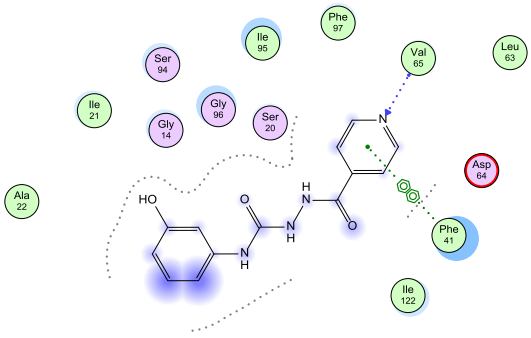


Figure (4.31) The interactions between the ligand (compound IX) and the 4TRO receptor active side

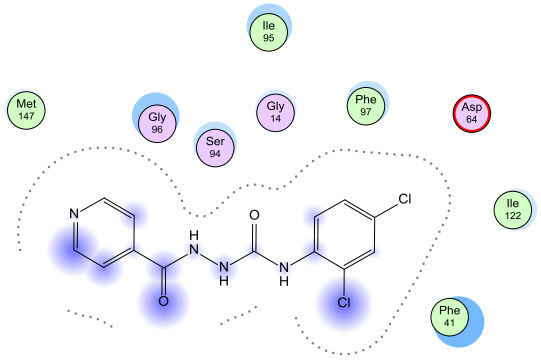


Figure (4.32) The interactions between the ligand (compound X) and the 4TRO receptor active side

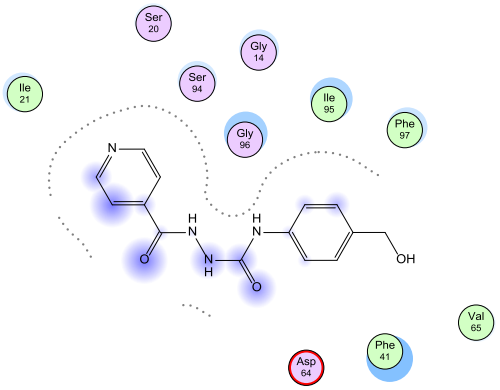


Figure (4.33) The interactions between the ligand (compound XI) and the 4TRO receptor active side

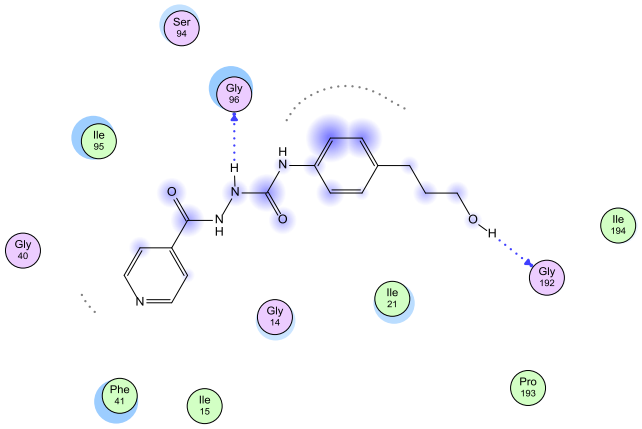


Figure (4.34) The interactions between the ligand (compound XII) and the 4TRO receptor active side

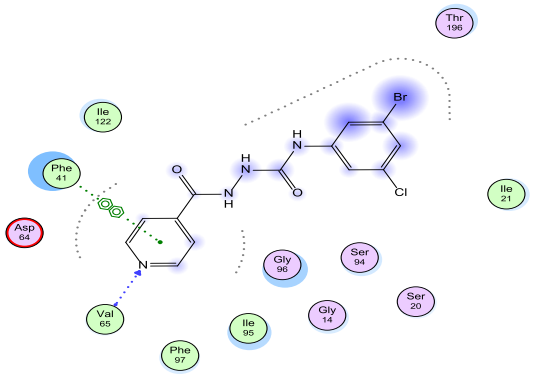


Figure (4.35) The interactions between the ligand (compound XIII) and the 4TRO receptor active side

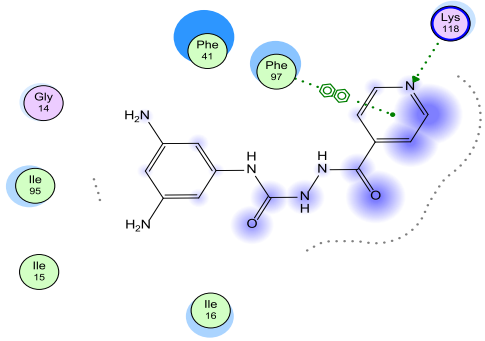


Figure (4.36) The interactions between the ligand (compound XIV) and the 4TRO receptor active side

