DEMOCRATIC AND POPULAR REPUBLIC OF ALGERIA
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

MOHAMED KHIDER UNIVERSITY OF BISKRA

FACULTY OF EXACT SCIENCES, SCIENCES OF NATURE AND LIFE

DEPARTMENT OF MATTER SCIENCES

DISSERTATION

Presented by

DJEBAILI Rachida

To obtain the degree of

Doctorate in Chemistry

Option:
Molecular chemistry

Entitled:

Structure-activity approaches for prediction of chemical reactivity and pharmacological properties of some heterocyclic compounds

Publicly defended on: 15/12/2022

In front of the jury composed of:

Mr. BELAIDI Salah	Prof.	University of Biskra	President
Mr. MELKEMI Nadjib	Prof.	University of Biskra	Supervisor
Mr. KENOUCHE Samir	MCA	University of Biskra	CO-Supervisor
Mr. HELLAL Abdelkader	MCA	University of Setif 1	Examiner
Mr. NOUAR Yacine	MCA	University of Setif 1	Examiner
Mr. DAOUD Ismail	MCA	University of Biskra	Examiner

This dissertation is dedicated to my dear parents
To my beloved brothers and sisters
To my precious uncles and aunts
To all those who are dear to me

Acknowledgments

In the name of Allah, the most gracious and the most merciful. First and foremost, I am thankful to Almighty ALLAH for giving me the strength, knowledge, ability, and opportunity to undertake this study and complete it satisfactorily.

I would like to acknowledge and give my warmest thanks to my respected teacher and supervisor Prof. MELKEMI Nadjib, first of all, I am very grateful for the countless noble human qualities with which you have supervised my dissertation over these years, thank you for your kindness, humility, patience, and encouragement in times of despair and weakness. Furthermore, I am very grateful for your scientific competence, efforts in supervision, availability at all times, confidence, and energy that motivates me through academic challenges to produce this dissertation and finish my doctorate. Indeed, it’s an honor for me that you are supervising my dissertation.

I would like to address my deep gratitude for my Co-supervisor Dr. KENOUCHE Samir, there are no words to describe how proud I am to work with you and learn from your vast knowledge, I learned from you to be persistent in my work, scientifically honest, and to not limit my knowledge to a single field. Thank you for your kindness, encouragement, diligence, rigor work, patience in answering my questions, valuable advice, and appreciated professional attitude in completing this dissertation. I will never forget your inspiring words, "C’est en forgeant qu’on devient forgeron".

I would like to thank my committee members, Prof. BELAIDI Salah, Dr. HELLAL Abdelkader, Dr. NOUAR Yacine, and Dr. DAOUD Ismail, for accepting to judge this work.

I would address a special thanks to the one who taught us the team's working spirit, my teacher, Prof. BELAIDI Salah, Head of Computational and Pharmaceutical Chemistry Group, LMCE Laboratory, University of Mohamed Khider, Biskra, for his guidance in my study pathway.

I thankfully acknowledge the support and inspiration that I received from Dr. HELLAL Abdelkader, who contributed to introducing me to the field of research. I would like to extend my deepest gratitude for your kindness, generosity in sharing knowledge, valuable skills, encouragement, and your profound belief in my abilities. Thank you for always finding the time to answer my questions.

I would like to express my deepest appreciation to the idol of ethics, science, and diligence, Dr. DAOUD Ismail, I am very grateful for your kindness, encouragement, academic/moral support, and generosity in sharing your expertise and answering all my questions. Thank you for your guidance, suggestions, and comments, as well as your valuable contribution to this dissertation.

Many thanks to my best friends and colleagues, METTAI Merzaka, ALMI Imane, BELKADI Ahlem, and SAGHIRI Khadijah. My dears, I want to express my gratitude for your unwavering support throughout my difficult times, your constant support throughout my PhD career, and your continuing motivation that gives me energy and positivity.

I would like to thank the Department of Matter Sciences, and MOHAMED KHIDER UNIVERSITY OF BISKRA for offering me the opportunity to pursue my PhD.

Finally, I would like to thank the best parents in the world, my parents, LAKHDAR and FATMA, and my beloved brother SALIM, for whom there are no words to express how grateful I am for their endless support and keen interest in my academic achievements. In particular, my dear father, my first supporter, my idol is ethics, love of knowledge, and discipline in work, who always looks forward to the moment of my graduation, I hope I made you proud of me, as I am of you.

Abstract

Benzodiazepine drugs are widely prescribed to treat many psychiatric and neurologic disorders. As its pharmacological action is exerted in a sensitive area of the brain; "the central nervous system", it is crucial to provide detailed reports on the chemistry of benzodiazepines, model the mechanism of action that occurs with $GABA_A$ receptors, identify the overlap with other modulators, as well as explore the structural requirements that better potentiate the receptor response to benzodiazepines. This dissertation consists of two original studies that consider the new lines of research related to benzodiazepines, particularly the identification of three new TMD binding sites. The first study provides, on the one hand, an overview of the chemistry of six Benzodiazepine basic rings starting from structural characteristics, electronic properties, Global/local reactivities, up to intermolecular interactions with long-range nucleophilic/electrophilic reactants. This was achieved by combining a DFT investigation with a quantitative MEP analysis on the vdW surface. On the other hand, the performed molecular docking simulations allowed identifying the best binding modes, binding interactions, and binding affinities with residues, which helped to validate the quantitative MEP analysis predictions. The second study was conducted on a dataset of [^{3}H]diazepam derivatives. First, molecular docking simulation was used to initially screen the dataset and select the best ligand/target complexes. Afterwise, the best-docked complexes were refined by performing molecular dynamics simulation for 1000 picoseconds. For both simulations, the binding modes, binding interactions, and binding affinities were thoroughly discussed and compared with each other and with outcomes collected from the literature. Additionally, the good pharmacokinetic properties (ADME prediction) as well as compliance with all drug-likeness rules were checked via in silico tools for all the dataset compounds. Finally, a QSAR analysis was carried out using an improved version of PLS regression. Briefly, the dataset is randomly split into 10 000 training and test sets that involve, respectively, 80% and 20% of chemicals. Subsequently, 10 000 statistical simulations were conducted that; after excluding outlying observations, yielded 10 000 best training models following the Bayesian Information Criterion. Among these 10 000 best models, the best predictors with the highest probability of occurrence were selected. As a consequence, the derived PLS regression equation explains 63.2% of the variability in BDZ activity around its mean. Furthermore, Internal and external validation metrics assure the robustness and predictability of the developed model. The developed model was interpreted based on literature investigations and a combination of implemented approaches.

Keywords: Benzodiazepine, $GABA_A$ receptor, allosteric modulation, DFT, molecular docking, molecular dynamic, MEP, QSAR, PLS, ADME.

الملـــــخص

توصف أدوية البنزوديازيبين على نطاق واسع لعلاج العديد من الاضطرابات النفسية والعصبية. نظرا لأن تأثيرها الدوائي يمارس في منطقة حساسة من الدماغ "الجهاز العصبي المركزي"، فمن الأساسي توفير تقارير مفصلة عن كيمياء البنزوديازيبينات، نمذجة آلية تأثيرها على مستقبلات $GABA_A$، تحديد تداخلاتها مع المُعدِّلات الأخرى، وكذلك تحديد المتطلبات الهيكلية الازمة لتعزيز استجابة المستقبل لها بشكل أفضل. تتكون هذه الرسالة من دراستين أصليتين تأخذان في الإعتبار خطوط البحث الجديدة المتعلقة بالبنزوديازيبينات، بالأخص مواقع الربط الثلاث التي تم تحديدها مؤخرا في بنية المجال عبر-الغشائي (TMD). تقدم الدراسة الأولى، من ناحية، نظرة عامة على كيمياء ست حلقات أساسية للبنزوديازيبين بدءًا من الخصائص الهيكلية، الخصائص الإلكترونية، التفاعلات العالمية/المحلية، وصولا الى التفاعلات الجزيئية مع المتفاعلات البعيدة الكارهة للكهرباء/المحبة للكهرباء. تم تحقيق ذلك عبر جمع نظرية الكثافة الوظيفية (DFT) مع التحليل الكمي للجهد الكهروستاتيكي الجزيئي (MEP) على سطح vdW. من ناحية أخرى ، سمحت عمليات محاكاة الالتحام الجزيئي المجراة بتحديد أفضل أوضاع الربط، تفاعلات الربط، وألفة الارتباط مع الأحماض الأمينية، مما ساعد على التحقق من صحة تنبؤات التحليل الكمي للجهد الكهروستاتيكي الجزيئي (MEP). أجريت الدراسة الثانية على مجموعة بيانات من مشتقات [^{3}H] ديازيبام. أولاً، تم استخدام محاكاة الالتحام الجزيئي لفحص مجموعة البيانات مبدئيا واختيار أفضل المعقدات جزيء/مستقبل. بعد ذلك ، تم تنقيح هذه الأخيرة عن طريق إجراء محاكاة الديناميكيات الجزيئية لمدة 1000 بيكو ثانية. لكلتا المحاكاتان، تمت مناقشة أوضاع الربط، تفاعلات الربط، وألفة الارتباط مع الأحماض الأمينية بشكل شامل ومقارنتها فيما بينها و بالنتائج التي تم جمعها من الأبحاث العلمية السابقة. أيضا، تم التحقق من ملائمة جميع مركبات مجموعة البيانات للخصائص الحركية الدوائية وكذلك مدى امتثالها لقواعد تشابه الأدوية عبر استعمال الأدوات الحسابية. أخيرًا، تم إجراء التحليل الكمي للعلاقة بنية-نشاط (QSAR) باستخدام نسخة محسنة من انحدار PLS. باختصار، تم تقسيم مجموعة البيانات بشكل عشوائي إلى 10000 مجموعة تدريب واختبار تضم كل منها، على التوالي، 80٪ و 20٪ من العناصر الكيميائية. بعد ذلك ، تم إجراء 10000 محاكاة إحصائية، و بعد استبعاد القيم المتطرفة تم الحصول على 10000 أفضل نموذج تدريب وفقًا لمعيار المعلومات البايزي (BIC). من بين هذه الأفضل 10000 نموذج، تم اختيار أفضل المتنبئات طبقا لأعلى احتمال لحدوثها. كنتائج، تشرح المعادلة المشتقة من انحدار PLS63% من التباين في نشاط البنزوديازيبينات حول متوسطه. علاوة على ذلك ، تضمن مقاييس التحقق الداخلية والخارجية متانة النموذج المطور وإمكانية التنبؤ به. تم تفسير النموذج المطور بناءً على تحقيقات الأبحاث العلمية السابقة ومجموعة الأساليب المنفذة.

الكلمات المفتاحية: البنزوديازيبين, مستقبلات الجابا (أ), الأنواع الفرعية, مجال عبر-الغشاء, المجال خارج الخلية, بطانة المسام, اللولب عبر الغشائي 2, القناة الأيونية, التعديل الخيفي, النقل العصبي, نظرية الكثافة الوظيفية, الإرساء الجزيئي, محاكاة الديناميكية الجزيئية، التحليل الكمي للجهد الكهروستاتيكي الجزيئي، التحليل الكمي للعلاقة بنية-نشاط, الإنحدار, إمتصاص-توزيع-معالجة-إطراح.

Contents

Chapter I

Overview on Neurotransmission, GABA receptors, and Benzodiazepines

Chapter II

Overview on computational methods: Conceptual-DFT, MEP analysis, QSAR analysis, and Molecular docking/Dynamic simulations

Chapter III

Combined Conceptual-DFT, Quantitative MEP Analysis, and Molecular Docking Study of Benzodiazepine Analogs

Chapter IV

Investigation of [3H]diazepam derivatives as allosteric modulators of GABAA receptor α1β2γ2 subtypes: combination of molecular docking, molecular dynamics simulation, pharmacokinetics/drug-likeness prediction, and QSAR analysis

List of works

International publications

1. Djebaili R, Kenouche S, Daoud I, et al (2022) Investigation of [3H]diazepam derivatives as allosteric modulators of $GABA_A$ receptor $\alpha_1\beta_2\gamma_2$ subtypes: combination of molecular docking/dynamic simulations, pharmacokinetics/drug-likeness prediction, and QSAR analysis. Struct Chem 1–33. https://doi.org/10.1007/s11224-022-02029-4

2. Djebaili R, Melkemi N, Kenouche S, et al (2021) Combined Conceptual-DFT, Quantitative MEP Analysis, and Molecular Docking Study of Benzodiazepine Analogs. Orbital Electron J Chem 13:301–315. https://doi.org/10.17807/ORBITAL.V13I4.1607

3. Kenouche S, Belkadi A, Djebaili R, Melkemi N (2021) High regioselectivity in the amination reaction of isoquinolinequinone derivatives using conceptual DFT and NCI analysis. J Mol Graph Model 104:107828. https://doi.org/10.1016/j.jmgm.2020.107828

4. Belkadi A, Kenouche S, Melkemi N, et al (2022) Molecular docking/dynamic simulations, MEP, ADME-TOX-based analysis of xanthone derivatives as CHK1 inhibitors. Struct Chem 33:833–858. https://doi.org/10.1007/s11224-022-01898-z

5. Belkadi A, Kenouche S, Melkemi N, et al (2021) K-means clustering analysis, ADME/pharmacokinetic prediction, MEP, and molecular docking studies of potential cytotoxic agents. Struct Chem 32:2235–2249. https://doi.org/10.1007/s11224-021-01796-w

International conferences

- R. Djebaili, N. Melkemi, I. Daoud, S. Kenouche, molecular docking study of 1,4-benzodiazepine derivatives against γ-aminobutyric acide type A ($GABA_A$), 13th international days of theoretical and computational chemistry, 02-03 February 2020, Biskra, Algeria.
- R. Djebaili, N. Melkemi, QSAR studies of 1,4-benzodiazepine derivatives in gas phase and in water, Fifth International Colloquium on Chemistry (CIC-5), 04-06 December 2018, Batna, Algeria.

National conferences

- R. Djebaili, R. Benakcha, N. Melkemi, phytochemical study on sesquiterpene lactones of tilia cordata plant, biological activity, 1st scientific day on the biology of medicinal plants (BMP-2020), 22th January 2020, Tebessa, Algeria.
- R. Djebaili, N. Melkemi, quantitative structure-activity relationships modeling of 1,4-benzodiazepine derivatives, first day of study on molecular physico-chemistry, 06 december 2017, el oued, Algeria.

- R. Djebaili, N. Melkemi, Atomic reactivity of benzodiazepine basic rings using condensed fukui function, 1st Study Days on Chemistry and its Applications (JECA-1-2019), 27 November 2019, Batna, Algeria.

Conferences organization

- The 13th international days of theoretical and computational chemistry, 02-03 February 2020, Biskra, Algeria.

List of main abbreviations

ADME: Absorption, Distribution, Metabolism, Elimination
BBB: blood-brain barrier
BDZ: benzodiazepine
BIC: bayesian information criterion
CADD: computer aided drug design
CNS: central nervous system
CV: cross-validation method
CYP: cytochrome enzyme
DFT: density functional theory
DZP: diazepam
ECD: extracellular domain
FDA: finite difference approximation
FF: Fukui function
FMO: frontier molecular orbitals approximation
GABA: neurotransmitter γ-aminobutyric acid
GABA$_A$: type ‘’A’’ γ-aminobutyric acid receptor
GI: gastrointestinal absorption
HPA: Hirshfeld population analysis
ICD: intracellular domain
MD: molecular dynamic
MEP: molecular electrostatic potential
NS: negative surface
OV: overall variance
PDB: protein data bank
PLS: partial least squares regression
PLZ: proflazepam
PS: positive surface
P-gp: Permeability Glycoprotein
QSAR: quantitative structure relationship
RMSD: root-mean-score deviation
SF: scoring function
TMD: transmembrane domain
U: potential energy

List of figures

Chapter I

Chapter II

Chapter III

Chapter IV

List of Tables

Chapter I

Chapter II

Chapter III

Chapter IV

GENERAL INTRODUCTION

Discovering, developing, and bringing new drugs to market is a long-term process that costs billions of dollars and involves many risks. Typically, the drug discovery and development cycle, from designing the initial concept to bringing it to market, takes about 14 years [1]. The entire process of discovering new therapeutic agents can summarize into three main phases:

- The discovery phase: involves the selection of targets as well as the manufacture of effective compounds.
- The development phase: includes all the preclinical and clinical research (in vitro and in vivo assays).
- The registry phase: consists of the Food and Drug Administration (FDA) approval, clinical application of drug, and post-market drug safety monitoring [2].

In 2014, published statistics reported that the cost associated with this course had risen in the past decade by over 150% [3]. In this context, the mean cost of developing a single new therapeutic substance indicates a sharp increase from $1.1 billion in 2003 to $2.8 billion in 2013 [4,5]. Other estimates suggest values between $0.314 billion and $2.1 billion over the same period [6–11]. Among a data set of 63 FDA-authorized drugs between 2009 and 2018, the median capitalized research and development (R&D) cost per product was estimated to be over $0.985 billion, including unsuccessful trials. The median estimates for each therapeutic area (for sectors with at least five drugs) varied from $0.7659 billion for nervous system agents to $2.7716 billion for antineoplastic and immune-modulating agents [12]. Similarly, the investigation performed by *Prasad and Mailankody* on the real costs of R&D of new cancer drugs conducted on a sample of 10 FDA-approved drugs between 2006 and 2015, estimated the median cost of bringing one cancer drug to market to be at approximately $0.648 billion. While, compared to a total R&D investment of $7.2 billion, the total income from sales of these 10 drugs over a median of 4 years since approval is estimated at $67.0 billion. This figure reflects the large profit margins of drug manufacturers and has implications for the current debate over drug pricing [13]. Unfortunately, failures that occur throughout the drug discovery and design pipeline consume approximately 75% of R&D expenditures, with 90% of therapeutic agents entering clinical trials failing to obtain FDA clearance and reach the consumer market [3]. According to recent statistics, the FDA approval rate for candidate drugs entering Phase 1, Phase 2, and Phase 3 is 13.8%, 35.1%, and 59.0%, respectively [14].

To handle challenges in the drug discovery research cycle, bioinformaticians have sought to seize the rapid evolution in the field of computer hardware, software, and algorithms by integrating drug screening and design processes into various computational methods that dramatically minimize drug R&D time and expense. Overall, bioinformatics may assist provide potential target proteins for drug screening and design by identifying essential genes from a vast amount of genomic data. Besides, biomolecular simulations, when combined with experiments, may yield reasonably accurate protein configurations and thermodynamic features that are useful for identifying drug binding locus and elucidating drug action mechanisms [15].

October 5th, 1981, was considered by some researchers to be the beginning of the real breakthrough for Computer-Aided Drug Design (CADD), following the release of a cover article by Fortune magazine headlined “Next Industrial Revolution: Designing Drugs by Computer at Merck” [16,17]. CADD serves as a “virtual shortcut” in the drug discovery pipeline. It is primarily applied in the discovery phase where the main objective is to ensure that the lead-candidate drug enters preclinical assays, hence minimizing the time and expenses associated with drug development. Moreover, CADD may be used as part of the "drug repurposing" strategy that seeks to identify side effects and find alternative applications for drugs that have received FDA approval or have reached the market [2]. CADD incorporates two main tools that are used based on the availability of target structure information; structure-based (SBDD) and ligand-based (LBDD) drug discovery. LBDD is developed as an alternative to SBDD when the experimental target structure information is missing or unable to be determined through computational methods [1]. Applications of SB-techniques, in particular molecular docking/dynamics simulations, allow direct ligands docking to targets, which in turn, permits conceptualization of the binding mechanism involved in ligand-target recognition, as well as, the estimation of the binding affinity by monitoring the dynamic behavior [18]. Otherwise, LB-techniques, such as QSAR modeling and pharmacophore mapping, exploit the structural information of a dataset of ligands with known biological responses to either select the most active ligands or to determine structural improvements required to enhance the considered biological response [19]. Importantly, when both target and ligand structures are available, it's practical to perform a hybrid approach based on combining SBDD and LBDD tools or methods based on End-Points [2].

To better accomplish its objectives, CADD uses a mechanical method that expresses the structure of drug molecules as a function of energy. By minimizing the energy function, the optimal geometric configuration for different conformations of molecules can be detected. The mechanical methods most used in CADD are quantum mechanics (QM) and molecular mechanics (MM) [20]. Thanks to its success in describing the quantum mechanical ground state of electrons in systems of interest, density functional theory (DFT) has established itself as a sufficiently rigorous and efficient QM method for solving pharmaceutical issues. Nowadays, DFT is a well-liked method for precisely and inexpensively describing physiologically relevant molecular systems [21]. Overall, discussions on its effectiveness in studying specific molecular features for drug design versus experimental outcomes have been extensively reported in the literature [20,22].

Since lead-drugs must traverse many complex biological barriers to reach their target in an organism, pharmacokinetics and drug-likeness screening are nowadays routinely performed along with CADD tools to improve the success rate of drug candidates, as well as avoid toxicity, low efficacy, off-target interactions, and damaging effects by entering significant metabolic pathways. Pharmacokinetic studies involve the assessment of absorption, distribution, metabolism, and excretion (ADME) [23,24]. Poor ADME profile and low bioavailability have been estimated to account for approximately 40% of all clinical trial drug failures. Therefore, ADME screening through in silico tools assists in excluding inadequate candidates at the early steps of the drug discovery pipeline [23].

Fast inhibitory neurotransmission in the brain is mainly mediated by the neurotransmitter γ-aminobutyric acid (GABA) and its synaptic target, the GABA receptor type A ($GABA_A$). In the mammalian brain, the pentameric isoforms of $GABA_A$ are assembled from eight distinct families of subunits with a total of twenty subtypes. This rich physiological structure has enabled $GABA_A$ to be the target for a wide range of therapeutic, illicit, and recreational drugs, including benzodiazepines, barbiturates, anesthetics, and ethanol. These agents act via distinct binding sites and, possibly, via overlapping pathways [25].

The various benzodiazepine derivatives are positive/negative allosteric modulators or antagonists of $GABA_A$ that function cooperatively with the orthosteric GABA-recognition site. Benzodiazepine drugs exert sedative/hypnotic, anxiolytic, muscle relaxant, and anticonvulsant effects, which explain their effectiveness in treating epilepsy, insomnia, anxiety, and panic disorder [26,27]. Later, the anesthetic behavior of some high-dose benzodiazepines, such as diazepam, was also reported. The researchers' efforts attributed this variation in pharmacological effects to the presence of two distinct classes of binding sites on the structure of $GABA_A$. The first class includes a high-affinity binding site inserted at the ECD α_1^+/γ_2^- interface. This class has been known since the discovery of the first benzodiazepine and is responsible for the positive modulation that is essential in treating anxiety and seizure disorders by inhibiting neuronal excitability in the brain. The second class consists of three low-affinity binding sites located at TMD interfaces and partially overlapping the binding sites of intravenous anesthetics, which explains why high doses of certain benzodiazepines induce a direct anesthetic effect. However, the structural mechanisms underlying this variation are still largely unknown to date due to the lack of reliable structural data for $GABA_A$ [28,29].

This work belongs to research devoted to the study of the chemical reactivity and physiological effects of benzodiazepines through computational tools. This is accomplished by implementing a hybrid approach that merges DFT investigation, quantitative MEP analysis, CADD methods, and pharmacokinetic/drug-likeness prediction. This dissertation is broken up into two parts with a total of four chapters to discuss our outcomes and impart our knowledge.