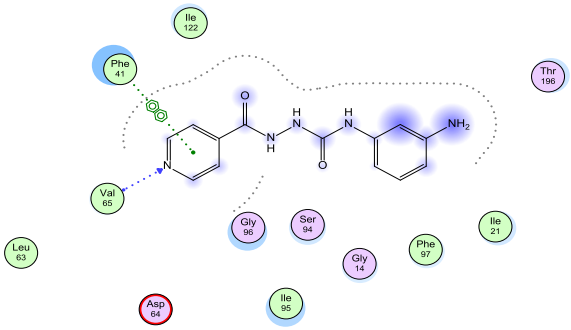


Figure (4.37) The interactions between the ligand (compound XV) and the 4TRO receptor active side

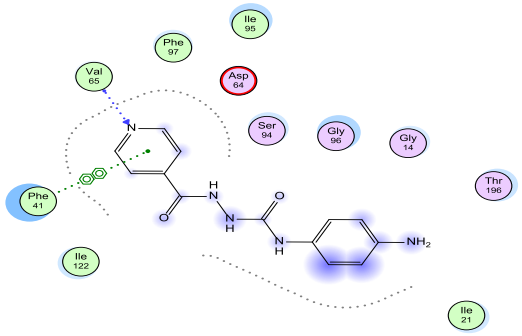


Figure (4.38) The interactions between the ligand (compound XVI) and the 4TRO receptor active side

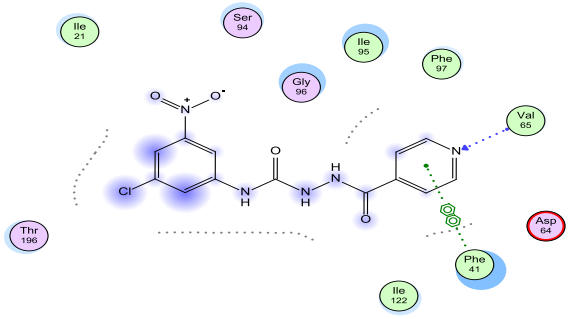


Figure (4.39) The interactions between the ligand (compound XVII) and the 4TRO receptor active side

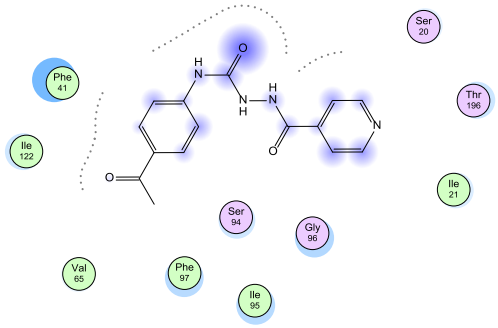


Figure (4.40) The interactions between the ligand (compound XVIII) and the 4TRO receptor active side

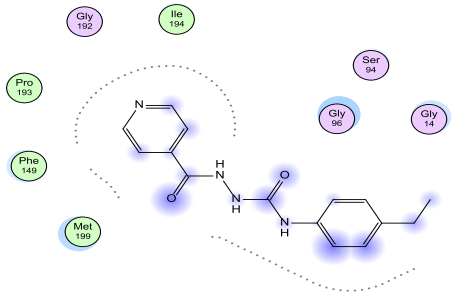


Figure (4.41) The interactions between the ligand (compound XIX) and the 4TRO receptor active side

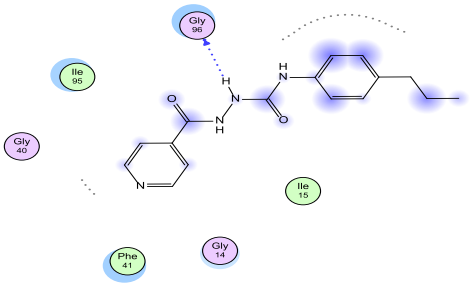


Figure (4.42) The interactions between the ligand (compound XX) and the 4TRO receptor active side

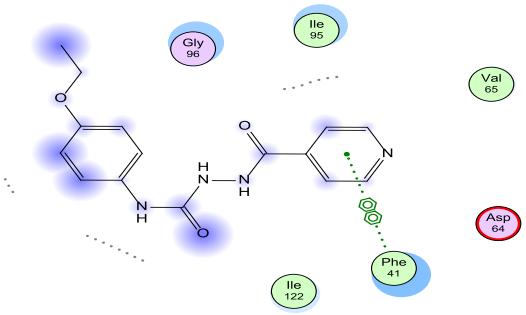


Figure (4.43) The interactions between the ligand (compound XXI) and the 4TRO receptor active side

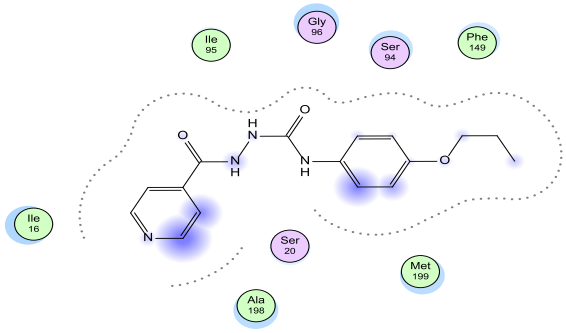


Figure (4.44) The interactions between the ligand (compound XXII) and the 4TRO receptor active side

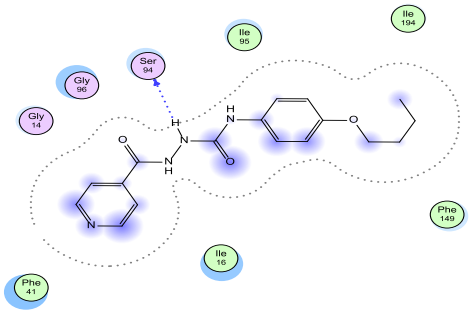


Figure (4.45) The interactions between the ligand (compound XXIII) and the 4TRO receptor active side

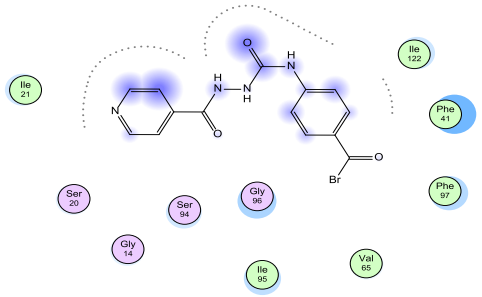


Figure (4.46) The interactions between the ligand (compound XXIV) and the 4TRO receptor active side

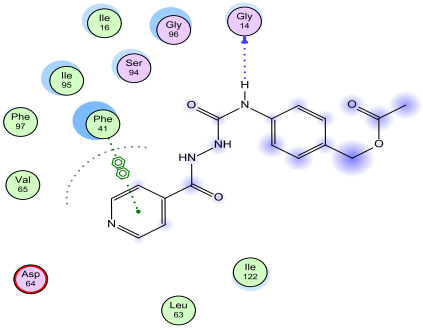


Figure (4.47) The interactions between the ligand (compound XXV) and the 4TRO receptor active side

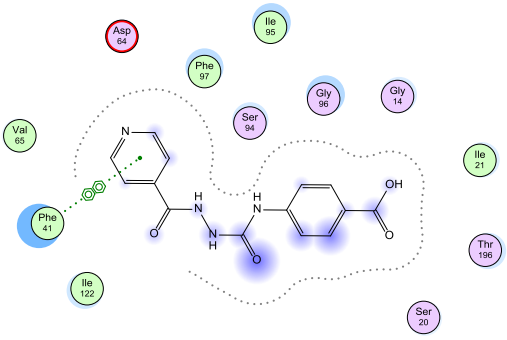


Figure (4.48) The interactions between the ligand (compound XXVI) and the 4TRO receptor active side

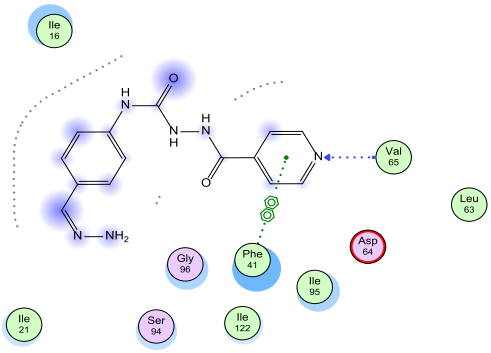


Figure (4.49) The interactions between the ligand (compound XXVII) and the 4TRO receptor active side

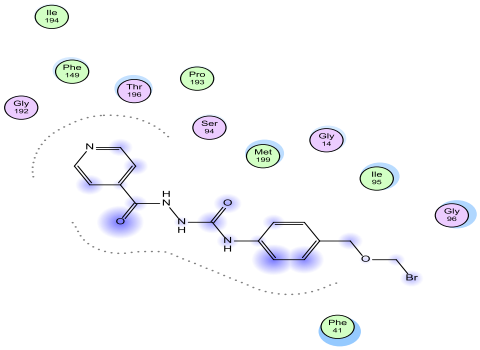


Figure (4.50) The interactions between the ligand (compound XXVIII) and the 4TRO receptor active side