Part I defines the study context and outlines the critical basic concepts required to understand the findings obtained in part II:

The first chapter provides a detailed review of the theoretical foundations of brain neurotransmission and neurotransmitter receptors, particularly those modulated by the neurotransmitter GABA, as well as it includes an in-depth description of the chemical/pharmacological properties and physiological effects of benzodiazepines under examination. Likewise, the second chapter provides an overview of the theoretical background of implemented computational methods. Initially, descriptions of global and local reactivity descriptors derived from conceptual-DFT and the quantitative MEP analysis are briefly mentioned. Next, CADD methods involving QSAR analysis, molecular docking, and molecular dynamics simulation were highlighted. Furthermore, the value of combining molecular docking/dynamic simulation is discussed.

Part II collects results and discussions:

The third chapter focused on investigating the chemical reactivity of six benzodiazepine basic rings by adopting a combinatorial approach based on conceptual-DFT, quantitative MEP analysis, and molecular docking simulation. First, statistical analysis was used to test the robustness of atomic charges to the basis sets, and then elucidate the differences in geometry and electronic properties. Afterwise, by computing global and local DFT-derived reactivity descriptors, the global and local reactivity of the considered rings have been determined and discussed. Quantitative MEP analysis on the van der Waals surface has also been defined and mapped to examine long-range intermolecular interactions. Finally, to confirm the outcomes of MEP, molecular docking simulations were carried out to predict the binding affinities of the issued molecules and estimate the binding poses into four $GABA_A$ receptor binding sites.

The fourth chapter explores how classical benzodiazepines modulate $GABA_A$ receptor $\alpha_1\beta_2\gamma_2$ subtypes and focuses on structural characteristics that enhance the receptor response better to benzodiazepine drugs. A dataset of $[^3H]$diazepam derivatives is the optimal choice for this purpose. Our investigation centered on four binding loci, three of them were recently identified, their pharmacological effect varies from causing neuronal inhibition at low doses to anesthetic effects at higher doses. A molecular docking simulation was used to initially screen the dataset, and determine the best complexes. With the use of molecular dynamics simulations and pharmacokinetics/drug-likeness predictions, the best-docked complexes were further refined. Accordingly, the topic of binding modes, binding interactions, and binding affinities has been extensively discussed. Finally, a QSAR analysis was implemented based on an improved version of PLS regression, leading to the development of a robust model. The model developed was interpreted based on the results of the literature review and the combination of the cited approaches.

Overall, this dissertation concludes with a general conclusion that summarizes the general workflow of the study, outlines the key findings, and specifies the hypotheses that will stimulate future experimental investigations and theoretical studies in the topic of benzodiazepines.

Bibliography

[1] S. OU-YANG, J. LU, X. KONG, Z. LIANG, C. LUO, H. JIANG, Computational drug discovery, Acta Pharmacol. Sin. 33 (2012) 1131–1140. https://doi.org/10.1038/aps.2012.109.

[2] F.D. Prieto-Martínez, E. López-López, K. Eurídice Juárez-Mercado, J.L. Medina-Franco, Computational Drug Design Methods—Current and Future Perspectives, in: Silico Drug Des., 2019: pp. 19–44. https://doi.org/10.1016/b978-0-12-816125-8.00002-x.

[3] S.P. Leelananda, S. Lindert, Computational methods in drug discovery, Beilstein J. Org. Chem. 12 (2016) 2694–2718. https://doi.org/10.3762/bjoc.12.267.

[4] J.A. DiMasi, H.G. Grabowski, R.W. Hansen, Innovation in the pharmaceutical industry: New estimates of R&D costs, J. Health Econ. 47 (2016) 20–33. https://doi.org/10.1016/j.jhealeco.2016.01.012.

[5] J.A. DiMasi, R.W. Hansen, H.G. Grabowski, The price of innovation: New estimates of drug development costs, J. Health Econ. 22 (2003) 151–185. https://doi.org/10.1016/S0167-6296(02)00126-1.

[6] C.P. Adams, V. Van Brantner, Estimating The Cost Of New Drug Development: Is It Really $802 Million?, Health Aff. 25 (2006) 420–428. https://doi.org/10.1377/hlthaff.25.2.420.

[7] C.P. ADAMS, V.V. BRANTNER, SPENDING ON NEW DRUG DEVELOPMENT, Health Econ. 19 (2010) 130–141. https://doi.org/10.1002/hec.1454.

[8] J.A. Dimasi, H.G. Grabowski, J. Vernon, R&D Costs and Returns by Therapeutic Category, Drug Inf. J. 38 (2004) 211–223. https://doi.org/doi:10.1177/009286150403800301.

[9] J.A. DiMasia, H.G. Grabowski, The Cost of Biopharmaceutical R&D: Is Biotech Different?, Manag. Decis. Econ. 28 (2007) 469–479. https://doi.org/10.1002/mde.1360.

[10] K. Jayasundara, A. Hollis, M. Krahn, M. Mamdani, J.S. Hoch, P. Grootendorst, Estimating the clinical cost of drug development for orphan versus non-orphan drugs, Orphanet J. Rare Dis. 14 (2019) 1–10. https://doi.org/10.1186/s13023-018-0990-4.

[11] S.M. Paul, D.S. Mytelka, C.T. Dunwiddie, C.C. Persinger, B.H. Munos, S.R. Lindborg, A.L. Schacht, How to improve RD productivity: The pharmaceutical industry's grand challenge, Nat. Rev. Drug Discov. 9 (2010) 203–214. https://doi.org/10.1038/nrd3078.

[12] O.J. Wouters, M. McKee, J. Luyten, Estimated Research and Development Investment Needed to Bring a New Medicine to Market, 2009-2018, JAMA - J. Am. Med. Assoc. 323 (2020) 844–853. https://doi.org/10.1001/jama.2020.1166.

[13] V. Prasad, S. Mailankody, Research and development spending to bring a single cancer drug to market and revenues after approval, JAMA Intern. Med. 177 (2017) 1569–1575. https://doi.org/10.1001/jamainternmed.2017.3601.

[14] C.H. Wong, K.W. Siah, A.W. Lo, Estimation of clinical trial success rates and related parameters, Biostatistics. 20 (2019) 273–286. https://doi.org/10.1093/biostatistics/kxx069.

[15] X. Lin, X. Li, X. Lin, A review on applications of computational methods in drug screening and design, Molecules. 25 (2020) 1–17. https://doi.org/10.3390/molecules25061375.

[16] S. Gupta, A. Basu, A. Jaiswal, N. Mishra, COMPUTATIONAL METHODS IN DRUG DISCOVERY, Int. J. Pharm. Siences Res. 9 (2018) 4102–4108. https://doi.org/10.13040/IJPSR.0975-8232.9(10).4102-08.

[17] G. Sliwoski, S. Kothiwale, J. Meiler, J. Edward W. Lowe, Computational methods in drug discovery, Pharmacol. Rev. 66 (2014) 334–395. https://doi.org/10.1124/pr.112.007336.

[18] A.C. Paweł Śledź, Protein structure-based drug design: From docking to molecular dynamics, Curr. Opin. Struct. Biol. 48 (2018) 93–102. https://doi.org/10.1016/j.sbi.2017.10.010.

[19] A. Tiwari, S. Singh, Computational approaches in drug designing, Bioinformatics. (2022) 207–217. https://doi.org/10.1016/b978-0-323-89775-4.00010-9.

[20] S. LaPointe, D. Weaver, A Review of Density Functional Theory Quantum Mechanics as Applied to Pharmaceutically Relevant Systems, Curr. Comput. Aided-Drug Des. 3 (2007) 290–296. https://doi.org/10.2174/157340907782799390.

[21] G. Sliwoski, S. Kothiwale, J. Meiler, J. Edward W. Lowe, Computational methods in drug discovery, Pharmacol. Rev. 66 (2014) 334–395. https://doi.org/10.1124/pr.112.007336.

[22] H. Tandon, T. Chakraborty, V. Suhag, A Brief Review on Importance of DFT In Drug Design, Res. Med. Eng. Sci. 7 (2019) 791–795. https://doi.org/10.31031/RMES.2019.07.00068.

[23] G. Bocci, E. Carosati, P. Vayer, A. Arrault, S. Lozano, G. Cruciani, ADME-Space: A new tool for medicinal chemists to explore ADME properties, Sci. Rep. 7 (2017) 25–27. https://doi.org/10.1038/s41598-017-06692-0.

[24] P.B. Reddy, M.B.M. Reddy, R. Reddy, S. Chhajed, P.P. Gupta, Molecular docking, PKPD, and assessment of toxicity of few chalcone analogues as EGFR inhibitor in search of anticancer agents, Struct. Chem. 31 (2020) 2249–2255. https://doi.org/10.1007/s11224-020-01571-3.

[25] S. Zhu, C.M. Noviello, J. Teng, R.M. Walsh, J.J. Kim, R.E. Hibbs, Structure of a human synaptic GABAA receptor, Nature. 559 (2018) 67–88.

https://doi.org/10.1038/s41586-018-0255-3.

[26] L. Richter, C. De Graaf, W. Sieghart, Z. Varagic, M. Mörzinger, I.J.P. De Esch, G.F. Ecker, M. Ernst, Diazepam-bound GABAA receptor models identify new benzodiazepine binding-site ligands, Nat. Chem. Biol. 8 (2012) 455–464. https://doi.org/10.1038/nchembio.917.

[27] K.R. Tan, A. Gonthier, R. Baur, M. Ernst, M. Goeldner, E. Sigel, Proximity-accelerated chemical coupling reaction in the benzodiazepine-binding site of γ-aminobutyric acid type A receptors: Superposition of different allosteric modulators, J. Biol. Chem. 282 (2007) 26316–26325. https://doi.org/10.1074/jbc.M702153200.

[28] S. Masiulis, R. Desai, T. Uchański, I. Serna Martin, D. Laverty, D. Karia, T. Malinauskas, J. Zivanov, E. Pardon, A. Kotecha, J. Steyaert, K.W. Miller, A.R. Aricescu, GABAA receptor signalling mechanisms revealed by structural pharmacology, Nature. 565 (2019) 454–459. https://doi.org/10.1038/s41586-018-0832-5.

[29] J.J. Kim, A. Gharpure, J. Teng, Y. Zhuang, R.J. Howard, S. Zhu, C.M. Noviello, R.M. Walsh, E. Lindahl, R.E. Hibbs, Shared structural mechanisms of general anaesthetics and benzodiazepines, Nature. 585 (2020) 303–308. https://doi.org/10.1038/s41586-020-2654-5.

Chapter I

Overview on Neurotransmission, GABA receptors, and Benzodiazepine.

1 Introduction

Anxiety disorder, bipolar disorder, schizophrenia, depression, autism spectrum disorder, and attention-deficit/hyperactivity disorder are among the growing psychiatric disorders worldwide. Throughout the years, several psychiatric disorders' symptoms, characteristics, and categories have been extensively documented in the literature. The pathophysiology of psychiatric disorders is influenced by many factors, including genetics, environment (such as infections, early traumas, and drugs), aging, and sociodemographics (e.g., ethnicity and socioeconomic status). Although the pathophysiology of psychiatric diseases is still poorly understood, disruptions in monoaminergic, disruptions in neurotransmitter systems (e.g., glutamatergic, purinergic, and GABAergic), and alterations in extremely complex and interconnected metabolic pathways have been reported as etiologies of many psychiatric diseases [1].

GABA, an inhibitory neurotransmitter, is estimated to be 1000 times more abundant in the brain than monoamine neurotransmitters. GABAergic inhibition deficiency underlies a large number of human disease states, including anxiety and stress disorders, musculoskeletal and pain disorders, insomnia and sleep disorders, addiction and drug-withdrawal syndromes, epilepsy and seizures, anesthesia, liver diseases and hepatic encephalopathy, cognition, learning and memory disorders, and hormonal disorders [2]. By strengthening synaptic activity, GABA regulates the functioning of different interconnected neuronal cells in the adult brain. Hence, the strength and polarity of GABAergic transmission are always modulated in both physiological and pathological states. The strength regulation of GABAergic transmission via $GABA_A$ receptors is accomplished by a variety of processes, including direct modulation of the $GABA_A$ receptors, adjustment of intracellular chloride concentration, and alteration of GABA metabolism [3]. Benzodiazepines are considered safe drugs for modulating GABAergic transmission during physiological and pathological conditions [4].

This chapter begins by collecting basic background information that will help readers better understand neurotransmission in the brain. Following that, it focuses on neurotransmitter receptors, starting with a brief description of their two main super-families: G-protein-coupled receptors and ligand-gated ion channel receptors, to a detailed description of the γ-aminobutyric acid receptors, in particular the type ‘’A’’ γ-aminobutyric acid receptor ($GABA_A$). Also, the GABA shunt (or GABA metabolism) and the orthosteric GABA binding sites in $\alpha_1\beta_2\gamma_2$ subtypes are well documented. Finally, the chemistry, pharmacological properties, and physiological effects of benzodiazepines were reviewed.

2 Neuronal transmission in the brain

The rapid transmissions of information inside the brain are ensured through communication between networks of nerve cells, known as neurons. Neurons have distinct features that set them apart from other cells in the body. Their morphology is tree-like, consisting of a cell body (soma) and specialized thin branches called dendrites and axons that allow them to perform their specific function in the central nervous system. Indeed, neurons receive signal inputs through dendrites and share them, with each other, via axons. This strategy is known as ‘’the interneuronal communication’’ (Figure I.1). The pre-synaptic neurons send electrical stimulations known as the "action potential" through their axons until it achieves the axon terminal where they enhance the concentration of calcium ions, which in turn, stimulates the vesicles to release chemical inputs called "neurotransmitters" into the synaptic gap. Meanwhile, the post-synaptic neurons receive the neurotransmitters via their dendrites where they bind to specific receptors and mediate the conduction of signals to the rest members of the neural network (Figure I.2). Chemical transmission is more common in the mammalian central nervous system than electrical transmission. The latter is produced by proteins known as ‘’connexins” that form ion-conducting pores that connect the intracellular compartments of neighboring neurons and enable direct ion passage from one cell to another [5,6].

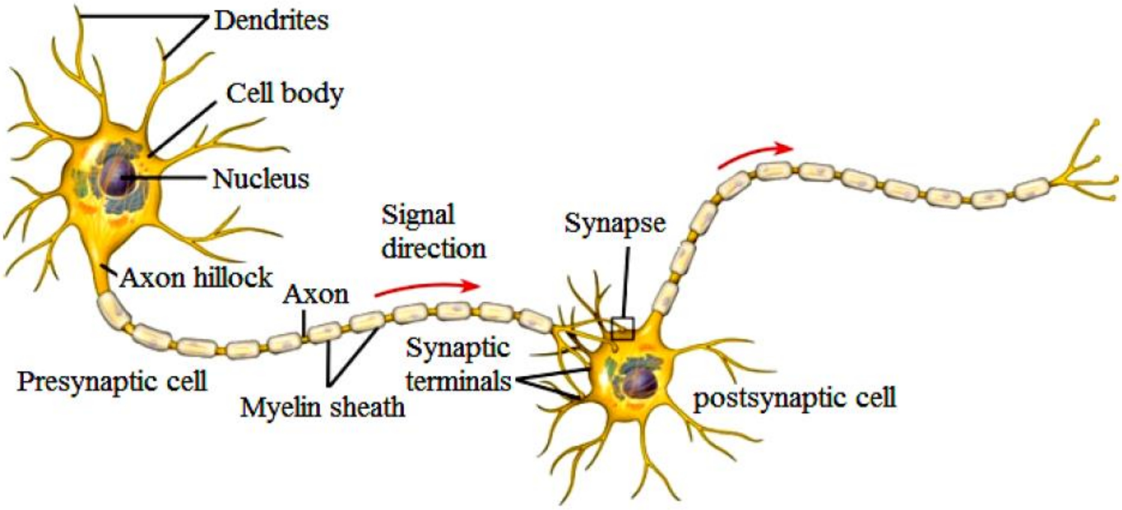


Figure I.1. Structure of a typical neuron [7].

The neuron must be excitable to perform its function, and this is achieved by the presence of an electrical voltage flow across the cell membrane. The electrical voltage results from the difference in concentration gradients of ions between the inner and outer medium of the neuron's membrane. In the resting state of the neuron, there is the so-called "resting membrane potential", which represents the ground on which the action potential of the nerve originates. The typical voltage value of the resting membrane potential is approximately -70 millivolts (mV) (numbered 1 in Figure I.3). This value maintains constant through the voltage-gated sodium/potassium pumps distributed, principally, across the soma and axon membranes. Potassium is the predominant intracellular ion and sodium is the predominant extracellular ion. Logically, both ions tend to flow through the pumps from the more concentrated to the less concentrated medium. The influx of positive ions out of the neuron generates negative membrane potential or hyperpolarization, and vice versa, the influx of positive ions into the neuron leads to depolarization [5].

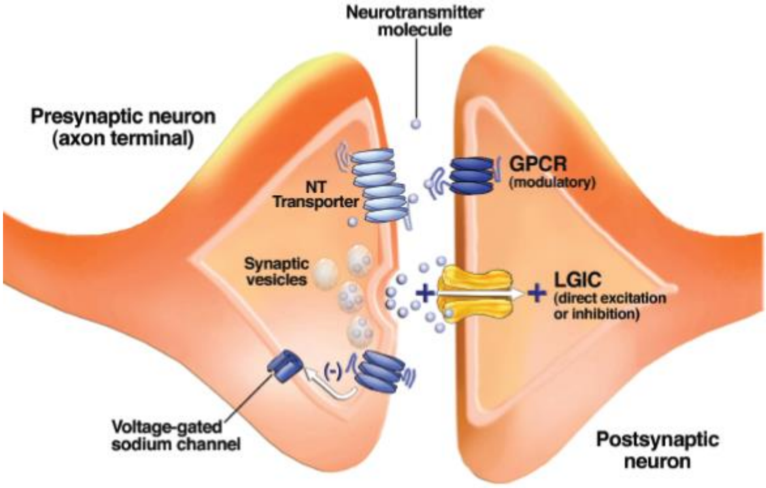


Figure I.2. Schematic representing a synapse between two neurons [5].

Upon stimulation, neurons receive signal inputs through dendrites and spread them via the soma. Excitatory signals activate the ligand-gated sodium channels and allow sodium to enter the cell. Thus, depolarizing the cell membrane. The influx of sodium within the neuron produces a current that travels to a trigger zone within the cell known as the "hillock". In hillock, if the input signal is excitatory and sufficiently strong, an action potential is generated by the voltage-gated ion channels and transmitted down the axon to the axon terminal. The voltage-gated ion channels are controlled by the membrane voltage. They activate at specific values of the membrane potential and deactivate at others. The threshold required to open the voltage-gated ion channels is about -55 mV (numbered 2 in Figure I.3). At this value, sodium-gated channels open at a faster rate than the potassium-gated channels. When sodium ions enter the cell, the inside becomes more positive than the outside, further depolarizing the cell membrane. Increasing the membrane voltage gradually activates all sodium channels and increases the rate of diffusion of ions into the cell. This phase corresponds to the rising state of the action potential. As the action potential reaches its peak (numbered 3 in Figure I.3), the falling phase begins with the progressive closing of sodium channels and the gradual opening of all potassium channels dispersed over the membrane. As a result, potassium ions exit the cell, and the voltage quickly returns to its resting value. Due to the slow closing of potassium channels, potassium remains in the cell for a bit longer, resulting in a negative overshoot known as hyper-polarization (numbered 4 in Figure I.3). The latter is gradually restored to the resting value via sodium/potassium pumps which return both ions to their respective positions through the membrane while awaiting a new stimulation [8,9].

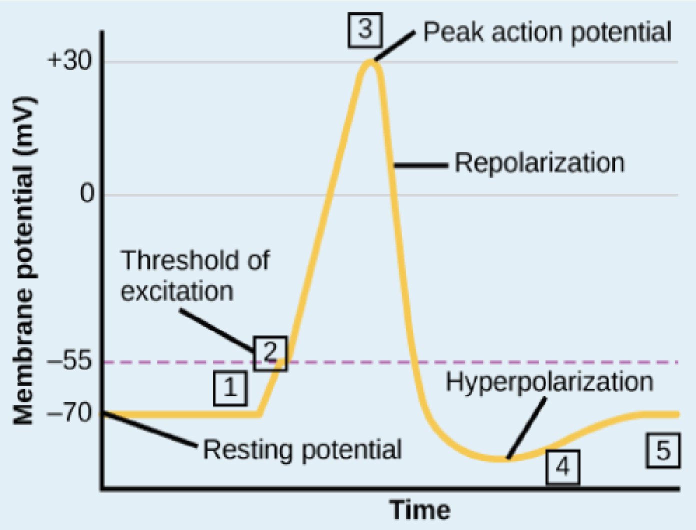


Figure I.3. Neuronal action potential diagram [5].

3 Neurotransmitter receptors

Neuroreceptors include two main super-families: the G-protein-coupled receptors (GPCRs) and the ligand-gated ion channel receptors (LGICRs) [5].

3.1 G-protein-coupled receptors

GPCRs are metabotropic receptors that are responsible for slow synaptic transmission. They are a unique superfamily of membrane proteins that mediate the majority of cellular events involving neurotransmitters, hormones, ions, small organic molecules, peptides, and photons, as well as having an impact on a wide range of biological processes including homeostasis, cell proliferation and migration, olfaction, taste, vision, growth, and mood [10–12]. Its overall architecture is defined by an extracellular N-terminus followed by seven transmembrane α-helices spanning the cell membrane seven times (Figure I.4). Each α-helix is linked to the next via an intracellular or extracellular loop. In the end, the intracellular C-terminus was connected to a heterotrimeric structure of three subunits arranged as follows: α, β, and then γ. Both α and γ subunits are attached to the cell membrane by lipidic anchors. This arrangement refers to the G-protein [10]. Upon activation through the ligand, the GPCR undergoes transmembrane and intracellular conformational changes. Thus it stimulates the exchange of GDP to GTP. Next, most commonly, α subunit completely detaches from the βγ dimer and heads to regulate the function of the target membrane protein (enzymes, ion channels…). As long as the ligand is bound to the extracellular interface, the α subunit remains searching for a target and repeatedly initiates a series of processes. At the end of the process, the body resorts to using a G-protein signaling regulator (RGS) that promotes hydrolysis of GTP to GDP, re-associates the α subunit to the βγ dimer, dissociates the ligand, and returns the GPCR to the inactive state [11].