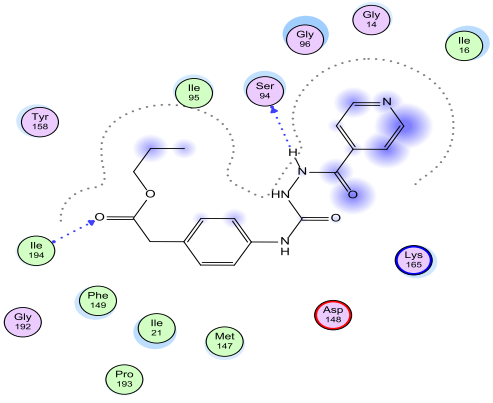


Figure (4. 51) The interactions between the ligand (compound XXIX) and the 4TRO receptor active side

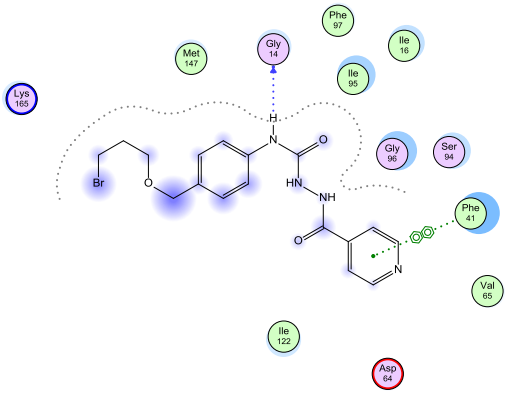


Figure (4.52) The interactions between the ligand (compound XXX) and the 4TRO receptor active side

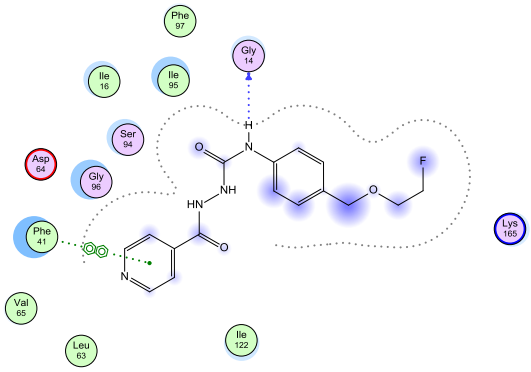


Figure (4.53) The interactions between the ligand (compound XXXI) and the 4TRO receptor active side

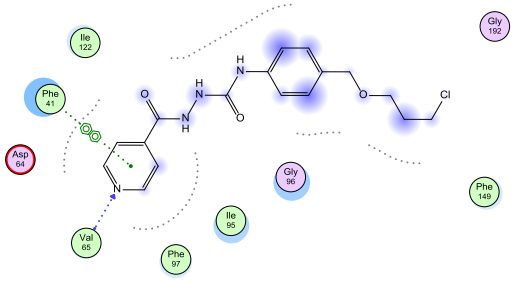


Figure (4.54) The interactions between the ligand (compound XXXII) and the 4TRO receptor active side

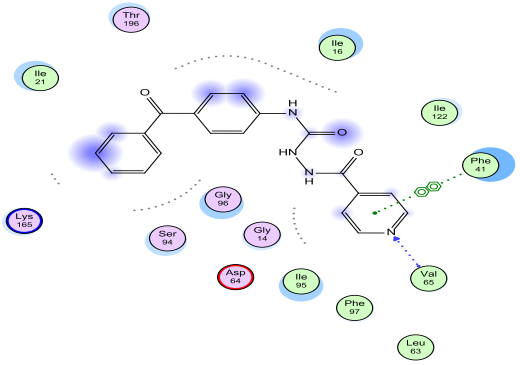


Figure (4.55) The interactions between the ligand (compound XXXIII) and the 4TRO receptor active side

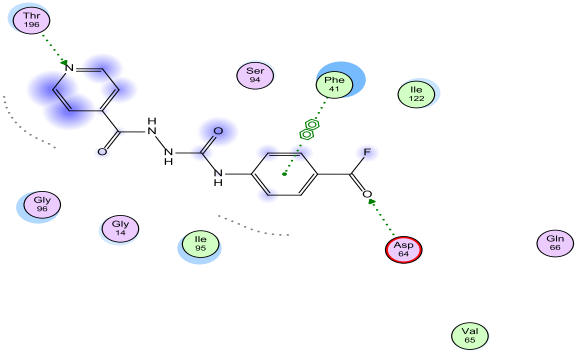


Figure (4.56) The interactions between the ligand (compound XXXIV) and the 4TRO receptor active side