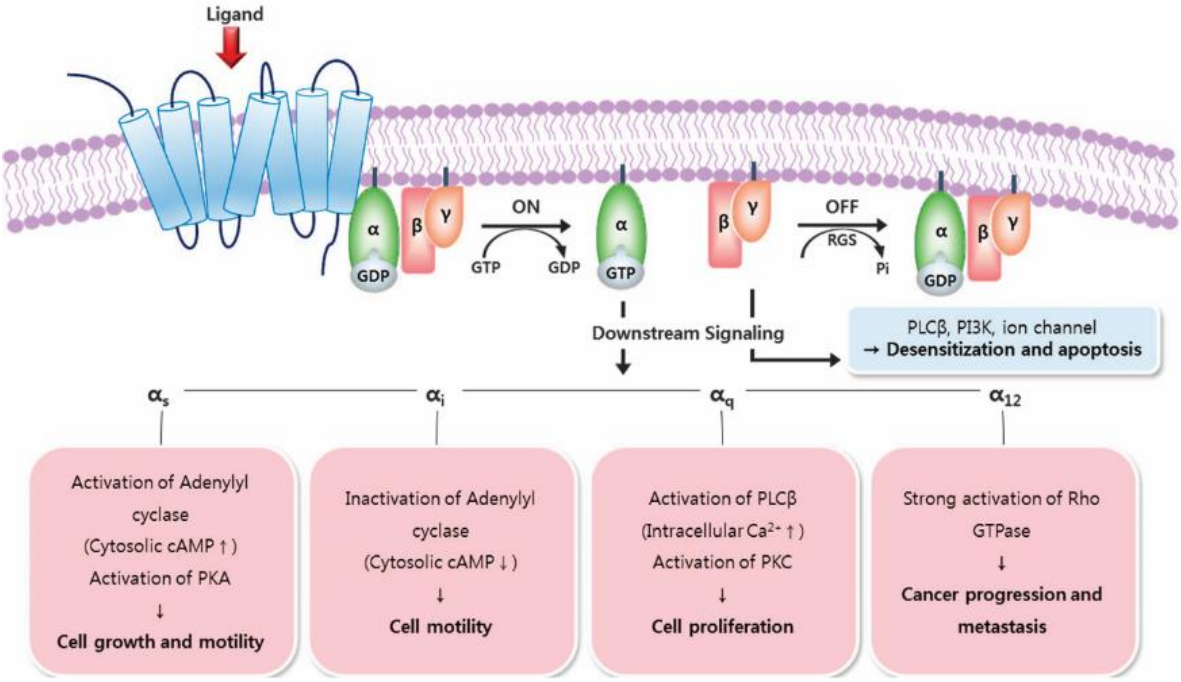


Figure I.4. Schematic diagram of GPCR signaling mediated by the activation of the G-protein α subunit [11].

In vertebrates, GPCRs include five families classified based on their sequence, structural, and functional similarities: rhodopsin, secretin, glutamate, adhesion, and Frizzled/Taste2 [10,11]. Despite GPCRs being required for many key cellular physiological activities, many of the structures and functions of the members of these families remain unknown due to the delaying of the GPCRs crystallization after isolation from the native membrane. Indeed, the significant advances in studies started in the year 2000 with the reporting of the first crystal structure of bovine rhodopsin [13].

3.2 Ligand-gated ion channel receptors

LGICs are ionotropic receptors that are responsible for the fast transmission of intercellular signaling. They are transmembrane integral proteins that open and close in response to the binding of an endogenous agonist like a neurotransmitter. Its morphology consists of multimeric subunits surrounding a selective ion channel that extends across the post-synaptic membrane (Figure I.5 (a)) [14]. When a neurotransmitter binds at the orthosteric binding site in the extracellular domain interface, a conformational change occurs, enabling the ion channel to open and therefore recording the inflow of Na^+, K^+, Ca^{2+}, or Cl^- ions into the post-synaptic cell on a rapid response (millisecond time scale). Furthermore, modulation of LGICRs can occur through endogenous/exogenous modulators. The latter stimulates the excitatory/inhibitory balance in the central nervous system by binding at allosteric binding sites that differ from the orthosteric binding sites of neurotransmitters [5,14]. LGICs are thus attractive candidates for new treatments. In vertebrates, genome and cDNA sequence analyses have divided the LGICs superfamily into three unrelated sub-families. Each sub-family is characterized by a specific architecture as follows (Figure I.5 (b)):

- ✓ The Cys-Loop receptors: assemble into pentameric isoform arranged around an ion channel gated by acetylcholine (ACh), γ-aminobutyric acid (GABA), glycine (Gly), and serotonin or 5-hydroxytryptamine (5-HT).
- ✓ The ionotropic glutamate receptors (iGluRs): assemble into tetrameric isoform arranged around nonselective cation channel (Na^+, K^+, Ca^{2+}) gated by glutamate.
- ✓ The P2X receptors (P2XRs): assemble into trimeric isoform arranged around ion channel gated by ATP [14].

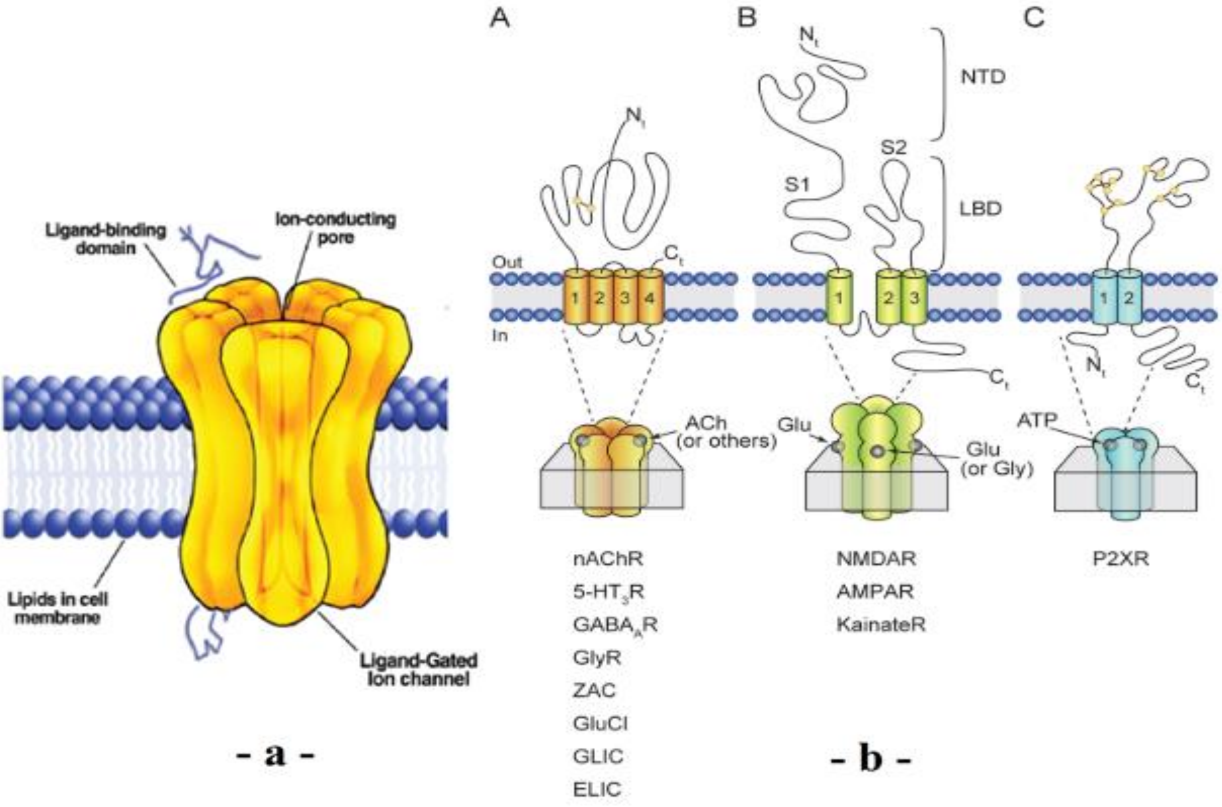


Figure I.5. a: structure of ligand-gated ion channel [5], **b:** subunit assembly of the three families of LGICs: (A) Cys-loop receptors, (B) iGluRs, and (C) P2X receptors [14].

3.2.1 Cys-Loop family

All members of the Cys-Loop family share homologous structures, with a pentameric isoform generated by five identical or related subunits arranged around a central ion channel (see ‘’A’’ in Figure I.5 (b)). Each subunit is divided into three parts: the large extracellular domain (ECD) composed of long hydrophilic N-terminal α-helix followed by β-strands folded into a β-sandwich containing the Cys-Loops that is enclosed by a disulfide bridge. The transmembrane domain (TMD) is formed by four membrane α-helices (M_1, M_2, M_3, and M_4) connected by three loops (short intracellular loop links M_1-M_2, short extracellular loop links M_2-M_3, and long cytoplasmic loop links M_3-M_4) and terminating with one small extracellular C-terminal. The five M_2 α-helices were assembled in the center to form the ion-channel liner. The intracellular domain (ICD) is mainly formed by the long M_3-M_4 loop (100–270 residues) [15]. The Cys-Loop family contains, in addition to the neurotransmitters orthosteric binding sites, the allosteric binding sites of many therapeutic agents such as benzodiazepines, anesthetics, alcohols, steroids, muscle relaxants, and a range of drugs that treat neurological disorders such as Alzheimer’s, epilepsy, anxiety, learning, attention deficit, and drug addiction [14]. The Cys-Loop family is the most abundant type of LGICs. This family contains the nicotinic acetylcholine receptors (nAChRs), serotonine receptors (5-HT3Rs), zinc-activated channel (ZAC), γ-aminobutyric acid ($GABA_A$/$GABA_C$), glycine receptors (GlyRs), glutamate-gated chloride channels (GluCl), and the prokaryotes proteins (e.g. ELIC and GLIC). Nicotinic acetylcholine receptors (nAChRs), 5-HT3 receptors, and zinc-activated channel (ZAC) mediate in fast excitatory neurotransmission while the gamma-aminobutyric acid ($GABA_A$/$GABA_C$) glycine receptors (GlyRs), and glutamate-gated chloride channels (GluCl) are mediated in fast inhibitory [14,16]. In this dissertation, we focused primarily to describe the γ-aminobutyric acid type ‘’A’’ receptor ($GABA_AR$).

4 γ-aminobutyric acid receptors

γ-aminobutyric acid receptors (GABARs) are types of membrane-bound receptors that respond to the inhibitory effect of GABA. This latter inhibits neuronal excitability by acting on the three known classes of GABARs: $GABA_A$, $GABA_B$, and $GABA_C$ [17]. As reported earlier, $GABA_A$ and $GABA_C$ are ionotropic receptors belonging to the LGICRs superfamily and they share the same overall architecture as all the members of the Cys-Loop sub-family. They are selective receptors for chlorine ions. When activated through GABA, they allow negative chloride ions to flow through the ion channel into the cell, thereby inhibiting neuronal excitability for a short time (phasic inhibition) or for a long time (tonic inhibition). Some studies consider the $GABA_CR$ to be one of the several isoforms of the $GABA_AR$

[18,19]. However, due to the several fundamental differences that were discovered between their function, structure, and pharmacology, starting in 2008, further use of this terminology is discouraged by the International Union of Basic and Clinical Pharmacology [20]. The main properties that differentiate $GABA_C$Rs from $GABA_A$Rs are shown in Table I.1. $GABA_B$, on the other hand, is a metabotropic receptor belonging to the glutamate family of the GPCRs superfamily which makes it significantly different from $GABA_A$ and $GABA_C$ receptors in structure, sequence, and function [17,19]. Its structure consists of two heterosubunits R1 and R2 conjugated to the pertussis toxin–sensitive G-protein Gαi/o class. It exerts its inhibitory role by stimulating Gαi/o dissociation into Gα and Gβγ. Gαi/o activates and inhibits adenylyl cyclase, while Gβγ dimer modulates Potassium/Calcium ions flow by activating and deactivating, respectively, the voltage-gated potassium ion channels and the voltage-gated calcium ion channels [21,22].

Table I.1. Differences between $GABA_A$ and $GABA_C$ receptors [23].

	$GABA_A$ receptors	**$GABA_C$ receptors**
Channel type	Cl^- channel	
Conductance	27-30 pS	7-8 pS
Mean channel open time	25-30 ms	150-200 ms
GABA concentrations	10-100 μM	1 μM
Channel composition	Heterooligomeric	Homooligomeric
Subunit composition	$\alpha_{1-6}, \beta_{1-4}, \gamma_{1-3}, \delta, \varepsilon, \pi, \theta,$	ρ_{1-3}
Chromosomal location of coding genes	Chromosomes 1, 4, 5, 15, and X	Chromosomes 3 and 6
Selective agonist	Not known	(+)-CAMP
Selective antagonist	Bicuculline	TPMPA
Potency order of common agonists	Muscimol > GABA > TACA	TACA > GABA > Muscimol
Modulators	benzodiazepines, barbiturates, and steroids	Not known
Anchoring protein	GABARAP	MAP-1B

TACA: trans-4-aminocrotonic acid, GABARAP: GABA Type A Receptor-Associated Protein, (+)-CAMP: 1*S*,2*R*-2-(aminomethyl)cyclopropanecarboxylic acid, TPMPA: 1,2,5,6-tetrahydropyridin-4-ylphosphinic acid, MAP-1B: micro-tubule-associated protein 1B.

4.1 Neurotransmitter γ-aminobutyric acid

Neurotransmitter γ-aminobutyric acid (GABA) is the main inhibitory neurotransmitter in the brain of vertebrates and invertebrates. Its inhibitory role exerts through the binding to GABA receptors located in the cellular membrane of pre-synaptic and post-synaptic neurons. GABA neurotransmission disruption leads to a variety of neurodegenerative diseases, notably epilepsy, anxiety disorder, Parkinson's disease, and Huntington's chorea [23].

4.1.1 GABA metabolism or GABA shunt

The GABA shunt refers to a closed-loop mechanism that produces GABA while also preserving its supply. GABA biosynthesis occurs only in *Gaba*ergic neurons, using glutamate as a precursor [24]. Thus, we will first briefly explain the biosynthesis of glutamate, and then we will discuss the biosynthesis of GABA from the latter. Figure I.6 depicted the metabolic cooperation between glutamatergic/*Gaba*ergic neurons and glial cells ‘’astrocytes’’. The detailed mechanism is illustrated in Figure I.7.

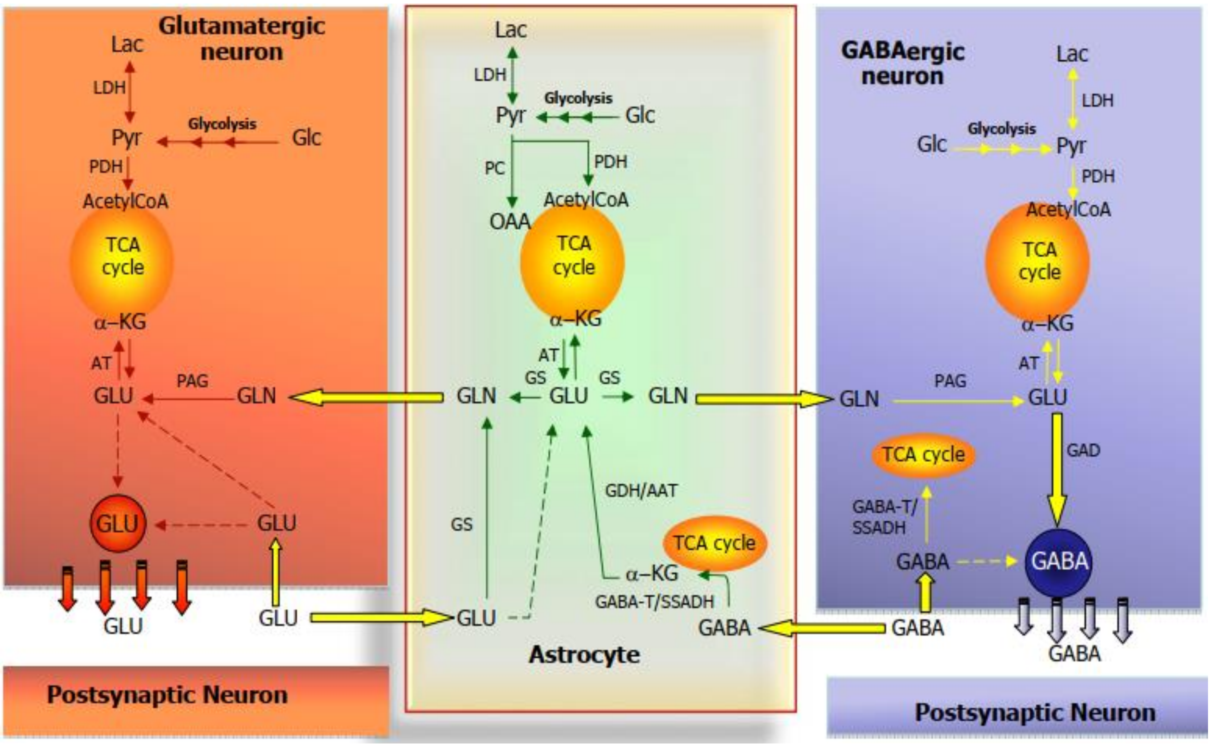


Figure I.6. The metabolic cooperation between glutamatergic neurons, *Gabaergic* neurons, and glial cells ‘’astrocytes’’. Abbreviations: LAC: lactate, LDH: lactate dehydrogenase, Pyr: pyruvate, Glc: glucose, PDH: pyruvate dehydrogenase, α-KG: α-ketoglutarate, AT: aminotransferases, GLU: glutamate, PAG: phosphate-activated glutaminase, GLN: glutamine, PC: pyruvate carboxylase, OAA: oxaloacetate, GS: glutamine synthetase, GDH:glutamate dehydrogenase, AAT: aspartate aminotransferase, GABA-T: GABA-transaminase, SSADH: succinic semialdehyde dehydrogenase, GAD: glutamate decarboxylase [25].

Glutamate biosynthesis is directly linked to glucose metabolism. The blood-brain barrier (BBB) isolates the brain from the vascular system, and while glucose and essential amino acids can penetrate the BBB via specialized transporters, the non-essential amino acids glutamate and its metabolite glutamine cannot. As a result, glutamate is formed in the brain from glucose. This process occurs through cytoplasmic glycolysis which is succeeded by the mitochondrial tricarboxylic acid (TCA) cycle. In this pathway, the key precursor to glutamate, α-ketoglutarate, is manufactured. Then, through transamination, α-ketoglutarate is transformed into L-glutamate. Subsequently, GABA biosynthesis occurs in the presynaptic cytoplasm by enhancing the decarboxylation of glutamate through the rate-limiting enzyme, L-glutamic acid decarboxylase (GAD), in association with the pyridoxine (vitamin B6) as a co-factor [24,25].

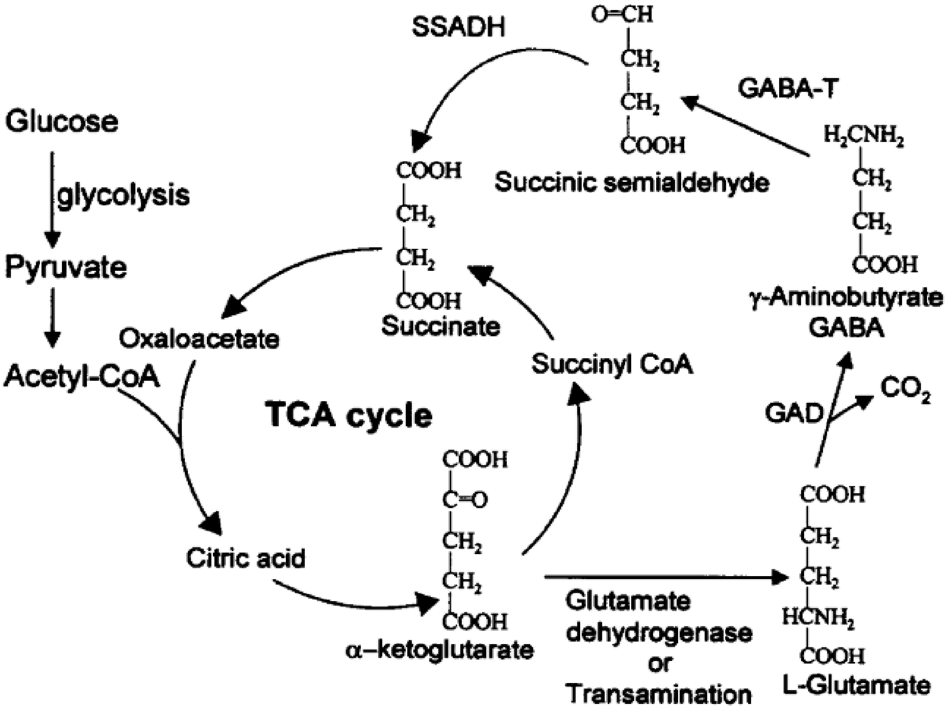


Figure I.7. General pathway of GABA and glutamate biosynthesis and degradation to succinate. GAD: L-glutamic acid decarboxylase, GABA-T: GABA transaminase, SSADH: succinic semialdehyde dehydrogenase [24].

The synthesized GABA is transported by the vesicular inhibitory amino acid transporter (vGAT) to the synaptic vesicles where it's stored until neurostimulation occurs. When the action potential attends the axon terminal, the voltage-gated calcium ion channels are activated and the calcium binds to synaptobrevin which promotes synaptic vesicles to fuse with the cell membrane and release their content into the synaptic gap (Figure I.8). Here, GABA mainly tends to diffuse into the post-synaptic cell to perform its inhibitory action by activating its target receptors [26].

When the neuronal transmission is completed, GABA is removed from the synaptic space by the GABA transporters (GAT) and transported into the presynaptic neurons and neighboring glial cells, where it's either uptake by the presynaptic cells and returned to the vesicles or catabolized to glutamine in astrocytes [26]. GABA catabolism occurs through a sequential process that begins with the conversion of GABA to succinic semialdehyde (SSA) in the presence of GABA transaminase (GABA-T). Then, the succinic semialdehyde dehydrogenase (SSADH) quickly oxidizes SSA to succinate, which enters the TCA cycle and is involved in the production of glutamine. GABA can eventually be regenerated from this latter [24,25].

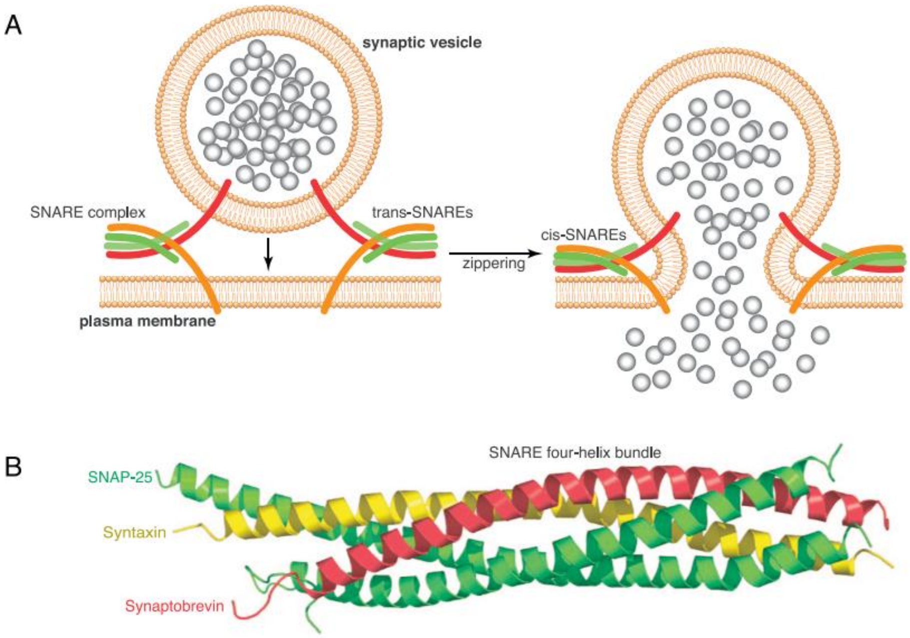


Figure I.8. A, Fusion of synaptic vesicles. **B**, SNARE structure: assembled from four coiledcoil domains provided by three different proteins situated on the vesicular and the cellular membranes: synaptobrevin, synataxin, and two SNAP-25 [27].

4.2 Type A γ-aminobutyric acid receptor

This type of GABA receptors mediates the majority of GABA's physiological functions. It is found in 20–50% of brain synapses [28]. It was distributed mainly on the dendrites and the soma of *Gaba*ergic neurons. Also, it localized with lower proportions at the presynaptic, perisynaptic, and extrasynaptic sites [29]. Its structure displays a high level of molecular diversity, owing to the heterogeneity of its subunits and the numerous ways in which they can be joined together to form heteropentameric isoforms [18]. Therefore, $GABA_AR$ contains large numbers of allosteric binding loci that make it among the most important drug targets in the central nervous system (CNS).

4.2.1 Overall architecture

As can be seen from Figure I.9 (a) and (b), $GABA_A$ receptors are characterized by a pentameric isoform generated by five identical or related subunits assembled around a central chlorine-ion channel. The general architecture of each mature subunit is constructed by the sequence of 450 amino acid residues in length [19]. 200–250 amino acids contribute to the ECD to form the hydrophilic N-terminal α-helix, the ten β-strands, and the Cys-loops. 85–255 amino acids contribute to TMD to form four membrane α-helices (named from M_1 to M_4), three loops, and the extracellular C-terminal [30]. The chloride-channel liner was constricted by the five M_2 α-helices, with a possible contribution from M_1 [20]. The components of each matter subunit have been modeled and given separately in part c of Figure I.9.

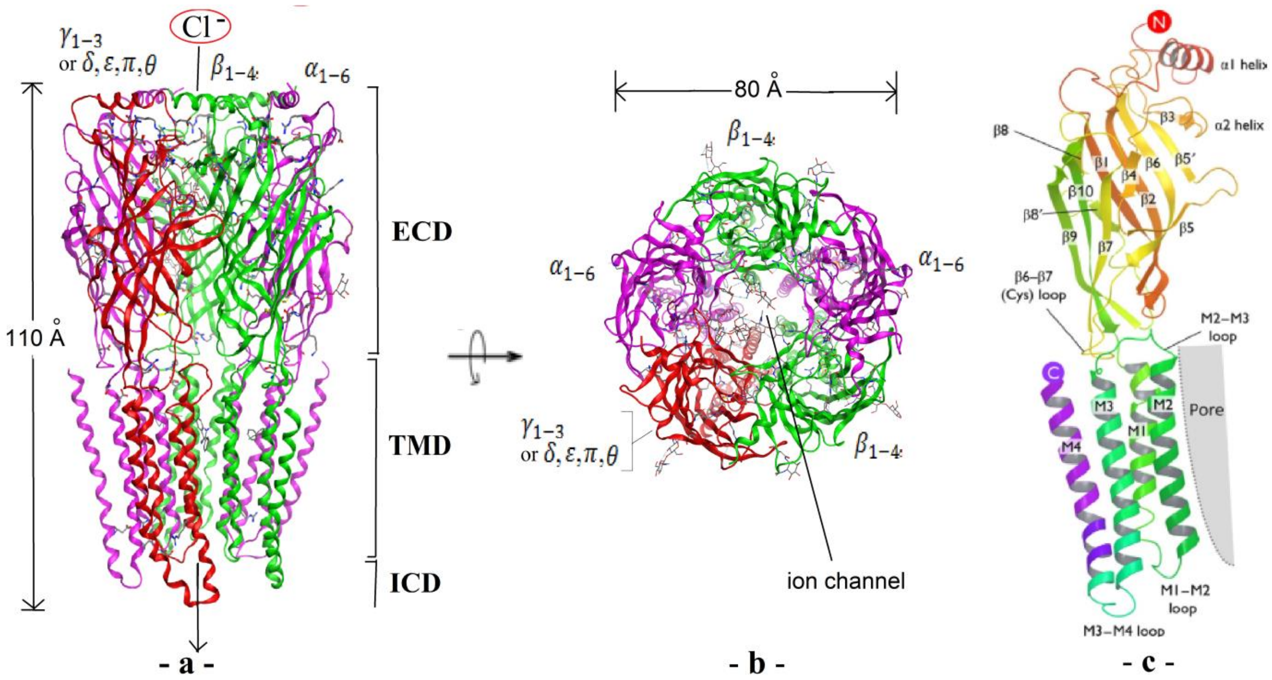


Figure I.9. Overall structure of GABA$_A$ receptor. **a, b,** side and top views representing the receptor configurations: the possible subunit arrangements, the extracellular domain (ECD), the transmembrane domain (TMD), the intracellular domain (ICD), the ion channel, and the direction of chlorine flow into the cell, adapted from PDB ID: 6x3x, using Moe 2014.0901. **c.** a single subunit structure, taken from [31].

4.2.2 Subunits

In the mammalian brain, GABA$_A$Rs are assembled from 8 different families of subunits containing a total of twenty subtypes: α_{1-6}, β_{1-4}, γ_{1-3}, δ, ε, π, θ, and ρ_{1-3} [29]. The existing GABA$_A$ receptor subunits and their sequence homologies are shown in Figure I.10 (a), and the genomic location of each subunit in human chromosomes is given in Table I.2.

Between subunit classes, there is roughly 30% sequence identity, and between subunit subtypes, there is approximately 70% sequence identity [23]. Furthermore, the alternative splicing of some genes has resulted in additional diversity in the sequence of subunits. For the subunits β_2, β_4, and γ_2, two types of splicing are observed, which are distinct from one another by the existence or lack of a short peptide in the long M_3-M_4 intracellular loop. For the β_3 subunit, it observed two distinct alternative forms as a result of the exon 1 splicing. Another example was found in the rat brain of the α_6 subunit, which cleaved around 20% of its transcripts, resulting in a loss of 10 amino acid residues at the N-terminus of unknown functional significance [29].

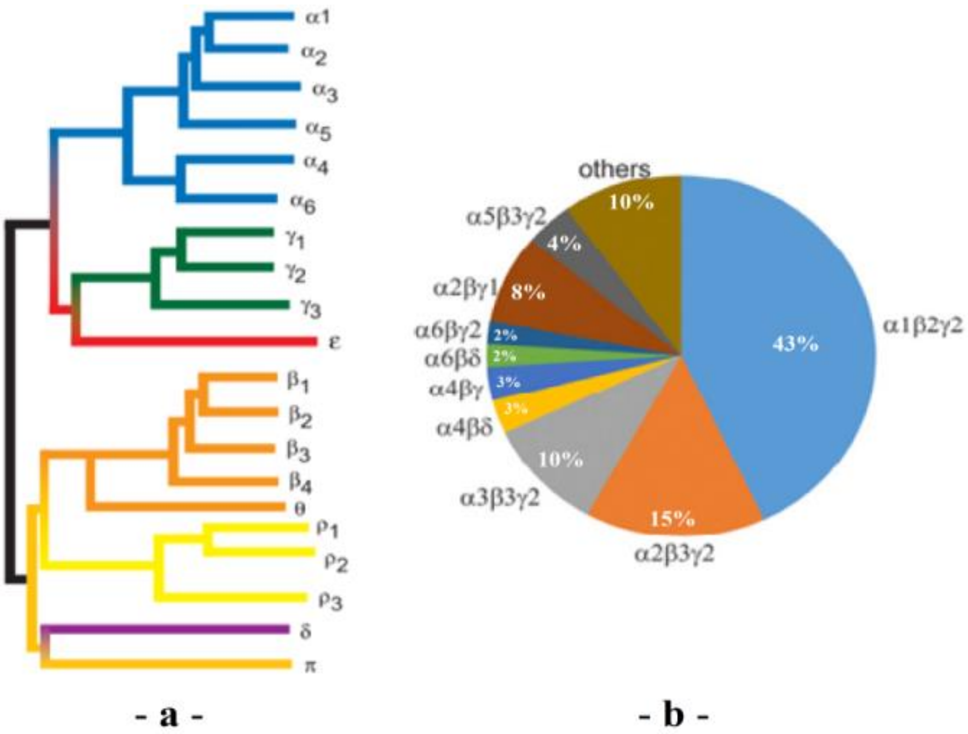


Figure I.10. a, Dendrogram representing the existing GABA$_A$ receptor subunits and their sequence homologies [32]. **b,** The estimated abundance of GABA$_A$R isoforms in the rat brain [18].

Table I.2. the genomic location of each subunit in human chromosomes [20].

Subunit	Gene	Human chromosome
α_1	GABRA1	5q34
α_2	GABRA2	4p12
α_3	GABRA3	Xq28
α_4	GABRA4	4p12
α_5	GABRA5	15q13.2
α_6	GABRA6	5q34
β_1	GABRB1	4p12
β_2	GABRB2	5q34
β_3	GABRB3	15q13.2
γ_1	GABRG1	4p12
γ_2	GABRG2	5q34
γ_3	GABRG3	15q13.2
δ	GARBD	1p36.3
ε	GABRE	Xq28
θ	GABRQ	Xq28
π	GABRP	5q35.1
ρ_1	GABRR1	6q15
ρ_2	GABRR2	6q15
ρ_3	GABRR3	3q12.1

The predominant $GABA_AR$ isoforms consist of two α-subunits, two β-subunits, and one γ-subunit, or one δ-subunit (Figure I.10 (b)) assembled in the alternating order: one α, one β, one α, one β, interconnected by one γ or δ subunit [18]. The three subunits δ, ε, and π can substitute the γ subunit in the given order. Meanwhile, the θ subunit can substitute the β subunit. In contrast, the ρ subunit is incompatible with the other subunits classes. It preferentially aggregates only with subunits of the same class to form homo- or hetero-oligomeric channels that define the structures and characteristics of $GABA_CR$ [33]. γ-containing receptors distribute mainly in the synaptic sites and account for approximately 90% of $GABA_ARs$ in the adult brain. In contrast, δ-containing are abundant in the extrasynaptic sites located in specific brain areas such as the hippocampus, amygdala, neocortex, thalamus, hypothalamus, and cerebellum. The $\alpha_1\beta_2\gamma_2$ combination constitutes approximately 43–60% of $GABA_ARs$ in the adult brain [18], which makes it the subject of many previous types of research attempting to provide high-resolution structural data that illuminate atomic mechanisms of drug recognition. The most important among them are the cryo-electron microscopy structures complemented with Mutagenesis, Electrophysiology, and molecular dynamics simulations proposed by *J.J. Kim and co-workers* (PDB ID: 6x3x) [28,34].

4.2.2.1 Regional distribution in CNS

The use of in situ hybridization and immunohistochemical experiments allowed researchers to precisely determine the density of distribution and abundance of $GABA_A$ subunits in the brain. Table I.3 summarizes the acquired results in brief. The individual subunits are distributed in diverse but overlapping regional and cellular patterns. The α_1, β_1, β_2, β_3, and $\gamma_2$subunits are dispersed across the brain with minimal variances, however, the α_2, α_3, α_4, α_5, α_6, γ_1, and δ subunits are more limited to specific brain regions. The most common and extensively distributed subunits in the brain are the α_1 and α_2. Most hypothalamic regions show colocalization of $\alpha_1$and α_2 with β_2 and β_3 subunits, indicating the possibility of preferential generation of $\alpha_{1/2}\beta_{2/3}$ receptors. Regarding the β and γ subunits, γ_2 and β_2 are the most abundant across all the CNS, whilst, β_1, β_3, γ_1, and γ_3 are relatively limited. The γ_1 is the least prevalent of the γ subunits and is exclusively found in the nervous system, whereas the γ_3 is expressed in the majority of brain areas [29,32,33].

Table I.3. Regional distribution of $GABA_A$ receptor subunits in the brain [33].

Region	α_1	α_2	α_3	α_4	α_5	α_6	β_1	β_2	β_3	γ_1	γ_2	γ_3	δ	ε	θ
Olfactory bulb glomerular															
Glomerular layer	xx	x	xx	o	x	-	-	xx	xx	-	xx	o	x	-	
Ext. plexiform layer	xxx	x	xx	o	x	-	xx	xxx	xxx	-	xxx	o	o	-	
Granular layer	xx	xx	o	x	xx	-	-	x	xx	-	xx	-	o	-	
Mitral cell layer	xx	-	o	-	xx	-	x	xx	-	-	xx	o	x	-	
Olfactory tubercle	x	x	-	xx	x	-	x	x	xx	-	x	x	x	-	
Cerebral cortex															
All layers	xx	x	x	xx	x	-	xx	xx	xx	-	xx	o	x		
Outer layers	xx	x	x	x	x	-	xx	xx	xx	-	xx	o	x	-	
Inner layers	xx	x	xx	x	x	-	xx	xx	xx	-	xx	o	x	-	
Hippocampus															
Molecular layer	x	xx	-	xx	x	-	xx	x	xx	-	xx	-	x	-	
Hilar neurons	xx	-	x	-	-	-	o	xx	-	-	xx	-	x	-	
Strat. oriens/radiatum	xx	xx	-	x	xx	-	xx	x	xx	-	xx	o	-	-	
Septum															
Medial	xx	x	x	-	o	-	o	xx	x	-	xx	o	-	x	x
Lateral	xx	xx	x	x	o	-	xx	x	x	x	xx	o	-		
Basal ganglia															
Striatum/n.accumbens	x	xxx	x	xx	xx	-	x	x	xxx	x	x	o	x	x	x
Globus pallidus	xx	o	o	x	o	-	o	xx	o	xx	xx	o	o		
Subst. nigra	x	x	x	o	x	-	x	x	-	x	x	x	o	xx	x
Thalamus															
Reticular nucleus	x	-	xx	x	x	-	xx	-	xx	-	xx	o	o		
Ventr. lat. geniculate	xxx	x	x	x	o	-	x	xx	x	-	x	o	o	xx	x
Dors. lat. geniculate	xxx	-	-	xxx	o	-	x	xxx	x	-	x	x	xxx		
Medial and central	x	xx	x	o	o	-	xx	xx	xx	xx	x	x	x	xx	x

Table I.3. Continued

Hypothalamus															
Ventromedial	x	xx	x	o	xx	-	xx	x	xx	x	xx	x	x	xx	xx
Supraopticus	xxx	xxx	x	x	o	-	xxx	xx	x	-	x	x	x		
Paraventricular	xx	xxx	-	-	x	-	xx	x	xx	-	x	x	x	xx	x
Arcuate	x	x	x	x	x	-	x	x	x	-	-	x	x	xx	x
Med. preoptic area	xx	xx	x	-	x	-	x	x	x	-	xx	x	x	xx	x
Amygdala															
Lateral	xx	xx	xx	x	o	-	xx	xx	xx	-	xx	x	o	x	x
Basolateral	xx	xx	xx	x	o	-	xx	xx	xx	-	xx	x	o		
Medial and central	x	xx	x	o	o	-	xx	xx	xx	xx	xx	x	x	x	x
Cerebellum															
Granule cell layer	xxx	x	o	o	x	xxx	x	xxx	xxx	x	xx	-	xxx		
Molecular layer	xx	xx	-	-	xx	-	x	x	-	-	x	o	-		
Midbrain/Pons															
Ventral tegmental area	xx	x	o	-	o	-	xx	x	xx	x	xx	xx	o	-	-
Raphe nuclei	xx	xx	x	-	o	-	xx	xx	x	-	xx	xx	xx	x	x
Inferior colliculus	xx	-	-	-	o	-	x	xx	o	-	o	o	o		
Olive superior	o	-	x	-	o	-	xx	-	x	-	x	o	x		
Medulla															
Trigeminal sensory complex	xx	-	xx	o	o	-	x	x	x	x	xx	x	x		
Dorsal cochlear nucleus	xx	x	xx	o	x	xx	o	x	o	o	x	x	xx		
Solitary tract nucleus	xx	o	xx	-	xx	-	x	o	x	-	xx	x	xx		

xxx: extremely high, xx: high, x: Low, o: very low.

4.2.3 Structural classification

4.2.3.1 Homo-Oligomeric $GABA_A$ receptors

The widely distributed Homo-Oligomeric $GABA_A$Rs are formed mostly from five ρ-type subunits while those formed from murine β_1 or β_3 and human γ_{2L} subunits are distributed less abundantly. Even in the absence of GABA, the gated-chlorine channels generated from murine or rat β_1 or β_3 subunits are self-opening. Picrotoxin, interestingly, inhibits these receptors by blocking the channel pore. This impact is species-dependent, as it is not seen with human or bovine β_1 subunits. In contrast, the feasibility of creating Homo-Oligomeric channels from α_1, β_2, γ_2, or δ subunits has yet to be determined [29,33]. Whereas some electrophysiological examinations [35–38] supported the existence of these channels, others [39,40] demonstrated their absence. Likewise, subunits ε and π appear to be unable to form Homo-Oligomeric channels.

4.2.3.2 Composed of two different subunits

$GABA_A$ receptors resulting from the assembly of two Hetero-subunits are formed more efficiently in the brain than Homo-Oligomeric receptors. Their channels are more sensitive to GABA impact and can be triggered even at low concentrations. Moreover, they provide a high chloride ion permeability into the cell compared to the Homo-Oligomeric channels [33].

In all analyzed systems, the assembly efficiency of GABA-activated channels that are constructed from αβ-subunit combinations seems to be extremely high. In contrast, investigations performed on the human embryonic kidney (HEK-293) cells yield conflicting results for the αγ or βγ subunit combinations. In this context, the aggregate efficiency of these produced from the $\alpha_1\gamma_2$ or $\beta_3\gamma_2$ combinations appears to be low. The assembly consisting of the $\beta_1\gamma_{2s}$, $\beta_2\gamma_{2s}$, and $\beta_3\gamma_{2s}$ subunits is created with efficiency corresponding to that of the homo-oligomeric β_{1-3} receptors, nevertheless with unique pharmacological aspects. Configurations of the $\alpha_1\gamma_2$ or $\beta_2\gamma_{2L}$ are maintained in the endoplasmatic reticulum (ER) and are unlikely to constitute functional receptors. On the other hand, it was agreed that the ρ-subunits can form active receptors by associating with γ_2-subunits and, in some cases, with subunits of glycine receptors. This co-assembling was detected mostly in the retinal cells. Furthermore, The co-assembling of $\alpha_1\varepsilon$, $\beta_1\varepsilon$, $\alpha_1\pi$, or $\beta_1\pi$ offers inactive chlorine-channels, while it is unclear whether or not the αδ, βδ, or γδ subunits can generate pentameric $GABA_A$ receptors [29,33].

4.2.3.3 Composed of three or more different subunit subtypes

$GABA_A$ receptors arranged from one sub-type of α, β, and γ subunit are the most prevalent in the brain and hence the most investigated. Theoretically, the possible existing stoichiometries are: 1α:1β:3γ, 2α:1β:2γ, 3α:1β:1γ, 1α:2β:2γ, 1α:3β:1γ, and 2α:2β:1γ. Among them, the predominant isoform in the adult mammalian brain is the 2α:2β:1γ. Else, arrangements with two different α, or β sub-types demonstrate different features than receptors with only a single sub-type of these subunits. A similar finding was also noted for the arrangements containing ($\alpha_1\beta_1\gamma_{2L}\delta$) and ($\alpha_1\beta_3\pi$ and $\alpha_1\beta_3\gamma_3\pi$), which display features different from those of ($\alpha_1\beta_1\gamma_{2L}$ and $\alpha_1\beta_1\delta$) and ($\alpha_1\beta_3$ and $\alpha_1\beta_3\gamma_3$), respectively. Moreover, the immunoprecipitation, electrophysiological, and fluorescence energy transfer studies do not support the arrangement of three identical subunit-types in the same $GABA_A$ receptor, whereas the possibility of combining five distinct subunits in a single receptor is still being debated [29,33].

4.2.4 Modulation

In addition to the neurotransmitter GABA, the $GABA_AR$ can be modulated by two endogenous neuromodulators (neurosteroids and the endocannabinoid 2-arachidonoylglycerol) [19] and a wide range of exogenous allosteric modulators that have significant roles in stimulating cognition, learning, and memory, as well as in the treatment of currently more prevalent psychiatric diseases such as anxiety, epilepsy, schizophrenia, depression, and insomnia [17,41]. GABA/muscimol, benzodiazepines, and t-butyl-bicyclophosphorothionate (TBPS)/picrotoxinin were the first binding sites revealed in the structure of the $GABA_A$ receptor. They were identified by *Sieghart* through radio-ligand binding studies [42]. Other allosteric binding sites were later added by *Johnston*, emphasizing the structural diversity of $GABA_A$ receptor-acting chemical patterns [43]. Indeed, subsequent researches have indicated that some of the following proposed sites, such as the benzodiazepine and neurosteroid sites, could be further split; which assumed the possibility of overlap between certain sites, which explains, in turn, why some drugs have more than one biological activity that depends on increasing or decreasing their concentration. Table I.4 presents some of the $GABA_AR$ ligand binding sites at the subunit interfaces.

1- Agonist binding sites that also interact with competitive agonists.
2- Picrotoxinin binding site that also recognizes γ-butyrolactones, caprolactams, and some insecticides
3- Sedative- hypnotic barbiturate binding sites

4- Neuroactive steroid binding sites
5- Benzodiazepine binding sites
6- Ethanol binding sites
7- Stereoselective binding sites for volatile anesthetics
8- Furosemide binding site
9- Zn^{2+} ion binding site
10- Divalent cation binding sites
11- La^{3+} ion binding site
12- Phospholipid binding sites
13- Phosphorylation sites involving specific protein kinase binding sites
14- Interacting sites of $GABA_A$ receptors and microtubules that may anchor receptor clusters at postsynaptic membranes.

Table I.4. Some $GABA_AR$ ligand binding sites at the subunit interfaces [44].

Interface	Extracellular domain	Trans-membrane domain
β^+/α^-	GABA	Etomidate
(2 copies/pentamer)		Propofol
		Volatiles anesthetics
α^+/β^-	EtOH (on δ)	Barbiturates
	Imidazo-BZ (on δ)	Propofol
	Pyrazoloquinolines	Octanol
γ^+/β^-	?	Barbiturates
		Propofol
		Octanol
α^+/γ^-	BZ	?
α^+/δ^-	?	?
δ^+/β^-	?	Barbiturates
		Propofol
		Octanol
?		Neuroactive steroids

4.2.5 Orthosteric GABA binding sites in $\alpha_1\beta_2\gamma_2$ isoform

GABA neurotransmitters bind to two binding sites in the $GABA_A$ structure; both are located at the ECD β^+/α^- interfaces (Figure I.11). The two GABA sites show nearly equal chemical selectivity for ligands, however, they differ significantly in the 3D organization and binding kinetics. In addition, the binding site inserted at the $\beta^+(C)/\alpha^-(D)$ interface was estimated to have a threefold higher affinity for GABA than the binding site at the $\beta^+(A)/\alpha^-(B)$ interface. Thus, the concept of collaboration between their functions is currently being debated. Indeed, it should be noted that different subunit combinations may alter sensitivity to GABA agonists and antagonists. The δ subunit, for example, demonstrated a higher affinity for several GABA agonists, notably muscimol and gaboxadol, whereas the γ subunit seemed to have a lower affinity [28,44,45].