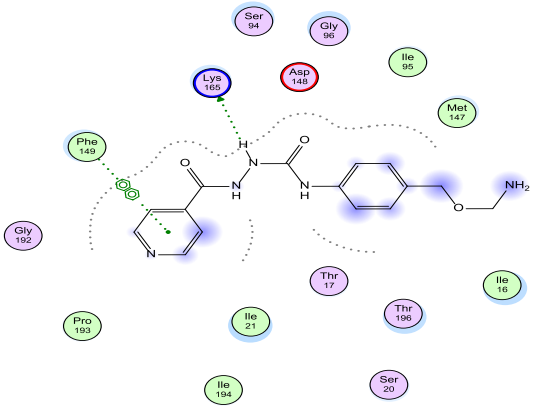


Figure (4.57) The interactions between the ligand (compound XXXV) and the 4TRO protein active side

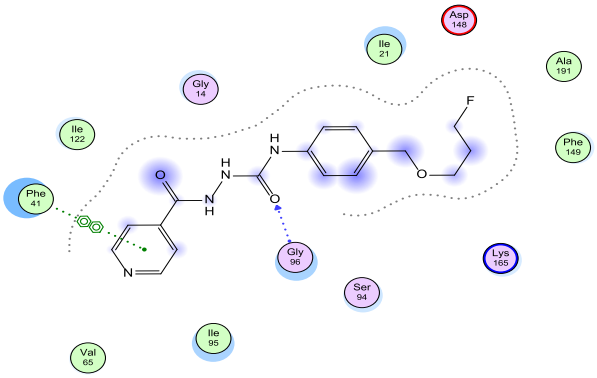


Figure (4.58) The interactions between the ligand (compound XXXVI) and the 4TRO receptor active side

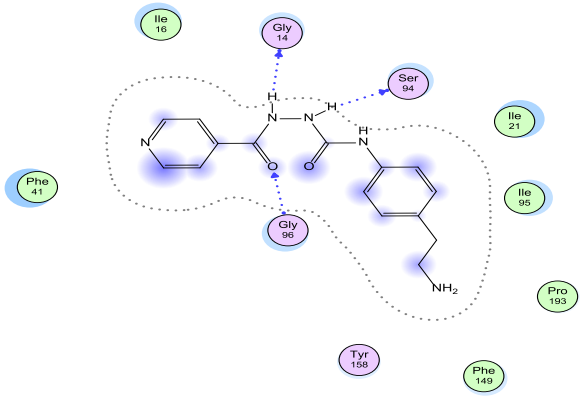


Figure (4.59) The interactions between the ligand (compoundXXXVII) and the 4TRO receptor active side

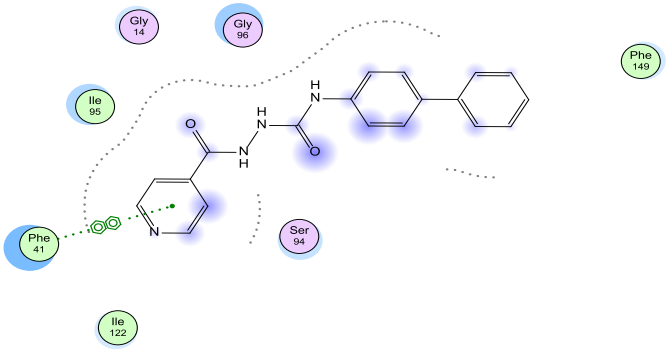


Figure (4.60) The interactions between the ligand (compound XXXVIII) and the 4TRO receptor active side

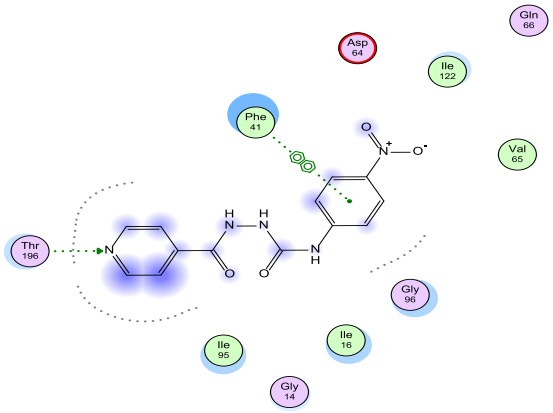


Figure (4.61) The interactions between the ligand (compound XXXIX) and the 4TRO receptor active side

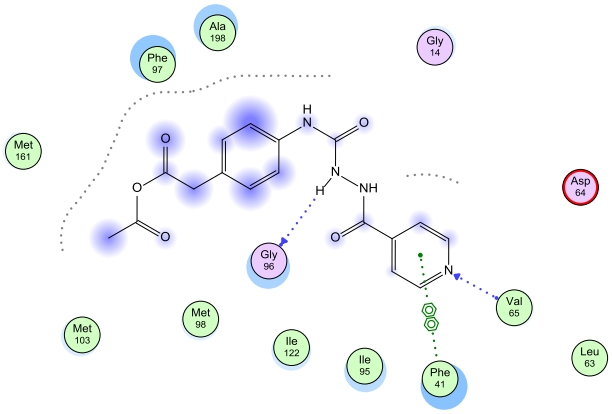


Figure (4.62) The interactions between the ligand (compound XL) and the 4TRO receptor active side