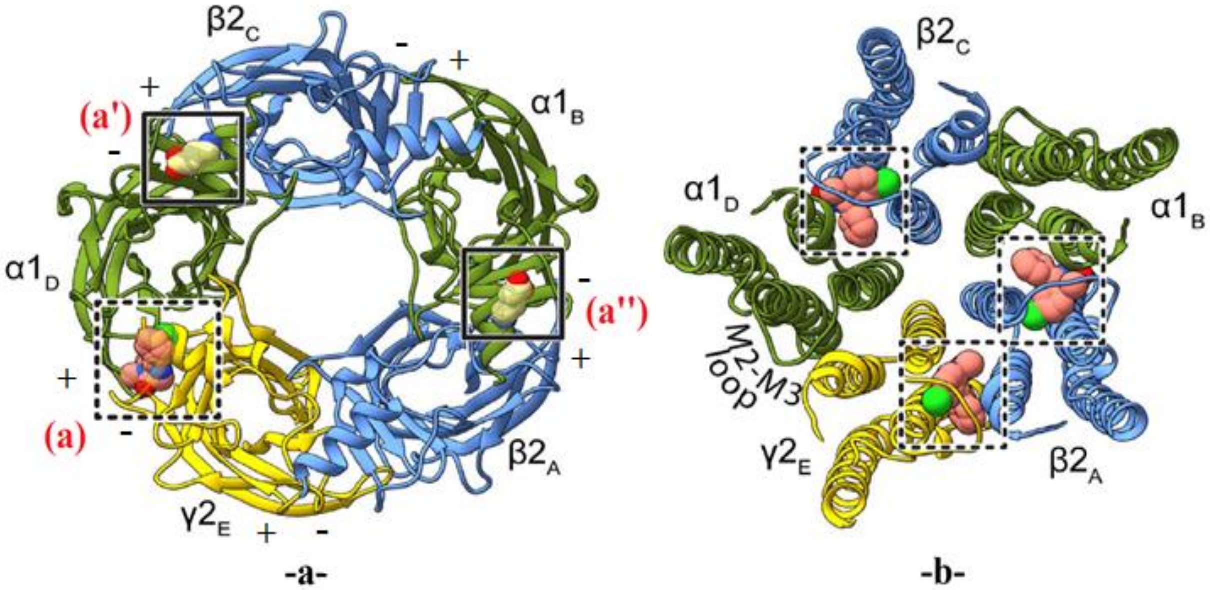
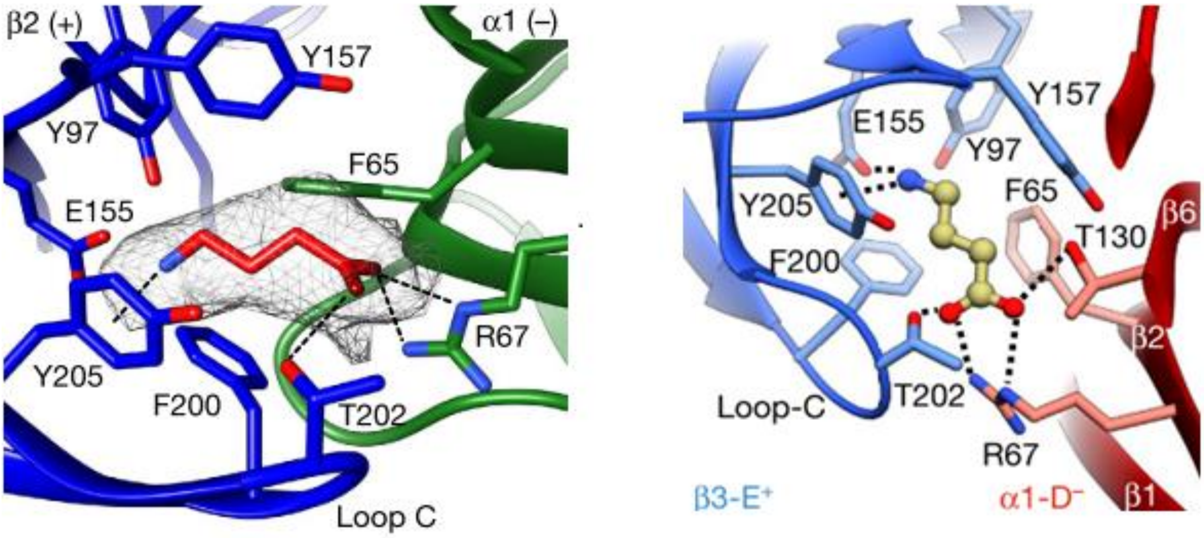


Figure I.11. Binding sites of the two endogenous agonists (GABA) and benzodiazepine. **a,** ECD interface for the binding of benzodiazepine (a) and GABA (a' and a''). **b,** the three TMD interfaces identified for the binding of benzodiazepine [34].

The GABA-binding modes observed in GABA$_A$/$\alpha_1\beta_2\gamma_2$ and GABA$_A$/$\alpha_1\beta_3\gamma_{2L}$ are compatible with each other. However, there are some differences in the proposed interactions (Figure I.12). In the GABA$_A$/$\alpha_1\beta_2\gamma_2$ subtypes, GABA has been found to participate in important hydrophobic interactions with α_1Phe65. The amino-nitrogen group contributes to favorable cation–π interactions with $\beta_2$phe200 and β_2Tyr205. Whereas, its carboxylate group forms a hydrogen bond with β_2Thr202 and β_2Thr130 and electrostatic interactions with the basic guanidinium group of α_1Arg67 [28]. Substitutions of β_2Tyr157 and β_2Tyr205 significantly reduced GABA sensitivity, confirming the relevance of these residues in GABA recognition [46]. Furthermore, β_2Glu155 and α_1Arg67 mutations reduce GABA potency [47]. In the GABA$_A$/$\alpha_1\beta_3\gamma_{2L}$ subtypes, the amino-nitrogen group maintained the cation-π interaction with β_3Tyr205 in addition to creating a network of hydrogen bonds with β_3Glu155, β_3Ser156, β_3Tyr157 and β_3Tyr97. The carboxylate group maintained hydrogen bonds with α_1Thr130 and β_3Thr202 while forming salt bridges with α_1Arg67 bonds instead of electrostatic interactions [45].



- a - **- b -**

Figure I.12. a, GABA-binding mode at β_2^+(C)/α_1^-(D) GABA$_A$/$\alpha_1\beta_2\gamma_2$ subtypes [28]. **b,** GABA-binding mode at β_3^+(C)/α_1^-(D) GABA$_A$/$\alpha_1\beta_3\gamma_{2L}$ subtypes [45].

5 Benzodiazepines

5.1 Chemistry

The term benzodiazepines (BDZ) refer to bicyclic heterocyclic compounds based on a benzene nucleus fused to a diazepine ring. The delocalization of nitrogen atoms in the diazepine ring divided the benzodiazepines into six basic rings: 5*H*-1,2-bdz, 1*H*-1,3-bdz, 3*H*-1,4-bdz, 3*H*-1,5-bdz, 5*H*-2,3-bdz, and 1*H*-2,4-bdz (Figure I.13) [48].

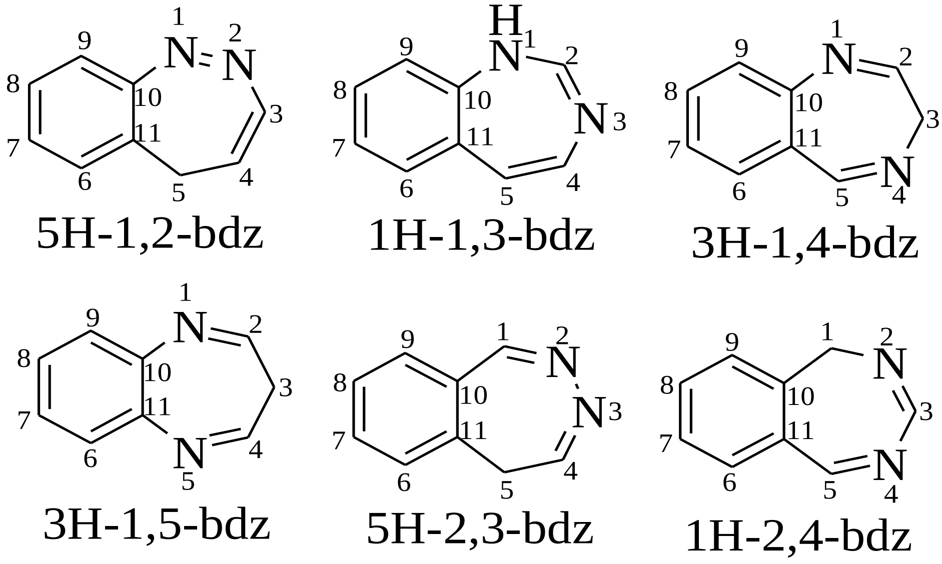


Figure I.13. Benzodiazepine basic rings [48,49].

5.2 Pharmacology

Clinically, benzodiazepines have widespread uses as anxiolytics, hypnotics, muscle relaxants, and anticonvulsants in epilepsy. Also, they are used as pre-anesthetic agents and as a withdrawal aid from alcohol and other drugs [50,51]. Their pharmacological effects are exerted by binding at the ECD and TMD binding interfaces in the presence of two GABA in their binding sites at the ECD (Figure I.11). The GABA neurotransmitter is responsible for opening the transmembrane channel that is permeable to chloride, and the presence of benzodiazepine potentiate its activity, thus increasing the conduction of the chloride channel and inhibiting the excitability of neurons [28]. As agreed, BDZs are agonists, inverse agonists, or antagonists of GABA_A receptors. BDZ receptor agonists are positive modulators that enhance the effect of GABA, inverse agonists are negative modulators that reduce the effect of GABA, while antagonists do not affect GABA function, but they limit the effect of benzodiazepines by blocking the binding locus (Figure I.14 (a)) [44]. Flumazenil is a common antagonist with a high affinity for the classical BDZ-binding site (Figure I.14 (b)). Its clinical applications include the treatment of BDZ overdoses and the reversal of diazepam-induced anesthesia [28].

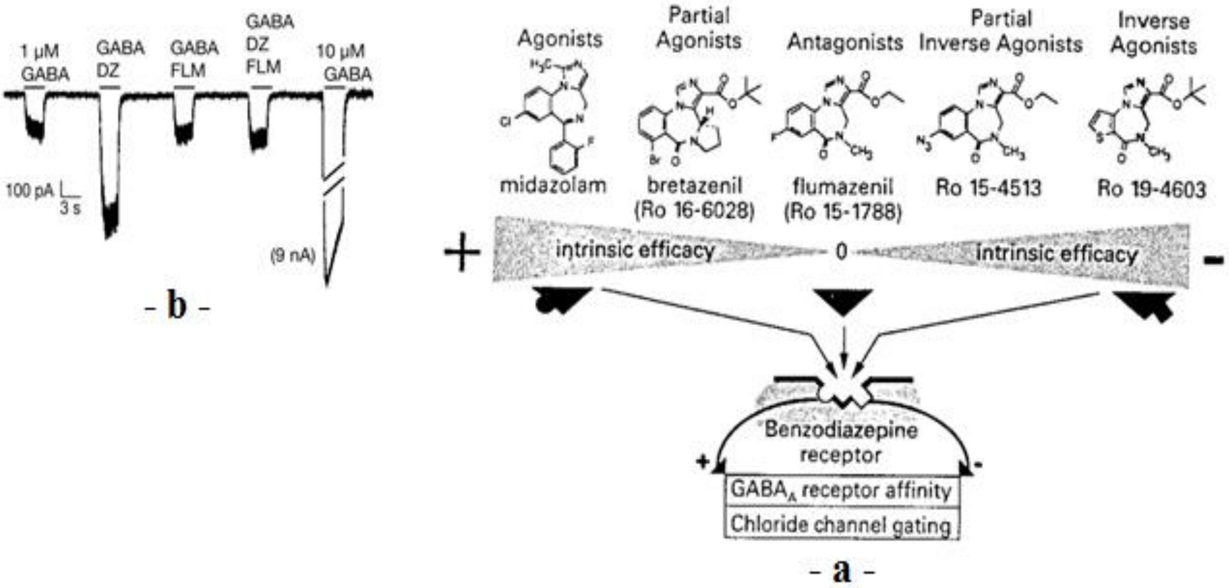
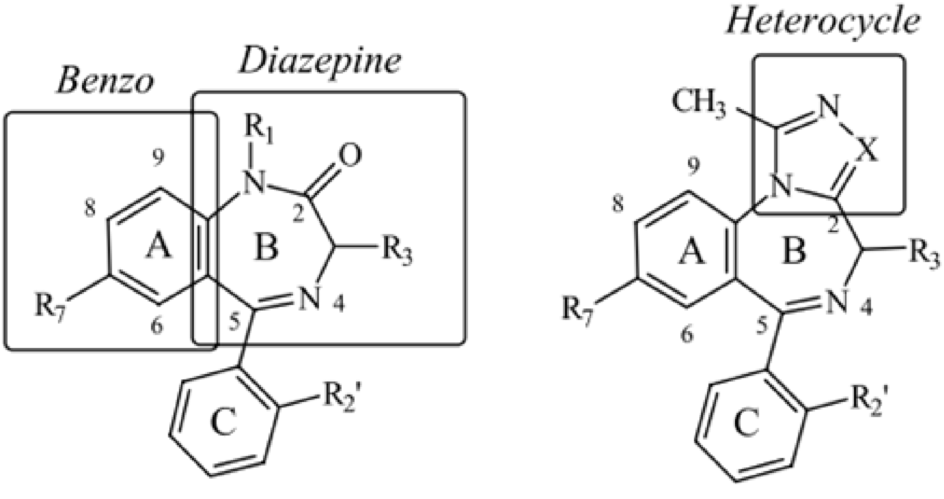


Figure I.14. a, An illustration of the spectrum of ligands and their allosteric effects on the GABA site, as well as their varying efficacies, positive or negative, at the BDZ binding site [52]. **b,** Electrophysiology of the Cryo-EM construct, showing the GABA_A receptor potentiation by the agonist diazepam (DZ), also showing that flumazenil (FLM) (3 µM) blocks GABA_A receptor potentiation by the agonist diazepam (DZ) (1 µM). n = 3 independent experiments [28].

Also, BDZ has wide pharmacological applications outside the CNS such as anticancer [53,54], antitrypanosomal [55,56], non-nucleoside inhibitors of HIV-1 reverse transcriptase [57], antimicrobial agent [58,59], antimalarial [60], antitumor agent [61], inhibitors of cholesterol absorption [62], inhibitors of the respiratory syncytial virus [63], and inhibitors of HCV NS5B polymerase [64].

5.3 1,4-dinitrogenetad BDZ

Among the six given basic rings, The 5-phenyl-1H-benzo[e][1,4]diazepin-2(3H)-one, also known as 5-Aryl-1,4-BDZ, serves as the skeleton of the most common psychotropic drugs currently on the marketed [65]. Otherwise, it can fuse with a triazole or imidazole ring to form, respectively, triazolobenodiazepines or imidazolobenzodiazepine (diazolobenzodiazepines) (Figure I.15). Importantly, the different drugs are distinguished by a variety of side groups that determine the degree of their binding to the GABA$_A$ receptor and thus may influence their efficacy, as well as their pharmacological and pharmacokinetic aspects [66].



- a -

-b -

Figure I.15. Structures of various 5-Aryl-1,4-BDZs. **a,** classical-BDZ structure, **b,** Imidazolo-benzodiazepines: X = CH, Triazolo-benzodiazepines: X = N [67].

5.3.1 Stereochemistry

Most 5-Aryl-1,4-based BDZs are devoid of chiral centers. Ring (B) might, however, adopt one of the two enantiomeric conformations (I) and (II) (Figure I.16). Importantly, conformation (I) is the most preferred for BDZ/target binding [67].

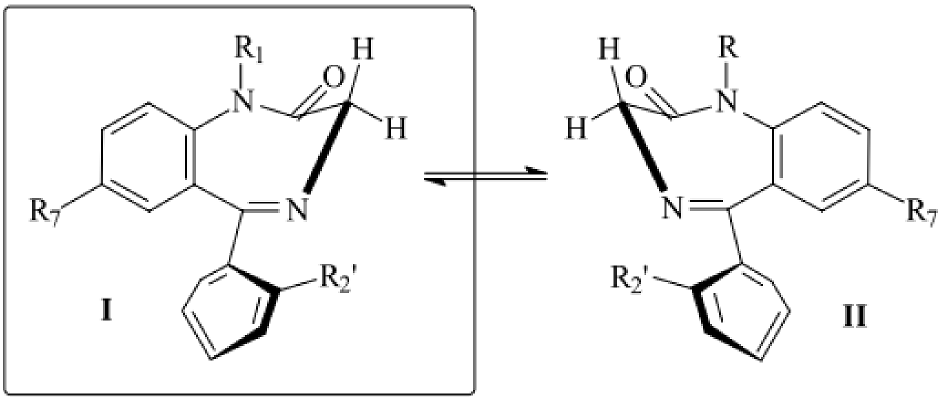


Figure I.16. Enantiomeric forms of 5-Aryl-1,4-based BDZs [67].

5.3.2 Structure-Activity relationship

- **Substitution at position 1:** alkylation by H, CH_3, or relatively small alkyl groups enhances the activity and produces prodrugs, ex: diazepam. The optimal functional groups for placing at position 1, in order from best to worst: OH > F > NH_2 > H > NHOH > Me > C1> CF_3 > Br > Et.
- **Substitution at position 2:** Although it can also be non-substitutable, ex: medazepam, the carboxylate-derived electronegative atom, preferably O or N, at this position enhances the activity.
- **Substitution at position 3:** If it is not substituted or has a hydroxyl group, the drug polarity will increase: glucuronoconjugation: faster elimination. Ex: lorazepam
- **Benzene ring at position 5:** optimal for activity. Contributes to BDZ-receptor binding through hydrophobic and steric interactions.
    - Substituted in *ortho* (2') by an electron-withdrawing group, preferably Cl or F, increases activity. Ex: flurazepam (F), clonazepam (Cl). The optimal functional groups for placing at position 2', in order from best to worst: NO_2 > F > CN > C1 > CF_3.
    - Substitution in *para* (4') may decrease activity.
    - It may also replace by another cycle. Ex: Cyclohexenyl (tetrazepam).
- **Substitution at position 7:** important in determining potency. A favorable position to increase activity, especially, when substituted by an electron-withdrawing group.
    - NO2: causes a hypnotic action, Ex: Clonazepam, nitrazepam.
    - X: causes an anxiolytic action, Ex: lorazepam, alprazolam.
    - The optimal 10 functional groups for placing at position 7, in order from best to worst: CH_2CF_3> I > Br > CF_3 > C1> $C(CH_3)_3$> NO_2 > F > N_3 > $CHCH_2$.
- **Substitution at positions 6, 8, and 9:** Here, any type of substitution decreases activity [68,69].

5.4 BDZ allosteric binding sites

5.4.1 ECD binding site

This binding locus is known as "the classical binding site" and is located exactly at the ECD α_1^+/γ_2^- interface. It has been recognized since the first BDZ was discovered in 1957. Therefore, it has been the subject of various researches targeted at identifying the mechanism of BDZ functioning. In this regard, the mutagenesis and cysteine crosslinking studies agreed on the importance of His102 in the recognition of classical benzodiazepines. Except for the

His102Cys mutant in the α_5 subunit, $GABA_A$Rs that contain αHis102 mutation to any other residue suffer from total insensitivity to the DZP and its analogs [28]. The topological organization of the $\alpha_4\beta\gamma_2$ and $\alpha_6\beta\gamma_2$ receptors reveals the presence of a natural substitution of His102 by the Arg residue, which leads to steric problems affecting the binding of classical BDZ at their correspondent binding locus. This could explain the selectivity of classical BDZ towards $\alpha\beta\gamma_2$Rs containing the α_1, α_2, α_3, and α_5 subunits rather than those containing the α_4 and α_6 subunits [45,70]. Else, the γ_2Phe77Tyr mutant affects less the binding affinity of DZP but more strongly reduces that of its analogs containing the chlorine substitutes at the pendant phenyl (C). This finding was explained by the difference in flexibility between the two pendant phenyls since the presence of the chlorine atoms possibly caused unfavorable steric clashes with the side chain of the tyrosine residue. Furthermore, several other mutation findings were surveyed in detail in the previous researches [71–73].

5.4.2 TMD binding sites

J.J. Kim and co-workers [34] demonstrated the presence of three additional BDZ binding sites in the TMD: two at β_2^+/α_1^- interfaces and a third at the γ_2^+/β_2^- interface (Figure I.11 (b)). β_2^+/α_1^- binding sites were also previously observed by *S. Masiulis and co-workers* [45] through their study on $\alpha_1\beta_3\gamma_2$Rs. The BDZ binding loci at both TMD β_2^+/α_1^- interfaces were identified to be also the binding sites for etomidate and propofol. As well as, the TMD binding locus located at the γ_2^+/β_2^- interface was estimated to overlap in part with that of the phenobarbital. Etomidate, propofol, and barbiturate are general intravenous anesthetics that differ from benzodiazepines in their ability -at higher doses- to directly activate $GABA_A$Rs without the need for GABA. In a mechanism similar to that of barbiturates, occupation of the γ_2^+/β_2^- interface by diazepam (DZP) leads to closing the gap presented at the interface between the two γ_2-β_2 subunits, which according to *J. J. Kim and co-workers* [34], may explain the anesthetic properties observed during the administration of a high dose of DZP. Currently, the [^{3}H]diazepam derivatives are the most common psychoactive drugs used to treat epilepsy, insomnia, muscle spasms, anxiety, alcohol withdrawal, and panic disorder [70].

5.5 Pharmacokinetics

The pharmacokinetic properties determine the onset of action and the duration of the drug effect. Pharmacokinetics specifically describes absorption, distribution, metabolism, and excretion. The side groups of BDZs influence their pharmacokinetic properties, particularly their solubility and metabolism [74].

5.5.1 Absorption

Due to their good absorption, BDZs are administered intramuscularly, intravenously, orally, sublingually, intranasally, or rectally. After oral administration, they are generally well absorbed from the gastrointestinal tract. After intravenous administration, they spread rapidly to the brain and CNS, making them effective in emergencies such as acute seizures [74–76].

5.5.2 Distribution

BDZs and their metabolites are highly protein-bound structures (70-99%). They are widely distributed in the organism and preferentially accumulated in lipid-rich areas such as the CNS and adipose tissue. They cross the BBB by passive diffusion. Generally, the more lipophilic the BDZ, the more it exhibits the highest absorption rates and the fastest onset of clinical effects [74–76].

5.5.3 Metabolism

Most BDZs are metabolized by cytochrome P450 enzymes. In phase (I), they undergo oxidation, hydroxylation, or dealkylation, while in phase (II); they conjugate with glucuronide or sulfate. Accordingly, the majority of BDZ drugs are excreted entirely in the urine. Besides, certain BDZs produce active metabolites that exert additional effects, increase drug action duration, and alter the elimination half-life. Ex: Diazepam: is a long-acting BDZ that produces three active metabolites: oxazepam, desmethyldiazepam, and temazepam. When prescribing BDZ drugs, especially for patients with extensive hepatic disease, these metabolites should be considered [74–76].

5.5.4 Elimination

BDZs are classified into the following categories based on their elimination half-life:

- **Short-acting drugs:** Elimination half-life <5 h. Mostly utilized as hypnotics due to their quick sleep onset. They have few residual effects and can cause rebound insomnia when disrupted, as well as amnesia and dependence problems. Ex: Midazolam and triazolam.
- **Intermediate-acting drugs.** Elimination half-life 5-24 h. Typically they are used for anxiety purposes. Might have next-day residual effects if used as a hypnotic. Ex: Alprazolam, lorazepam, lormetazepam.
- **Long-acting drugs.** Elimination half-life >24 h. They present risk of accumulation, especially in the elderly or patients with metabolic disease. Ex: Diazepam, clorazepate [77].

Bibliography

[1] Z. Kovács, D.P. D’Agostino, D. Diamond, M.S. Kindy, C. Rogers, C. Ari, Therapeutic potential of exogenous ketone supplement induced ketosis in the treatment of psychiatric disorders: Review of current literature, Front. Psychiatry. 10 (2019) 1–15. https://doi.org/10.3389/fpsyt.2019.00363.

[2] P.V. Ramachandran, A. Shekhar, Welcome to “GABAergic drugs,” Future Med. Chem. 3 (2011) 139–140. https://doi.org/10.4155/fmc.10.301.

[3] G. Deidda, I.F. Bozarth, L. Cancedda, Modulation of GABAergic transmission in development and neurodevelopmental disorders: Investigating physiology and pathology to gain therapeutic perspectives, Front. Cell. Neurosci. 8 (2014) 1–23. https://doi.org/10.3389/fncel.2014.00119.

[4] B. Dell’osso, M. Lader, Do benzodiazepines still deserve a major role in the treatment of psychiatric disorders? A critical reappraisal, Eur. Psychiatry. 28 (2013) 7–20. https://doi.org/10.1016/j.eurpsy.2011.11.003.

[5] D.M. Lovinger, Communication Networks in the Brain Neurons, Receptors, Neurotransmitters, and Alcohol, Alcohol Res. Heal. 31 (2008) 196–214.

[6] D.S. Faber, A.E. Pereda, Two forms of electrical transmission between neurons, Front. Mol. Neurosci. 11 (2018) 1–11. https://doi.org/10.3389/fnmol.2018.00427.

[7] D. A.Torse, R. R. Maggavi, S. A. Pujari, Nonlinear Blind Source Separation for EEG Signal Pre-processing in Brain-Computer Interface System for Epilepsy, Int. J. Comput. Appl. 50 (2012) 12–19. https://doi.org/10.5120/7838-0911.

[8] B.P. Bean, The action potential in mammalian central neurons, Nat. Rev. Neurosci. 8 (2007) 451–465. https://doi.org/10.1038/nrn2148.

[9] G.J. Kress, S. Mennerick, Action potential initiation and propagation: Upstream influences on neurotransmission, Neuroscience. 158 (2009) 211–222. https://doi.org/10.1016/j.neuroscience.2008.03.021.

[10] D.M. Rosenbaum, S.G.F. Rasmussen, B.K. Kobilka, The structure and function of G-protein-coupled receptors, Nature. 459 (2009) 356–363. https://doi.org/doi: 10.1038/nature08144.

[11] M. Jo, S.T. Jung, Engineering therapeutic antibodies targeting G-protein-coupled receptors, Exp. Mol. Med. 48 (2016) 1–9. https://doi.org/10.1038/emm.2015.105.

[12] A.J. Venkatakrishnan, X. Deupi, G. Lebon, C.G. Tate, G.F. Schertler, M. Madan Babu, Molecular signatures of G-protein-coupled receptors, Nature. 494 (2013) 185–194. https://doi.org/10.1038/nature11896.

[13] K. Palczewski, T. Kumasaka, T. Hori, C.A. Behnke, H. Motoshima, B.A. Fox, I. Le Trong, D.C. Teller, T. Okada, R.E. Stenkamp, M. Yamamoto, M. Miyano, Crystal

structure of rhodopsin: A G protein-coupled receptor, Science (80-.). 289 (2000) 739–745. https://doi.org/10.1126/science.289.5480.739.

[14] D. Lemoine, R. Jiang, A. Taly, T. Chataigneau, A. Specht, T. Grutter, Ligand-gated ion channels: New insights into neurological disorders and ligand recognition, Chem. Rev. 112 (2012) 6285–6318. https://doi.org/10.1021/cr3000829.

[15] A.J. Thompson, H.A. Lester, S.C.R. Lummis, The structural basis of function in Cys-loop receptors, 2010. https://doi.org/10.1017/S0033583510000168.

[16] C.A. Briggs, M. Gopalakrishnan, Ion channels-ligand gated, in: Compr. Med. Chem. II, 2006: pp. 877–918. https://doi.org/10.1016/b0-08-045044-x/00067-5.

[17] V.R. Solomon, V.J. Tallapragada, M. Chebib, G.A.R. Johnston, J.R. Hanrahan, GABA allosteric modulators: An overview of recent developments in non-benzodiazepine modulators, Eur. J. Med. Chem. 171 (2019) 434–461. https://doi.org/10.1016/j.ejmech.2019.03.043.

[18] S.H. Chuang, D.S. Reddy, Genetic and molecular regulation of extrasynaptic GABA-A receptors in the brain: Therapeutic insights for epilepsy, J. Pharmacol. Exp. Ther. 364 (2018) 180–197. https://doi.org/10.1124/jpet.117.244673.

[19] E. Sigel, M.E. Steinmann, Structure, function, and modulation of GABAA receptors, J. Biol. Chem. 287 (2012) 40224–40231. https://doi.org/10.1074/jbc.R112.386664.

[20] R.W. Olsen, W. Sieghart, International Union of Pharmacology. LXX. Subtypes of γ-aminobutyric acidA receptors: Classification on the basis of subunit composition, pharmacology, and function. Update, Pharmacol. Rev. 60 (2008) 243–260. https://doi.org/10.1124/pr.108.00505.

[21] M. Terunuma, Diversity of structure and function of GABA B receptors : a complexity of GABA B -mediated signaling, Proc. Japan Acad. Ser. B. 94 (2018) 390–411.

[22] B. Bettler, K. Kaupmann, J. Mosbacher, M. Gassmann, Molecular structure and physiological functions of GABAB receptors, Physiol. Rev. 84 (2004) 835–867. https://doi.org/10.1152/physrev.00036.2003.

[23] M. Chebib, G.A.R. Johnston, GABA-Activated Ligand Gated Ion Channels: Medicinal Chemistry and Molecular Biology, J. Med. Chem. 43 (2000) 1427–1447. https://doi.org/10.1021/jm9904349.

[24] M. Watanabe, K. Maemura, K. Kanbara, T. Tamayama, H. Hayasaki, GABA and GABA receptors in the central nervous system and other organs, Int. Rev. Cytol. 213 (2002) 1–47. https://doi.org/10.1016/S0074-7696(02)13011-7.

[25] A. Schousboe, L.K. Bak, H.S. Waagepetersen, Astrocytic control of biosynthesis and turnover of the neurotransmitters glutamate and GABA, 4 (2013) 1–12. https://doi.org/10.3389/fendo.2013.00102.

[26] N.M. Rowley, K.K. Madsen, A. Schousboe, H. Steve White, Glutamate and GABA

synthesis, release, transport and metabolism as targets for seizure control, Neurochem. Int. 61 (2012) 546–558. https://doi.org/10.1016/j.neuint.2012.02.013.

[27] R.S. Zucker, D.M. Kullmann, P.S. Kaeser, Release of Neurotransmitters, Third Edit, Elsevier Inc., 2014. https://doi.org/10.1016/B978-0-12-397179-1.00015-4.

[28] S. Zhu, C.M. Noviello, J. Teng, R.M. Walsh, J.J. Kim, R.E. Hibbs, Structure of a human synaptic GABAA receptor, Nature. 559 (2018) 67–88. https://doi.org/10.1038/s41586-018-0255-3.

[29] D. Berezhnoy, M.C. Gravielle, D.H. Farb, Pharmacology of the GABA A Receptor, in: Handb. Contemp. Neuropharmacol., 2007. https://doi.org/10.1002/9780470101001.hcn012.

[30] P.S. Miller, A.R. Aricescu, Crystal structure of a human GABAA receptor, Nature. 512 (2014) 270–275. https://doi.org/10.1038/nature13293.

[31] H.C. Chua, M. Chebib, GABAA Receptors and the Diversity in their Structure and Pharmacology, in: Adv. Pharmacol., 1st ed., Elsevier Inc., 2017: pp. 1–34. https://doi.org/10.1016/bs.apha.2017.03.003.

[32] D. Nutt, GABA A Receptors : Subtypes, Regional Distribution, and Function, (1999).

[33] W. Sieghart, G. Sperk, Subunit Composition, Distribution and Function of GABA, Curr. Top. Med. Chem. 2 (2002) 795–816.

[34] J.J. Kim, A. Gharpure, J. Teng, Y. Zhuang, R.J. Howard, S. Zhu, C.M. Noviello, R.M. Walsh, E. Lindahl, R.E. Hibbs, Shared structural mechanisms of general anaesthetics and benzodiazepines, Nature. 585 (2020) 303–308. https://doi.org/10.1038/s41586-020-2654-5.

[35] T.A. Verdoorn, A. Draguhn, S. Ymer, P.H. Seeburg, B. Sakmann, Functional properties of recombinant rat GABAA receptors depend upon subunit composition, Neuron. 4 (1990) 919–928. https://doi.org/10.1016/0896-6273(90)90145-6.

[36] A. Pöltl, B. Hauer, K. Fuchs, V. Tretter, W. Sieghart, Subunit composition and quantitative importance of GABAA receptor subtypes in the cerebellum of mouse and rat, J. Neurochem. 87 (2003) 1444–1455. https://doi.org/10.1046/j.1471-4159.2003.02135.x.

[37] L.A. C Blair, E.S. Levitan, J. Marshall, V.E. Dionne, E.A. BARNARDt, Single Subunits of the GABAA Receptor Form Ion Channels with Properties of the Native Receptor, Science (80-.). 242 (1988) 577–579. https://doi.org/10.1126/science.2845583.

[38] D.B. Pritchett, H. Sontheimer, C.M. Gorman, H. Kettenmann, P.H. Seeburg, P.R. Schofield, Transient expression shows ligand gating and allosteric potentiation of GABAA receptor subunits, Science (80-.). 242 (1988) 1306–1308. https://doi.org/10.1126/science.2848320.

[39] E. Sigel, R. Baur, G. Trube, H. Möhler, P. Malherbe, The effect of subunit composition of rat brain GABAA receptors on channel function, Neuron. 5 (1990) 703–711. https://doi.org/10.1016/0896-6273(90)90224-4.

[40] P.A. Davies, E.F. Kirkness, T.G. Hales, Modulation by general anaesthetics of rat GABA(A) receptors comprised of α1β3 and β3 subunits expressed in human embryonic kidney 293 cells, Br. J. Pharmacol. 120 (1997) 899–909. https://doi.org/10.1038/sj.bjp.0700987.

[41] R. Bergmann, K. Kongsbak, P.L. Sørensen, T. Sander, T. Balle, A Unified Model of the GABAA Receptor Comprising Agonist and Benzodiazepine Binding Sites, PLoS One. 8 (2013). https://doi.org/10.1371/journal.pone.0052323.

[42] W. Sieghart, Structure and pharmacology of γ-aminobutyric acid A receptor subtypes, Pharmacol. Rev. 47 (1995) 181–234.

[43] G.A.R. Johnston, GABAA receptor pharmacology, Pharmacol. Ther. 69 (1996) 173–198. https://doi.org/10.1016/0163-7258(95)02043-8.

[44] R.W. Olsen, GABAA receptor: Positive and negative allosteric modulators, Neuropharmacology. 136 (2018) 10–22. https://doi.org/10.1016/j.neuropharm.2018.01.036.

[45] S. Masiulis, R. Desai, T. Uchański, I. Serna Martin, D. Laverty, D. Karia, T. Malinauskas, J. Zivanov, E. Pardon, A. Kotecha, J. Steyaert, K.W. Miller, A.R. Aricescu, GABAA receptor signalling mechanisms revealed by structural pharmacology, Nature. 565 (2019) 454–459. https://doi.org/10.1038/s41586-018-0832-5.

[46] J. Amin, D.S. Weiss, GABAA receptor needs two homologous domains of the beta-subunit for activation by GABA but not by pentobarbital, Nature. 366 (1993) 565–569. https://doi.org/10.1038/366565a0.

[47] M. Mortensen, F. Iqbal, A.P. Pandurangan, S. Hannan, R. Huckvale, M. Topf, J.R. Baker, T.G. Smart, Photo-antagonism of the GABA A receptor, Nat. Commun. 5 (2014). https://doi.org/10.1038/ncomms5454.

[48] G.A. Archer, L.H. Sternbach, The chemistry of benzodiazepines, Chem. Rev. 68 (1968) 747–784. https://doi.org/10.1021/cr60256a004.

[49] B. Ahmed, M. Rashid, Chemistry and biological activity of benzodiazepines, Org. Chem. an Indian J. 4 (2008) 486–507. https://doi.org/10.1021/cen-v056n035.p026.

[50] B. Hayhoe, J. Lee-Davey, Tackling benzodiazepine misuse, BMJ. 362 (2018) k3208. https://doi.org/10.1136/bmj.k3208.

[51] B. Moosmann, V. Auwärter, Designer benzodiazepines: Another class of new psychoactive substances, in: Handb. Exp. Pharmacol., 2018: pp. 383–410. https://doi.org/10.1007/164_2018_154.

[52] E.A. BARNARD, P. SKOLNICK, R.W. OLSEN, H. MOHLER, W. SIEGHART, G. BIGGIO, C. BRAESTRUP, A.N. BATESON, S.Z. LANGER, International Union of Pharmacology. XV. Subtypes of ⫽-Aminobutyric AcidA Receptors: Classification on the Basis of Subunit Structure and Receptor Function, Pharmacol. Rev. 50 (1998) 291–313.

[53] C. Cortes Eduardo, H.-O. Simon, R. Apan Teresa, N. Camacho Antonio, I. V. Lijanova, M.-G. Marcos, Anticancer Activity and Anti-inflammatory Studies of 5-Aryl-1,4-benzodiazepine Derivatives, Anticancer. Agents Med. Chem. 12 (2012) 611–618. https://doi.org/10.2174/187152012800617713.

[54] A. Kamal, K.S. Reddy, M.N.A. Khan, R.V.C.R.N.C. Shetti, M.J. Ramaiah, S.N.C.V.L. Pushpavalli, C. Srinivas, M. Pal-Bhadra, M. Chourasia, G.N. Sastry, A. Juvekar, S. Zingde, M. Barkume, Synthesis, DNA-binding ability and anticancer activity of benzothiazole/benzoxazole-pyrrolo[2,1-c][1,4]benzodiazepine conjugates, Bioorganic Med. Chem. 18 (2010) 4747–4761. https://doi.org/10.1016/j.bmc.2010.05.007.

[55] N. Maheshwari, A. Goyal, S. Jain, 2D-QSAR study of 1,4-benzodiazepine-2-ones as potent anti-trypanosomal agents, Med. Chem. Res. 22 (2013) 6002–6009. https://doi.org/10.1007/s00044-013-0592-6.

[56] J. Spencer, R.P. Rathnam, A.L. Harvey, C.J. Clements, R.L. Clark, M.P. Barrett, P.E. Wong, L. Male, S.J. Coles, S.P. MacKay, Synthesis and biological evaluation of 1,4-benzodiazepin-2-ones with antitrypanosomal activity, Bioorganic Med. Chem. 19 (2011) 1802–1815. https://doi.org/10.1016/j.bmc.2011.01.010.

[57] V. Wiwanitkit, Analysis of binding energy activity of TIBO and HIV-RT based on simple consideration for conformational change, African J. Biotechnol. 6 (2007) 188–189. https://doi.org/10.5897/AJB2007.000-2003.

[58] Y.S. An, Z.F. Hao, X.J. Zhang, L.Z. Wang, Efficient Synthesis and Biological Evaluation of a Novel Series of 1,5-Benzodiazepine Derivatives as Potential Antimicrobial Agents, Chem. Biol. Drug Des. (2016) 110–121. https://doi.org/10.1111/cbdd.12739.

[59] S. Mor, P. Pahal, B. Narasimhan, Synthesis, characterization, biological evaluation and QSAR studies of 11-p-substituted phenyl-12-phenyl-11a,12-dihydro-11H-indeno[2,1-c][1,5] benzothiazepines as potential antimicrobial agents, Eur. J. Med. Chem. 57 (2012) 196–210. https://doi.org/10.1016/j.ejmech.2012.09.003.

[60] L. Nallan, K.D. Bauer, P. Bendale, K. Rivas, K. Yokoyama, C.P. Hornéy, P.R. Pendyala, D. Floyd, L.J. Lombardo, D.K. Williams, A. Hamilton, S. Sebti, W.T. Windsor, P.C. Weber, F.S. Buckner, D. Chakrabarti, M.H. Gelb, W.C. Van Voorhis, Protein farnesyltransferase inhibitors exhibit potent antimalarial activity, J. Med. Chem. 48 (2005) 3704–3713. https://doi.org/10.1021/jm0491039.

[61] D. Antonow, M. Kaliszczak, G.D. Kang, M. Coffils, A.C. Tiberghien, N. Cooper, T. Barata, S. Heidelberger, C.H. James, M. Zloh, T.C. Jenkins, A.P. Reszka, S. Neidle,

S.M. Guichard, D.I. Jodrell, J.A. Hartley, P.W. Howard, D.E. Thurston, Structure-activity relationships of monomeric C2-aryl pyrrolo[2,1- c][1,4]benzodiazepine (PBD) antitumor agents, J. Med. Chem. 53 (2010) 2927–2941. https://doi.org/10.1021/jm901722v.

[62] Y.A. Kesäniemi, T.A. Miettinen, Inhibition of cholesterol absorption by neomycin, benzodiazepine derivatives and ketoconazole, Eur. J. Clin. Pharmacol. 40 (1991) 65–67. https://doi.org/10.1007/BF01409412.

[63] E.A. Henderson, D.G. Alber, R.C. Baxter, S.K. Bithell, J. Budworth, M.C. Carter, A. Chubb, G.S. Cockerill, V.C.L. Dowdell, I.J. Fraser, R.A. Harris, S.J. Keegan, R.D. Kelsey, J.A. Lumley, J.N. Stables, N. Weerasekera, L.J. Wilson, K.L. Powell, 1,4-Benzodiazepines as inhibitors of respiratory syncytial virus. The identification of a clinical candidate, J. Med. Chem. 50 (2007) 1685–1692. https://doi.org/10.1021/jm060747l.

[64] D. McGowan, O. Nyanguile, M.D.C.S. Vendeville, P.R. Koen Vandyck, Walter Van den Broeck, Carlo W. Boutton, Hendrik De Bondt, Ludo Quirynen, Katie Amssoms, Jean-François Bonfanti, Stefaan Last, Klara Rombauts, Abdellah Tahri, Lili Hu, Frédéric Delouvroy, Katrien Vermeiren, Geneviève Vandercruyssen, Liesbet, 1,5-Benzodiazepine inhibitors of HCV NS5B polymerase, Bioorganic Med. Chem. Lett. 19 (2009) 2492–2496. https://doi.org/10.1016/j.bmcl.2009.03.035.

[65] P. Aastha, K. Navneet, A. Anshu, S. Pratima, K. Dharma, 1 , 5 Benzodiazepines : Overview of Properties and Synthetic Aspects, Res. J. Chem. Sci. 3 (2013) 90–103. https://doi.org/10.1002/CHIN.201422238.

[66] R.L. Hagan, Clarification of Benzodiazepine Structural Classes, J. Anal. Toxicol. 19 (1995) 58–59.

[67] J. DeRuiter, GABA Receptor Ligands and the Benzodiazepines, Princ. Drug Action 2. Fall (2004) 1–27.

[68] D.J. Maddalena, G.A.R. Johnston, Prediction of Receptor Properties and Binding Affinity of Ligands to Benzodiazepine/GABAA Receptors Using Artificial Neural Networks, J. Med. Chem. 38 (1995) 715–724. https://doi.org/10.1021/jm00004a017.

[69] S.S. So, M. Karplus, Genetic neural networks for quantitative structure-activity relationships: Improvements and application of benzodiazepine affinity for benzodiazepine/GABAA receptors, J. Med. Chem. 39 (1996) 5246–5256. https://doi.org/10.1021/jm960536o.

[70] L. Richter, C. De Graaf, W. Sieghart, Z. Varagic, M. Mörzinger, I.J.P. De Esch, G.F. Ecker, M. Ernst, Diazepam-bound GABAA receptor models identify new benzodiazepine binding-site ligands, Nat. Chem. Biol. 8 (2012) 455–464. https://doi.org/10.1038/nchembio.917.

[71] K.R. Tan, A. Gonthier, R. Baur, M. Ernst, M. Goeldner, E. Sigel, Proximity-accelerated chemical coupling reaction in the benzodiazepine-binding site of γ-

aminobutyric acid type A receptors: Superposition of different allosteric modulators, J. Biol. Chem. 282 (2007) 26316–26325. https://doi.org/10.1074/jbc.M702153200.

[72] J.M.C. Derry, S.M.J. Dunn, M. Davies, Identification of a residue in the γ-aminobutyric acid type A receptor α subunit that differentially affects diazepam-sensitive and -insensitive benzodiazepine site binding, J. Neurochem. 88 (2004) 1431–1438. https://doi.org/10.1046/j.1471-4159.2003.02264.x.

[73] U. Rudolph, H. Möhler, Analysis of GABAA Receptor Function and Dissection of the Pharmacology of Benzodiazepines and General Anesthetics Through Mouse Genetics, Annu. Rev. Pharmacol. Toxicol. 44 (2004) 475–498. https://doi.org/10.1146/annurev.pharmtox.44.101802.121429.

[74] C.E. Griffin, A.M. Kaye, F. Rivera Bueno, A.D. Kaye, Benzodiazepine pharmacology and central nervous system-mediated effects, Ochsner J. 13 (2013) 214–223.

[75] T.I. Saari, M. Uusi-Oukari, J. Ahonen, K.T. Olkkola, Enhancement of GABAergic activity: Neuropharmacological effects of benzodiazepines and therapeutic use in anesthesiology, Pharmacol. Rev. 63 (2011) 243–267. https://doi.org/10.1124/pr.110.002717.

[76] R. Elisabeth, J.M. Tréluyer, G. Pons, Pharmacokinetic optimisation of benzodiazepine therapy for acute seizures. Focus on delivery routes, Clin. Pharmacokinet. 36 (1999) 409–424. https://doi.org/10.2165/00003088-199936060-00003.

[77] E.B. Rocafort, Clarification of Benzodiazepine Structural Classes, University of Barcelona, 2018.

Chapter II

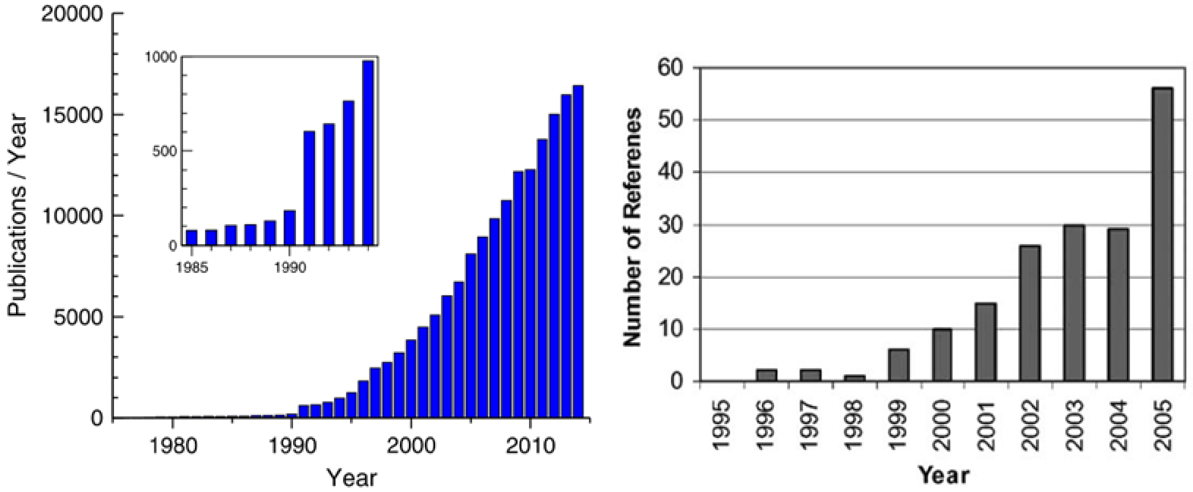
Overview on computational methods: Conceptual-DFT, MEP analysis, QSAR analysis, and Molecular docking/Dynamic simulations.