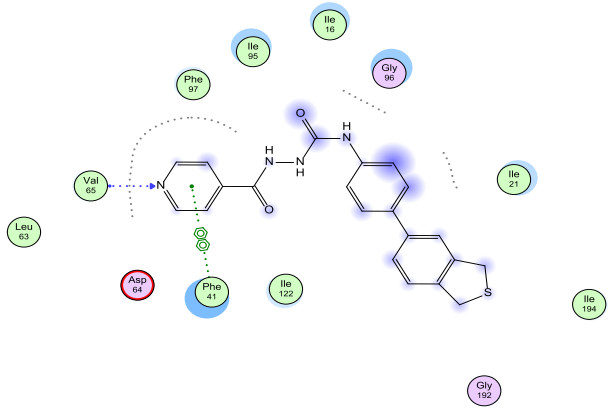


Figure (4.63) The interactions between the ligand (compound XLI) and the 4TRO receptor active side

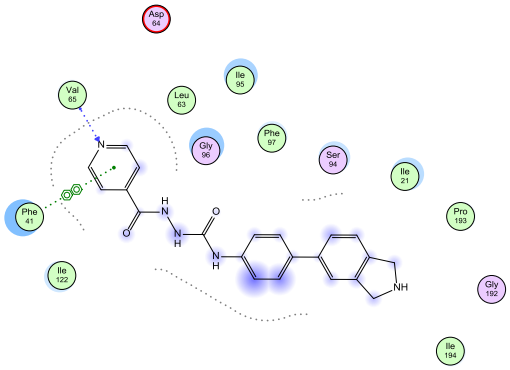


Figure (4.64) The interactions between the ligand (compound XLII) and the 4TRO receptor active side

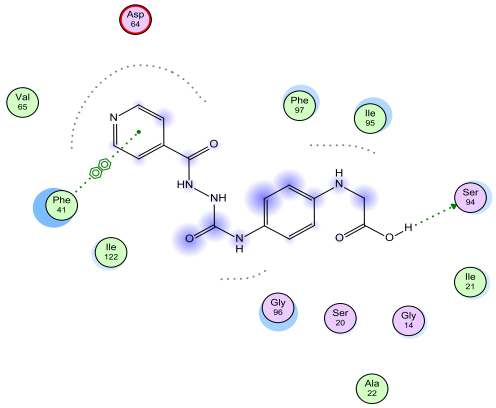


Figure (4.65) The interactions between the ligand (compound XLIII) and the 4TRO receptor active side

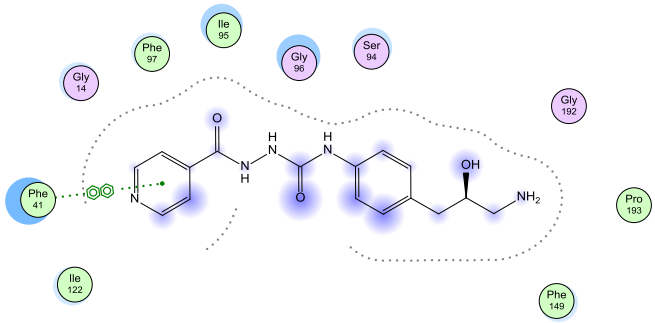


Figure (4.66) The interactions between the ligand (compound XLIV) and the 4TRO receptor active side

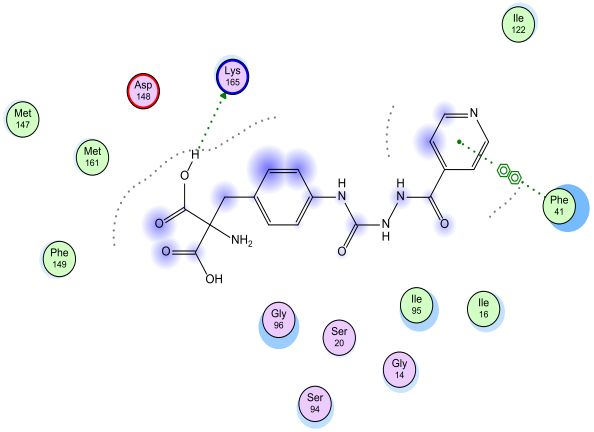


Figure (4.67) The interactions between the ligand (compound XLV) and the 4TRO receptor active side