1 Introduction

Describing the chemical reactivity of a particular molecular system involves investigating how it responds to the approach of different types of chemical reagents. To establish this response, the inherent chemical reactivity of the entities is described by adopting their electronic structure in the isolated state as a reference point, and then the effects of attacking the reagents on this state are investigated [1]. Recently, chemical reactivity theory has emerged as an essential technique for determining chemical processes. Therefore, theoretical chemists set out to develop several theories to describe the stability, selectivity, and reactivity of molecular entities. Accordingly, the density-based theory, DFT, provided a sound basis for the development of computational strategies for obtaining information on the energetics, structure, and properties of atoms and molecules at much lower costs, excellent level of accuracy, and with lesser computational time than other existing methods [2].

Drug designing is an essential phase in the investigation of new medications that require multiscale simulations. As per published works [3], DFT is a well-accepted method for the precise study of molecular properties and description of physiologically important systems. Hence, DFT-hybrid functions, in particular B3LYP, are more appropriate for this purpose than LDA and GGA functions [3,4]. Figure II.1 (a) shows that between 1975 and 2014, the number of publications in the disciplines of "Density Functional" and "Density Functional Theory (DFT)" increased dramatically [5], and part (b) exhibits an increase in the number of publications implementing DFT for drug design between 1995 and 2005 [3]. QSAR modeling exploits DFT-derived descriptors to encode the molecular properties accountable for the relevant activity of the target chemicals and then establishes the quantitative correlation. Using reliable models, it is feasible to create hypothetical structures that either do not exist or have never been synthesized. Selecting the pertinent molecular descriptors from the many available is still until now the most challenging task in QSAR implementation [4].

Furthermore, understanding the mechanism of interaction between a potential therapeutic agent and its target is essential for drug design. This can be accomplished via molecular docking and dynamic simulations. The latter are powerful approaches to identifying all potential drug/target interactions: covalent bonds, ionic contacts, ion-dipole interactions, dipole-dipole interactions, hydrogen bonding, charge transfer, and hydrophobic interactions. When assessing the effectiveness of DFT in computer-aided drug design, its robustness in estimating the strengths of these interactions has drawn attention. Specifically, DFT succeeded in predicting covalent, ionic, and hydrogen bonds, while it suffered from hurdles in predicting weaker bonds: ion-dipole, dipole-dipole, charge transfer, and hydrophobic interactions [4,6].



- a -

- b -

Figure II.1. a, Number of publications per year (1975–2014) on the terms ‘’Density Functional’’ and ‘’Density Functional Theory (DFT) [5]. **b,** Number of publications per year (1995– 2005) that used DFT for drug design [3].

Accordingly, the integration of DFT with MEP analysis and CADD methods incorporating QSAR analysis, molecular docking, and molecular dynamics simulations is not a randomized protocol in computational drug discovery and development research. The background knowledge in this chapter gives readers a better understanding of these approaches which, in next chapters, have yielded trustworthy outcomes in the investigation of the biological process of benzodiazepines.

2 Conceptual-DFT

Conceptual DFT, as a branch of DFT, is intended to develop chemical reactivity theory based on the concepts of DFT. Thus, DFT-derived physical and mathematical notions are reused and merged to generate descriptors for interpreting and predicting chemical phenomena. Conceptual DFT-based descriptors have proven to be effective tools for describing and analyzing the chemical reactivity and selectivity of molecular systems [7]. At present, several global and local reactivity descriptors have been offered, as well as physical definitions, based on the hierarchy of the electronic energy derivatives developed in the canonical ensemble $E[N; v(r)]$ [8]. In the following sections, we will outline the fundamental descriptors we used in this dissertation.

2.1 Global reactivity descriptors

The proposed global descriptors provide information on the global behavior of chemical species as a whole. They consider the reference point of the DFT of chemical reactivity by defining the notion of chemical potential (μ) and the negative of electronegativity (χ) as the first partial derivative of the total energy concerning the number of electrons (N_e), when the external potential $v(r)$ remains constant [9]:

$$\mu = -\chi = (\partial E/\partial N_e)_{v(r)} \tag{1}$$

The chemical potential estimates the propensity of electrons to escape from the molecular system [10]. More extensively, electrons tend to redistribute between regions of higher and lower chemical potentials until space equilibrium is reached [11]. Thus, the DFT-chemical potential specifies the inverse of the ability of chemical systems to attract electrons, which in turn is known by the concept of electronegativity [12].

Similarly, the second partial derivative of the total energy was defined by *Parr and Pearson* [13] as the concept of chemical hardness (η):

$$\eta = (\partial^2 E/\partial N_e^2)_{v(r)} \tag{2}$$

In a molecular system, chemical hardness refers to the degree of compaction of the electron cloud surrounding the nucleus. So, it is an estimation of its resistance to electrons shift toward other systems [14]. Except in the case of bulk metals, where it yields zero values, chemical hardness is always positive [15]. The reciprocal of the chemical hardness is defined as its chemical softness (S). This latter determines the relative diffusion of electron density between two interacting entities [14].

$$S = 1/2\eta \tag{3}$$

On another side, the combination of global hardness (η) with the chemical potential (μ) is given by *Parr et al* [16] as the electrophilic index (ω). Indeed, a system with a high ω value will naturally be more prone to acquire electronic density from the environment and thus be highly reactive [17].

$$\omega = \mu^2/2\eta \tag{4}$$

Since there is no electronic molecular stabilization along the subtraction of the electron density of a molecular system, the notions of electrophilicity and nucleophilicity cannot be related together by a single definition. In fact, when electrons are detached from a molecule, the molecular electronic energy constantly increases. Therefore, a variational approach like electrophilicity cannot be used to determine the nucleophilic index [14].

Accordingly, researchers held that nucleophilicity is defined as the negative value of gas ionization potentials (IP), N = -IP, and expressed as follows:

$$N = \xi_{HOMO(Nu)} - \xi_{HOMO\ (TCE)} \tag{5}$$

TCE: tetracyanoethylene. TCE was chosen as a reference since it exhibits the lowest HOMO energy among a large number of compounds [18].

According to the electrophilicity and nucleophilicity classification scales proposed by *L.R. Domingo et al* [18–20], electrophiles can be classed as strong ($\omega \geq 1.50$ eV), moderate ($1.50 < \omega < 0.80$ eV), and marginal ($\omega \leq 0.80$ eV). Likewise, nucleophiles can be classified as strong ($N \geq 3.00$ eV), moderate ($3.00 < N < 2.00$ eV), and marginal ($N < 2.00$ eV).

By using the finite difference approximation (FD), Eqs. (1) and (2) becomes as follows:

$$\mu = -\chi = 1/2\ \ (A + I) \tag{6}$$

$$\eta = 1/2\ \ (I - A) \tag{7}$$

A is the vertical electronic affinity, defined as the quantity of energy released as a result of the capture of an electron by a neutral atom (or molecule) to form a negative ion. The property **I** is the vertical ionization potential, defined as the quantity of energy that must be supplied to a neutral atom (or molecule) to remove the most loosely bound electron and form a positive ion [21]. Eqs below give the arithmetic forms of **A** and **I**:

$$A = E_N - E_{N+1} \tag{8}$$

$$I = E_{N-1} - E_N \tag{9}$$

Where, E_N, E_{N+1}, and E_{N-1} are the total ground-state energies of neutral, anionic, and cationic states, determined at the neutral geometry.

2.2 Local reactivity descriptors

The local descriptors give a profound understanding of the reactivity of a specific atomic site in a molecular system during chemical interactions or excitations [14]. The variation of the total energy as a function of $\nu(r)$ is quantified by the electronic density $\rho(r)$ at first order and by the Fukui function $f(r)$ at second order [22]:

$$\rho(r) = \left(\frac{\delta E}{\delta \nu(r)}\right)_{N_e} \tag{10}$$

$$f(r) = \left(\frac{\partial \rho(r)}{\partial N_e}\right)_{\nu(r)} = \left(\frac{\delta \mu}{\delta \nu(r)}\right)_{N_e} \tag{11}$$

μ, refers to the chemical potential.

Electronic density is best suited for application when examining the atomic reactivity of charged species. While Fukui functions (FFs) are more convenient when dealing with neutrals [22]. Due to the discontinuity of the derivative of Eq. (11), three alternative forms of FFs have been proposed based on the use of its right and left derivatives with respect to N_e. Through the use of FD approximation, they are determined by the formulas [9]:

$$f^{+}(r) = \left(\frac{\partial \rho(r)}{\partial N_e}\right)^{+}_{\nu(r)} = \rho_{N+1}(r) - \rho_N(r) \qquad \text{for nucleophilic attack} \qquad (12)$$

$$f^{-}(r) = \left(\frac{\partial \rho(r)}{\partial N_e}\right)^{-}_{\nu(r)} = \rho_{N}(r) - \rho_{N-1}(r) \qquad \text{for electrophilic attack} \qquad (13)$$

$$f^{\circ}(r) = \frac{1}{2}\,[f^{+}(r) + f^{-}(r)] \qquad \text{for radical attack} \qquad (14)$$

Where, $\rho_N(r)$, $\rho_{N-1}(r)$, and $\rho_{N+1}(r)$ are the total electronic densities at point **r** for neutral (N), cationic (N-1), and anionic (N+1) systems, respectively.

Generally, high FFs values imply large variations in electron density and therefore, the most reactive molecular sites when the number of electrons is changed. In terms of the frontier molecular orbitals approximation (FMO), FFs are evaluated through the formulas [23]:

$$f^{+}(r) \approx \rho_{LUMO}(r) = |\phi_{LUMO}\ (\vec{r})|^2 \qquad (15)$$

$$f^{-}(r) \approx \rho_{HOMO}(r) = |\phi_{HOMO}\ (\vec{r})|^2 \qquad (16)$$

$$f^{\circ}(r) = \frac{1}{2}[|\phi_{LUMO}\ (\vec{r})|^2 + |\phi_{HOMO}\ (\vec{r})|^2] \qquad (17)$$

This approximation provides a fast way to calculate the FFs, but neglects the orbital relaxation effect [8]. Moreover, the condensed-to-atoms FFs forms can be calculated using the procedure proposed by *Yang and Mortier* [24] in the term of electronic population as:

$$f^{+}(r) = p^{k}_{N+1} - p^{k}_{N} \qquad (18)$$

$$f^{-}(r) = p^{k}_{N} - p^{k}_{N-1} \qquad (19)$$

$$f^{\circ}(r) = \frac{1}{2}(p^{k}_{N+1} - p^{k}_{N-1}) \qquad (20)$$

Where p^{k}_{N}, p^{k}_{N-1}, and p^{k}_{N+1} are the electronic populations of the k^{th} atom in the neutral, cationic, and anionic systems, respectively. The electronic population is the difference between the atomic number Z_K and the atomic charge Q_K. The atomic charges can be calculated according to various population analysis schemes (Mullikan, natural, electrostatic …etc) [25].

Unfortunately, literature is replete with numerous published investigations with incorrect FF plots probably due to the unfamiliarity of researchers with the problem of degeneracy in the frontier orbitals or, in rare cases, in the ground state. The degeneracy of HOMO and LUMO arises due to the symmetric behavior of molecular systems. For example, if electrons are removed or added to only one of the frontier orbitals, symmetry breaking will occur in the corresponding FF. As a consequence, the local reactivity is affected. Hence, FF must maintain molecular symmetry. Once this challenge is overcome, mapping FFs rather than computing the condensed values becomes the ideal process for regioselectivity evaluation [26].

The dual descriptor ($f^{(2)}(r)$ or $\Delta f(r)$) is another useful function used to reveal reactive sites. This descriptor has the advantage to locate simultaneously the electrophilic and nucleophilic sites in a volume element centred at **r**. Regions with the most positive values ($f^{(2)}(r) > 0$) are the preferred sites for nucleophilic attacks, and regions with the most negative values ($f^{(2)}(r) < 0$) are the preferred sites for electrophilic attacks, while a value close to zero suggests that the molecular site is neither electrophilic nor nucleophilic [22,27]. As for the Fukui functions, through the FD and FMO approximations, this descriptor is calculated using the formulas:

$$f^{(2)}(r) = f^{+}(r) - f^{-}(r) = \rho_{N+1}(r) - 2\rho_{N}(r) + \rho_{N-1}(r) \tag{21}$$

$$f^{(2)}(r) \approx \rho_{LUMO}(r) - \rho_{HOMO}(r) \tag{22}$$

$f^{(2)}(r)$ has been demonstrated to be more robust than FF to correctly predict the preferable regions for nucleophilic and electrophilic attacks. As a consequence, in latest years, researchers have focused their attention on the degeneracy and quasi-degradation states of HOMO-LUMO orbitals, yielding in the development of a more precise operational formula to calculate and visualize $f^{(2)}(r)$ within the framework of FMO theory [28–33].

3 Quantitative molecular electrostatic potential analysis

The electrostatic potential (ESP) is an assessment of the electrostatic interaction between a unit point charge deposited at **r** and the system in question. ESP is formulated directly from Coulomb’s law and expressed as:

$$V_{Total}(r) = V_{Nuc}(r) + V_{Elec}(r) = \sum_{A} \frac{Z_A}{|r-R_A|} - \int \frac{\rho(r')}{|r-r\prime|} dr' \tag{23}$$

Where, Z and $\rho(r')$ denote, respectively, the charge on nucleus **A** located at $\mathbf{R_A}$ and the total electronic density.

As the equation shows, a positive $V_{Total}(r)$ value reflects the dominance of nuclear charges in the considerable site. Otherwise, the negative value indicates the dominance of electronic charges [34,35].

Since molecules always tend to approach each other in a complementary manner of ESP, this latter is commonly used to predict electron-rich/deficient regions and the mode of recognition of molecular species, as between drugs and receptors [36]. Hence, the repulsive and attractive electrostatic interactions also must be considered to provide a comprehensive knowledge of the reactive behavior. When $f^{(2)}(r)$ and ESP of a particular chemical reactivity yield conflicting predictions, a quantitative analysis is required to determine the dominant effect. In this context, electrostatics is always dominant when the reagents are separated by a large distance, while electron-transfer effects (effectively captured by $f^{(2)}(r)$) are usually dominant when the reagents are close together [37].

Quantitative ESP analyses are frequently carried out on the molecular van der Waals (vdW) surface by setting the iso-surface density to $\rho = 0.001\ e/bohr^3$ [38], since this value reflects the specific properties of molecular species, such as lone pairs and π-electrons, strained bonds, etc. MEP positive (negative) regions are electron-deficient (electron-rich) sites, susceptible to react with nucleophilic (electrophilic) reactants. A set of statistically-based quantities have been proposed to define the whole surface potentials of molecular entities, two of these are site-specific: the minima and maxima of ESP on the vdW surface abbreviated by ($V_{s,min}$) and ($V_{s,max}$), respectively [39]. The magnitude and distributions of $V_{s,min}$ and $V_{s,max}$ enabled the prediction and explanation of the strength and orientation of several intermolecular non-covalent interactions including hydrogen bonding, halogen bonding, and π-hole bonding [36]. As well, four quantities are global in nature: the local polarity (Π), the total ESP variance (σ_{tot}^2), the degree of charge balance (ν), and the average local ionization energy ($\bar{I}$) [34,35,39]. Mathematical formulas are outlined in eqs. (24-27).

$$\Pi = \frac{1}{n}\sum_{i=1}^{n}|V_S(r_i) - \bar{V}_S| \tag{24}$$

$$\sigma_{tot}^2 = \sigma_+^2 + \sigma_-^2 = \frac{1}{\alpha}\sum_{j=1}^{\alpha}\left[V_S^+(r_j) - \bar{V}_S^+\right]^2 + \frac{1}{\beta}\sum_{k=1}^{\beta}[V_S^-(r_k) - \bar{V}_S^-]^2 \tag{25}$$

$$\nu = \frac{\sigma_+^2\sigma_-^2}{(\sigma_{tot}^2)^2} \tag{26}$$

$$\bar{I}(r) = \frac{\sum_i \rho_i(r)|\varepsilon_i|}{\rho(r)} \tag{27}$$

$V_S(r_i)$ and $\bar{V}_S$denote, respectively, the value of MEP at point **i** on the surface and its overall average value. Similarly, $V_S^+(r_j)$, $V_S^-(r_k)$, $\bar{V}_S^+$, and $\bar{V}_S^-$are correspond to the positive and negative values of MEP at points **j** and **k** on the surface and their averages (see eqs. (28)-(30)). While, $\rho_i(r)$ and ε_i are the electron density and orbital energy of the ith molecular orbital, respectively.

$$\bar{V}_S = (\alpha\bar{V}_S^+ + \beta\bar{V}_S^-)/(\alpha + \beta) \tag{28}$$

$$\bar{V}_S^+ = \frac{1}{\alpha}\sum_{j=1}^{\alpha} V_S^+(r_j) \tag{29}$$

$$\bar{V}_S^- = \frac{1}{\beta}\sum_{k=1}^{\beta} V_S^-(r_k) \tag{30}$$

4 Quantitative structure–activity relationship

Quantitative structure-activity relationship (QSAR) modeling is a ligand-based drug design method proposed initially by *Hammett* in 1930s and developed later in the mid-1960s by *Hansch and Fujita* [40]. Its mathematical framework is based on the implementation of chemometric techniques to generate statistically-derived models correlating the independent variables (or descriptors) of systems to their dependent variables. The independent variables in drug design refer to all structural attributes of chemical entities (the physicochemical and biological properties) and the dependent variables refer to their functions (including experimental biological/biochemical responses such as binding affinity, activity, toxicity, rate constants, etc.) [41]. The generic QSAR model formula is expressed as follows:

Predicted Biological Activity = Function (Chemical Structure)

QSAR has been incorporated as a key analytical tool in several fields of chemistry, including medicinal chemistry, agricultural chemistry, environmental chemistry, and toxicology. In pharmaceutical chemistry, QSAR analysis has proven to be a model means for drug invention and is currently a standard component of all industrial drug design software packages [4]. Its recent trends in the drug discovery and development pipeline can be summarized in two points: the first is the development of reliable models that accurately predict and classify the biological responses of potential leads. The second is the application of these models in the design of new chemical entities (NCEs) and the screening of large libraries or datasets of compounds to select new hits with desired attributes. However, in both cases, the predictions should be validated experimentally [42,43]. By selecting promising hits, QSAR analysis reduces the number of costly experiments and thus reduces costly failures of candidate drugs [44].

4.1 Recent QSAR approaches

Since its inception, QSAR has evolved rapidly, leading to the development of several methodologies including 2D-QSAR, 3D-QSAR, 4D-QSAR, and Hologram-QSAR

(HQSAR). Classical 2D-QSAR approaches are based on correlating the physicochemical features of 2D molecular substituents (or fragments) with biological responses. Afterwise, QSAR has improved significantly, with *Cramer et al* [45] proposing comparative molecular field analysis (CoMFA), the first innovative 3D-QSAR approach, in 1988. The CoMFA paved the way for the development of alternative 3D-QSAR methods, such as CoMSIA, SOMFA, and CoMMA, as well as multidimensional (nD)-QSAR methods, such as 4D-QSAR, 5D-QSAR, etc., to address well-known 3D-QSAR challenges like subjective molecular alignment and bioactive conformation problems [44]. Nowadays, three modern QSAR methods have been developed by reformulating the conventional methods cited above with refined mathematical tools to yield well-designed theoretical models [4]:

4.1.1 Fragment-Based 2D-QSAR (FB-QSAR)

This method divides prospective drug structures into various fragments depending on the substitute under consideration. By generating two sets of coefficients identifying the structural fragments and the physicochemical properties, the physicochemical properties of the fragments are then correlated to the biological activities of the prospective drugs [4].

4.1.2 Multiple Field 3D-QSAR (MF-3D-QSAR)

In this method, an additional thermodynamic and non-thermodynamic field of molecular potential has been inserted into the CoMFA approach using two sets of coefficients representing the Cartesian 3D-space position and the potential field [4].

4.1.3 Amino Acid-Based Peptide or Protein Prediction (AABPP)

This method is concerned with the evaluation of peptide/protein activity. Analogously to prior methods, the relationship correlating the biological activities of targets to the physicochemical properties of concerned amino acid residues is given by using two sets of coefficients, one representing the physicochemical properties of concerned amino acids and the other representing the residues in the target chain [4].

4.2 General workflow of QSAR Studies

QSAR process can be generally conducted through three steps: data preparation, data analysis/model development, and model validation [46]. Figure II.2 illustrates the General workflow of QSAR model development. For model development, the chemicals acquired from literary sources could be split into training and test set. The training set is used to build the model, whereas the test set is required for external validation [47]. The success of any QSAR model depends on many factors, including the accuracy of the input data, the selection of relevant descriptors and practical statistical techniques for model building, and the validation strategies. The validity of models is assessed by an internal validation process usually based on the use of the cross-validation method (CV). Whereas, the predictive power can be estimated using independent test data that was not included in the model generation [42].

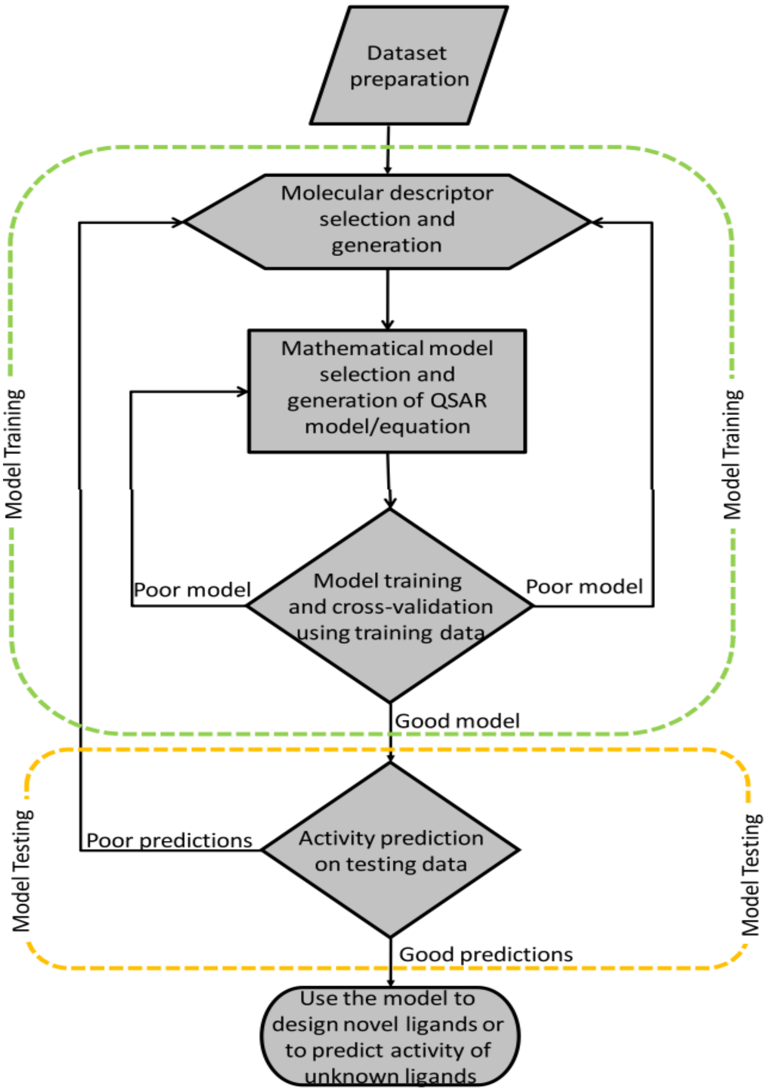


Figure II.2. General workflow of QSAR model development [44].