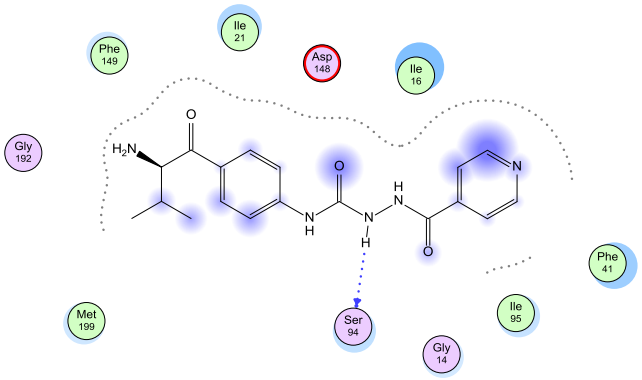


Figure (4.68) The interactions between the ligand (compound XLVI) and the 4TRO receptor active side

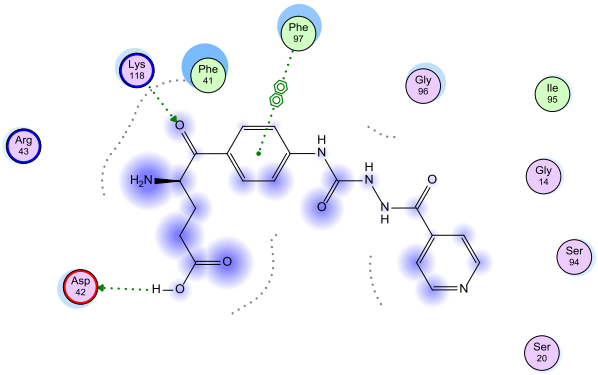


Figure (4.69) The interactions between the ligand (compound XLVII) and the 4TRO receptor active side

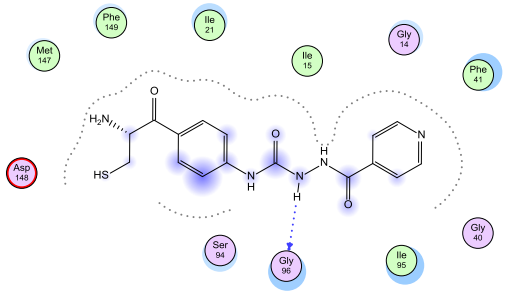


Figure (4.70) The interactions between the ligand (compound XLVIII) and the 4TRO receptor active side

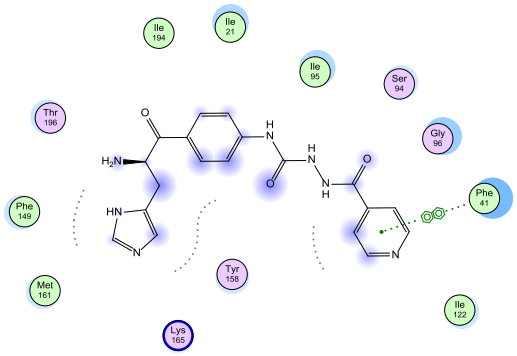


Figure (4.71) The interactions between the ligand (compound XLIX) and the 4TRO receptor active side

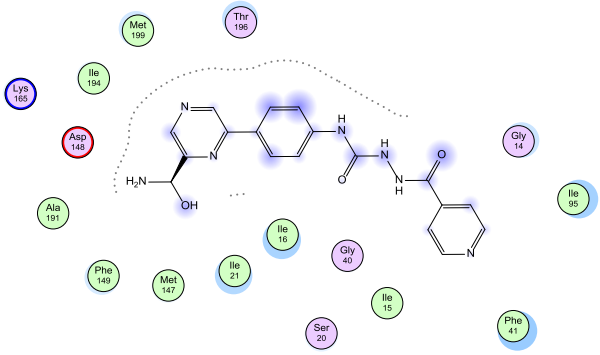


Figure (4.72) The interactions between the ligand (compound L) and the 4TRO receptor active side

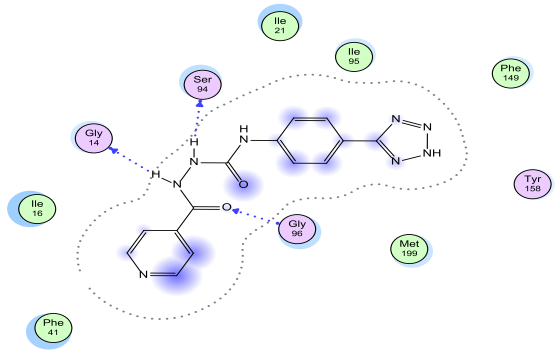


Figure (4.73) The interactions between the ligand (compound LI) and the 4TRO receptor active side

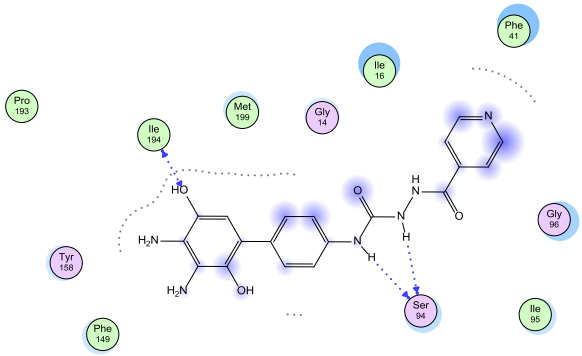


Figure (4.74) The interactions between the ligand (compound LII) and the 4TRO receptor active side

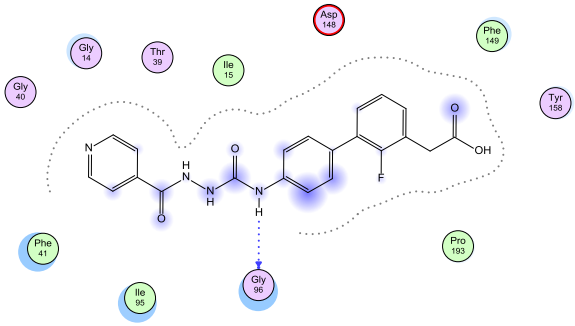


Figure (4.75) The interactions between the ligand (compound LIII) and the 4TRO receptor active side

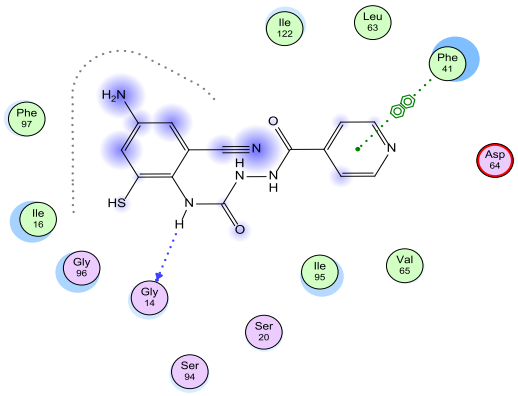


Figure (4.76) The interactions between the ligand (compound LIV) and the 4TRO receptor active side

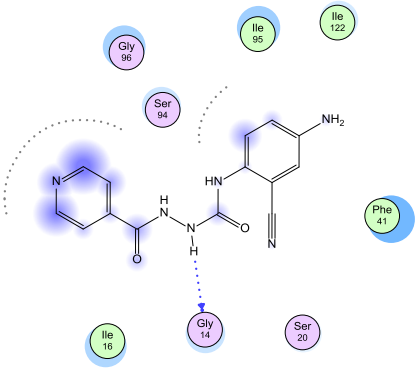


Figure (4.77) The interactions between the ligand (compound LV) and the 4TRO receptor active side

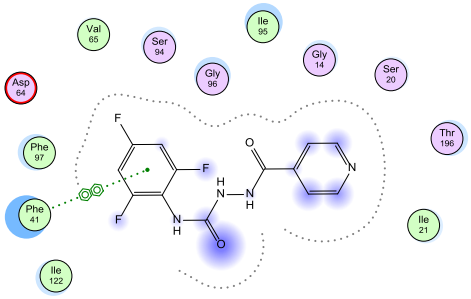


Figure (4.78) The interactions between the ligand (compound LVI) and the 4TRO receptor active side

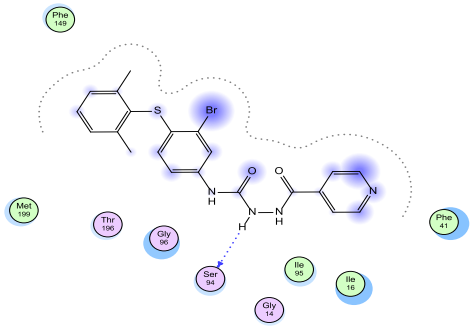


Figure (4.79) The interactions between the ligand (compound LVII) and the 4TRO receptor active side

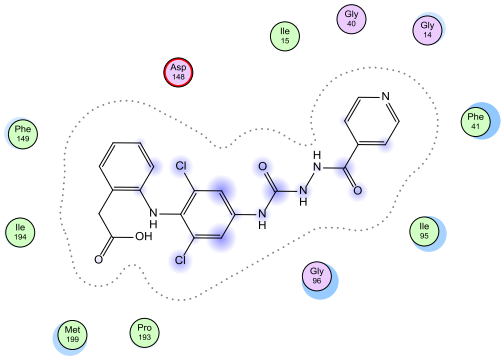


Figure (4.80) The interactions between the ligand (compound LVIII) and the 4TRO receptor active side