4.2.1 Data preparation

The first preliminary step in the development of a QSAR model includes the curation of the relevant dataset for study and the computation/selection of molecular descriptors [46].

4.2.1.1 Datasets curation

A wide variety of chemical and biological data collections are available in research papers and publicly or commercially accessible electronic databases, e.g., PubChem, BindingDB, ChEMBL, ChemSpider, NIMH, etc. Hence, the appropriate dataset is not selected randomly; indeed, it is guided by a set of rules, including:

- The dataset should come from the same bioassay protocol to prevent data inconsistencies and inter-laboratory variability, and it is encouraged to use data obtained from uniform source (organism/tissue/cell/protein) or a single lab; since irrelevant data points will damage the proper structure/activity correlation.
- The dataset should be large enough to assure statistical stability of models.
- All compounds of the dataset should have the same mechanism of action and same/comparable binding mode.
- The unit of measurement for activity data should be standardized for all compounds.
- Activities of chemicals should be dispersed uniformly over the range of activity.
- Training sets should include structurally diverse entities to expand the area of study of the required biological response.
- The relationship dose/response should be smooth.
- The potency (or affinity) should be reproducible [46].

Dataset curation implies checking the accuracy, consistency, and reproducibility of the reported experimental data. This step is essential for the reliability of cheminformatics studies, notably QSAR modeling. It can be accomplished using a variety of database curation tools included in commercial and open-source software. The overall process for chemical and biological data curation is depicted in Figure II.3. The main steps for chemical structures curation consists of:

- Correcting structural errors, e.g., ring aromatization, detection of valence violations or extreme bond lengths and angles, while maintaining their proper representations in the data set.
- Eliminate structures consisting of atoms for which there are no coefficients for descriptor computation, structures with multiple fragmentations, salts, confusing structures (inorganics, organometallics, counterions, biologics, mixtures), and duplicate structures.
- Addressing the problem of isomerism.
- Structures exhibiting unreliable target properties, such as high variability in experimental activity value, are removed from the training set [48–50].

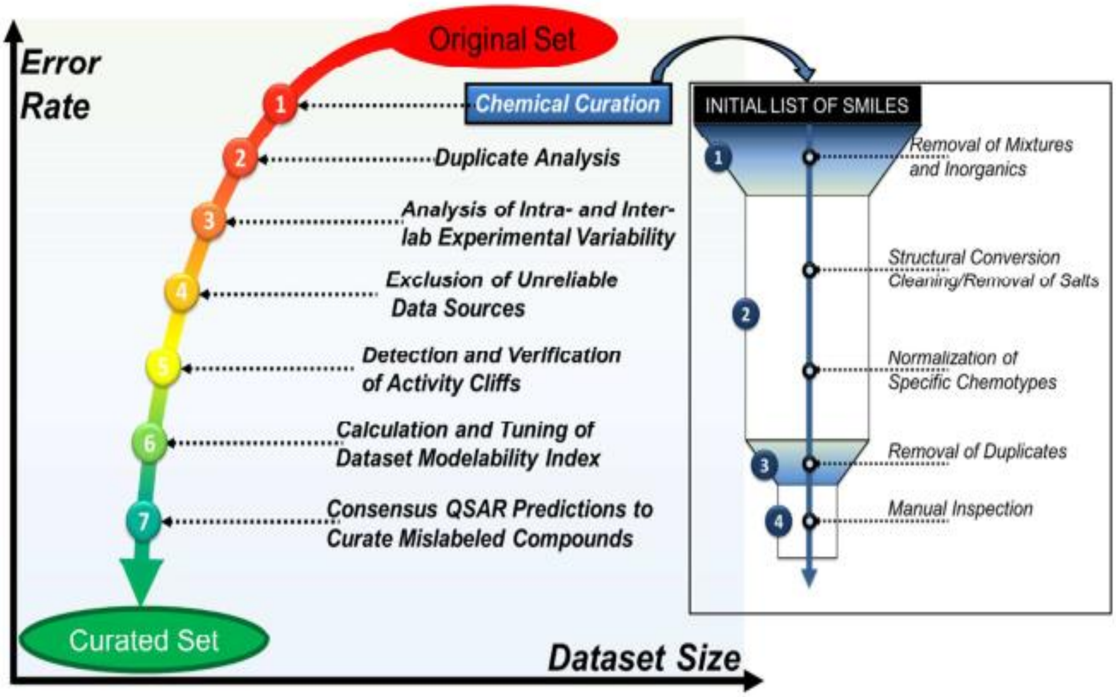


Figure II.3. An integrated workflow for chemogenomics datasets curation [49].

4.2.1.2 Molecular descriptors: selection and generation

Descriptors, variables, or predictors are the numerical forms of the chemical characteristics encoded within a molecular structure. Some of these include topological, geometrical, and constitutional indices, as well as physicochemical parameters (Figure II.4). The information encoded by the descriptors is influenced by two factors: The dimensions of the molecule and the calculation algorithm [47,51].

Figure II.4. Representation of molecular descriptors used in QSAR analysis [51].

4.2.1.2.1 Types of molecular descriptors

Molecular descriptors are classified as follows:

- **0D-descriptors:** are also known as constitutional descriptors. The most prevalent among them are: the number of atoms, bond count, atom type, ring count, and molecular weight (MW).
- **1D-descriptors:** This class contains the physicochemical features of substances as well as their constitutional parameters. These descriptors are independent of molecular topology and are often used in combination with other descriptors, since their values, for example, the number of distinct molecular groups, are vulnerable to being identical between different chemicals.
- **2D-descriptors:** These are frequently used to describe chemical spaces. They are also independent of molecular topology and invariant of molecular graphs, e.g., connectivity indices, information indices, counts of paths and walks, etc.
- **3D-descriptors:** these refer to the geometrical spatial information of chemicals and, thus, provide an accurate description of molecular structures, e.g., CoMFA descriptors and various CoMFA-like descriptors. Despite being computationally costly, these descriptors are useful in classifying isomers.
- **4D descriptors:** these provide all the features of 3D descriptors with the advantage of analyzing multiple structural conformations, simultaneously.
- **Fingerprints or structural keys:** These terms refer to descriptors that take values of zero or one depending on the presence or absence of pre-defined molecular properties (or fragments) such as atoms (oxygen, nitrogen, etc.), rings (aromatic, etc.), bonds (double, triple, etc.), halogens, etc. [4,46].

4.2.1.2.2 Selection of relevant descriptors

Out of numerous available types of descriptors supplied by commercial and open-source software, the selection of related types that reflect structural variation is critical and constitutes a necessary step for deriving a robust QSAR model that satisfies the purpose of the analysis [44]. Descriptor processing entails excluding irrelevant variables (uninterpretable variables, duplicated pairs, correlated pairs, and variables having the same value for all compounds in the dataset) [46]. A statistically significant subset of descriptors should contain a minimum number of descriptors, since most learning algorithms, as in the training algorithm and production steps, become computationally intractable when the number of variables is large. Also, the use of high-throughput data in statistical modeling poses a challenge to accurate prediction. Moreover, due to a large amount of inherent noise and variation in samples and their high dimensionality, the risk of overfitting is always posed [4,48,52]. Thus, descriptor selection assures the interpretability of the final model, increases model performance, and minimizes overfitting. It can be performed through statistical tools involved in molecular modeling programs. The working principle of these tools is based on different algorithms, such as the Genetic Algorithm (GA), which uses the evolutionary rules of natural selection to choose the most suitable subset of descriptors for the studied case [48]. Methods used for descriptors selection are illustrated in the sub-section ‘’Variable selection methods’’.

4.2.2 Data analysis/model development

4.2.2.1 Statistical methods

Model development entails finding the relationship between molecular descriptors and the biological response, through the application of statistical or chemometric techniques. For this purpose, linear and non-linear approaches have been developed and incorporated into various modeling programs to build robust QSAR models. Table II.1 gives an illustration of some of these approaches [41,48].

Table II.1. Statistical methods used in QSAR analysis [41].

Linear Regression Analysis (RA)
Simple linear regression (SLR)
Multiple linear regression (MLR)
Stepwise multiple linear regression Multivariate
Multivariate data analysis
Principal component analysis (PCA)
Principal components regression (PCR)
Partial least square analysis (PLS)
Genetic function approximation (GFA)
Genetic/partial least squares (G/PLS)
Pattern recognition
Cluster analysis
Artificial neural networks (ANNs)
k-nearest neighbor (*k*NN)

If the inputs (compound dataset and descriptors) are of a small number, SLR or MLR are recommended. In contrast, if the number of inputs is enormous, PLS or PCR are the optimal choices. In addition, alternative approaches can be obtained by combining the best features of the approaches mentioned in Table II.1 with various variable selection methods, For example:

- ✓ A combination of genetic algorithm (GA) with multiple linear regression (MLR), PLS, and Artificial Neural Networks (ANNs).
- ✓ A combination of Artificial Neural Networks (ANNs) and MLR/PLS under the name of Polynomial Neural Network (PNN) [48].

4.2.2.1.1 Partial least square analysis

PLS regression is a useful method for multivariate data containing correlated molecular descriptors. This method based on dimension reduction technique, builds orthogonal components, often called factors or latent variables, as linear combinations of the original predictor variables [53]. PLS constructs these components while considering the observed response values, leading to a parsimonious model with reliable predictive power [54].

4.2.2.2 Variable selection methods

Variable selection methods have been developed to specify the reliable subset of independent variables that must be entered into the analysis to ensure the best prediction of the outcomes. Typically, calculating all possible regression models that include all possible combinations of variables, then among these models, selecting the best model according to statistical criteria is considered the most efficient method.

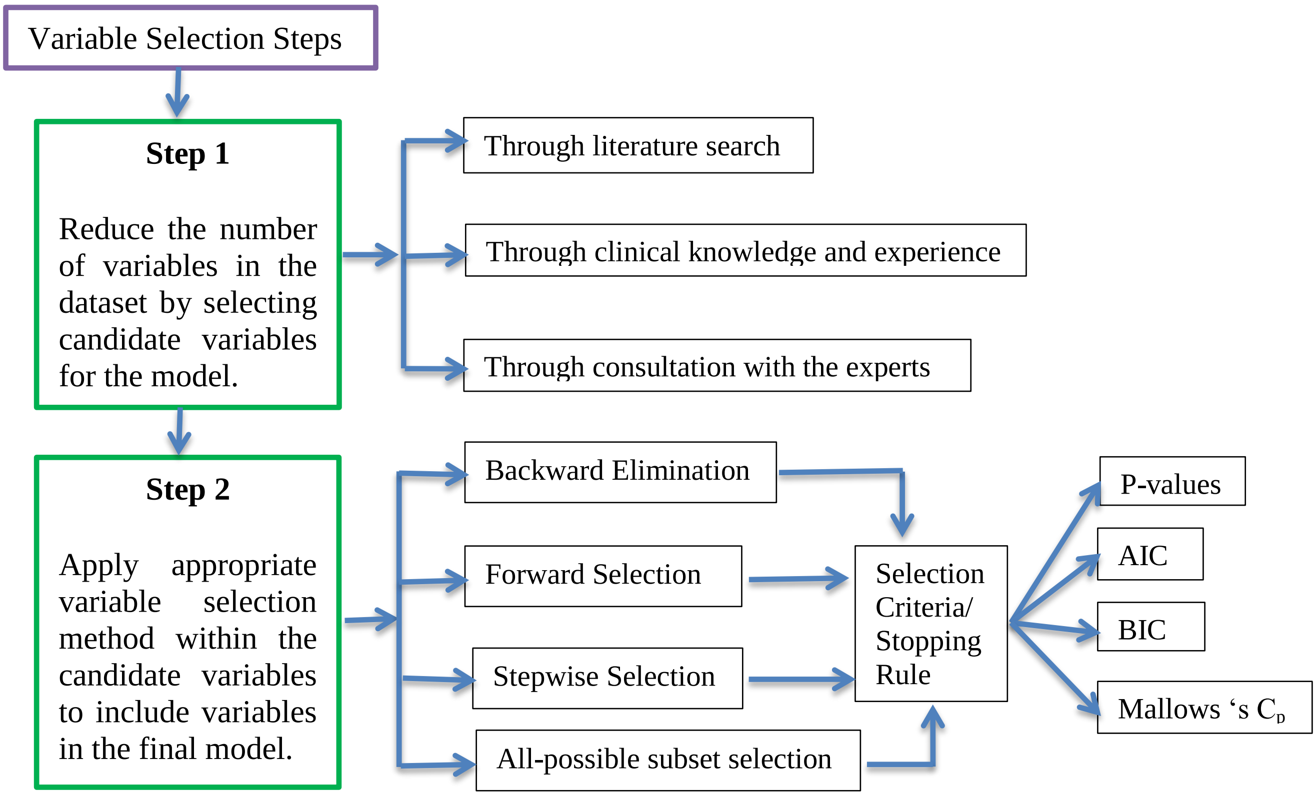


Figure II.5. Variable selection steps [55].

Unfortunately, even with modern ultra-fast computers, this method still suffers from computational drawbacks that consume time, effort, and money whenever the number of input variables is large. To deal with these shortcomings, efforts have been devoted to developing computing algorithms that take advantage of previously performed computations of subsets, for example, the leaps-and-bounds algorithm of *Furnival and Wilson* [56]. Currently, Variable selection methods involve two approaches: The best subset selection and stepwise regression selection [55]. Variable selection steps are illustrated in Figure II.5.

4.2.2.2.1 Best subset selection

This method is based on generating all 2^P possible models from different combinations of **P** independent variables (fit all models with one variable, all models with two variables,..., all models with P variables). Afterwise, the potential models are reduced from 2^P to (P+1) by selecting the best model for each subset size depending on the percentage of variance (R^2) or the residual sum of squares (RSS) values, with the best-fit model being the one with the highest R^2 or the lowest RSS value. Finally, the overall best model among these (P+1) selections was identified using statistical criteria such as cross-validated prediction error, Mallow’s Cp (AIC), BIC, or adjusted R^2 [57].

Despite its simplicity, best subset selection suffers from computational and statistical limits when the number of models under consideration is large. The large search space enhances the probability of finding better models on the training dataset, but it can also lead to overfitting and high variance of the predicted coefficients. Later, stepwise approaches have been proposed as viable solutions to circumvent these limitations [58].

4.2.2.2.2 Stepwise methods

Stepwise methods construct the best-fitting model in several iterative steps, each step requiring either adding or excluding one variable at a time according to its statistical significance [59]. Stepwise methods involve three variable selection methods used in regression: forward selection, backward elimination, and stepwise selection [58].

4.2.2.2.2.1 Forward selection

Among the 2^P possible models, forward selection requires generating only $1 + \sum_{K=0}^{P-1}(\mathrm{P} - \mathrm{K}) = 1 + \mathrm{P}(\mathrm{P} + 1)/2$ models, where **K** denotes the degree of iteration [57]. The selection procedure proceeds by sequentially adding variables to the null model based on their statistical significance. The first step focuses on determining the first relevant variable to enter the model. Hence, it is necessary to generate all the simple linear regressions that correlate each of the P independent variables and the outcome variable, and from which calculate the bivariate r^2 values. Next, select the variable that displays the highest value of r^2 to be the first element to enter the null model. In the following steps, the remaining variables are added, successively, to the obtained model according to their impact on its R^2 or RSS, where, the most significant variable is the one that produces the largest increase in R^2 or the largest decrease in RSS. By default, adding variables continues until all variables have been included in the model or a stopping role is reached [57,58].

Forward selection is an effective alternative to best subset selection. Nevertheless, it does not ensure the selection of the best subset of variables out of all the 2^p possibilities. For instance, assume a sample of three variables, X_1, X_2, and X_3, where the best size-one subset contains X_1 and the best size-two subset contains (X_2 and X_3). As per the forward selection process, if the best subset of size-one involves X_1, the best subset of size two should also involve X_1 along with one other variable. Here, forward selection fails to recognize the true best subset of size-two. Since it starts from the null model, the forward selection is the most appropriate choice in the case where the number of variables exceeds the sample size [57].

4.2.2.2.2.2 Backward elimination

Similar to the forward selection, backward selection is an effective alternative to best subset selection. It allows building $1 + P(P + 1)/2$ models. Nevertheless, it does not ensure the selection of the best subset of variables out of all the 2^p possibilities. Since it starts from the full model, the backward selection is the optimal choice in the case where the sample size exceeds the number of variables. Overall, its selection procedure begins with the full model and then performs successive removal of the independent variables as a function of their statistical significance. The variable suitable for deletion at each step, its deletion should cause the slightest decrease in R^2 or the slightest increase in RSS of the model under investigation. This process terminates when all variables are removed or a stopping role is attained [57].

4.2.2.2.2.3 Stepwise selection

Nowadays, statistical software allows users to merge forward selection and backward elimination into a single hybrid strategy that retains both of its computational features; the stepwise selection. While this alternative is more computationally intensive than forward and backward selections, it offers the advantage of checking more candidate subset models before settling on a model for each subset size. In this context, it considers the impact of adding or deleting a variable on the contributions of other variables to the model. The selection procedure begins by sequentially adding variables to the null model, and whenever a new variable is incorporated, the relevance of all previously included variables is re-checked. If any of the included variables no longer have any contribution to improving model fit, then backward elimination is launched. Accordingly, the variables are excluded, by one at each step, until all the remaining variables meet the minimum statistical requirements to keep in the model. Following that, the forward selection is restarted in a new iteration. This hybrid process ends when it reaches a stopping role or when no more variables can be added or removed [55,60].

4.2.2.2.2.4 Stopping rules

Stopping roles are used during the implementation of stepwise methods to specify when the selection process should end. Out of the numerous stopping rules available in the literature [61], Rules based on Student test (t-test) are the most applied in stepwise methods to test the hypothesis that input variables should be added to, or removed from, the model. Since the Fisher-test (F-test) follows the t-distribution, it can be used as an alternative to t-test when comparing two models. [62]. Its arithmetic form is given by:

$$F = \frac{\frac{SSE_r - SSE_f}{df_r - df_f}}{\frac{SSE_f}{df_f}} \tag{31}$$

SSE is the sum of squared errors, defines as:

$$SSE = \sum_{i=1}^{n_0} (y_i - \hat{y}_i)^2 \tag{32}$$

Accordingly, $\mathrm{SSE_r}$ and $\mathrm{SSE_f}$ are the sum of squared errors of the reduced and full models, respectively. During forward selection the reduced model contains all of the input variables currently in the model and the full model contains these variables and the input variable under consideration. During backward elimination the reduced model contains all of the variables currently in the model except the variable under consideration and the full model contains all of the variables currently in the model. Similarly, df_r and df_f are the degrees of freedom of the reduced and full models, respectively. The degrees of freedom of a model are equal to the difference of the number of observations n_o and the number of variables in the model $n_\emptyset$. The difference $df_r - df_f$ is called the numerator degrees of freedom and df_f is called the denominator degrees of freedom. In eq (32), y_i is the i^{th} observation and $\hat{y}_i$ is the i^{th} model output [63].

F-statistic requires two user-specified thresholds to enter, or exclude, variables from the model. The model-entry criterion is expressed by the term "F-to-enter" or "significance level to enter (SLE)". Similarly, the model-retention criterion is expressed by the term "F-to-stay" or "significance level to stay (SLS)" [55,62] .Thresholds can be a fixed value or determined through statistical criteria such as AIC or BIC [60]. At each step of forward selection, the F-statistic values of all the out-of-model variables are calculated, and then the largest F-statistic value among them is selected and compared with the F-to-entry value. If the largest F-statistic is greater than or equal to F-to-enter (or if the lowest p-value is less than the F-to-enter value), then the corresponding variable is added to the model. This process continues until no remaining variable satisfies the input criterion. In parallel, in backward elimination, if the smallest F-statistic value of the variables within the model is less than the F-to-stay value, then the corresponding variable is eliminated from the model. This process terminates when all variables remaining in the model satisfy the criterion to stay. On the other side, to achieve a stopping rule in the stepwise selection, both the F-to-enter and F-to-stay criteria must be met. In this context, variables continue to be added and dropped until none of the out-of-model variables have an F-value greater than F-to-enter and none of the incorporated variables have an F-value smaller than F-to-stay. Typically, F-to-enter and F-to-stay criteria are not required to be identical, however, to prevent an infinite loop in the stepwise selection, F-to-enter should be lower than F-to-stay [55,57,59].

4.2.2.2.3 Choosing the optimal model

While implementing best subset selection and stepwise methods, the RSS and R^2 statistics, given in equations (33) and (34) respectively, are reliable criteria for determining the best model among a subset of models of the same size (containing the same number of variables). However, they become computationally impractical when selecting the best overall model among a set of models of different sizes [57].

$$RSS = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{33}$$

Where, y_i is the ith observed response value, and $\hat{y}_i$ is the ith response value predicted by the linear model.

$$R^2 = 1 - \frac{RSS}{TSS} \tag{34}$$

Where, RSS is the residual sum of squares defined in eq (33), and TSS is the total sum of squares; $TSS = \sum(y_i - \bar{y})^2$, $\bar{y}$ is the sample mean $\bar{y} \equiv \frac{1}{n}\sum_{i=1}^{n} y_i$.

Since these statistics are associated with training error, RSS decreases and R^2 increases as even more variables are added to the models. In light of this, it stands to reason that the model with all the variables will logically show the lowest RSS value and the highest

R^2 value. A poor training error, on the other hand, does not ensure a poor test error. So, when needs to define the model with the low test error, other goodness-of-fit criteria are used [55,57]: Mallow's Cp, Akaike information criterion (AIC), Bayesian information criterion (BIC), and Adjusted R^2.

4.2.2.2.3.1 Cp, AIC, BIC, and R^2_{Adj}

In the following, the arithmetic formulas for Cp, AIC, BIC, and $\mathrm{R}^2_{\mathrm{Adj}}$ are presented in the case of fitted least squares model with ''d'' variables and ''n'' observations:

$$C_p = \frac{1}{n}(RSS + 2d\hat{\sigma}^2) \tag{35}$$

$$AIC = \frac{1}{n\hat{\sigma}^2}(RSS + 2d\hat{\sigma}^2) \tag{36}$$

$$BIC = \frac{1}{n\hat{\sigma}^2}(RSS + \log(n)d\hat{\sigma}^2) \tag{37}$$

$$R^2_{Adj} = 1 - \frac{RSS/(n-d-1)}{TSS/(n-1)} \tag{38}$$

Where, $\hat{\sigma}^2$ is an estimate of the variance of the error ϵ associated with each response measurement in the standard linear model, RSS is the residual sum of squares, TSS is the total sum of squares, and (n − d – 1) is the degree of freedom [57].

For least squares models, Cp and AIC are proportional to each other. Although BIC is derived from a Bayesian perspective, it is also similar to CP and AIC. Cp, AIC, and BIC statistics exhibit a small value for a model with a low test error. Hence, when identifying any set of models is the best, it must choose the model that contains the lowest CP, AIC, or BIC value [60]. Unlike Cp, AIC, and BIC, a high value of R^2_{Adj} implies a model with a low test error [57].

4.2.2.3 Check for Outliers

The occurrence of outliers is a common issue while developing a QSAR model. The term "outlier" refers to chemicals that do not adhere to the fundamental principle underpinning all QSAR studies: compounds with analogous structures have similar bioactivities or properties [64]. In other words, outliers are compounds having unexpected biological activity, acting on the target via distinct mechanisms or modes, and are unable to fit into a QSAR model [65]. Thus, Outliers in the dataset, if not eliminated during model construction, might affect predictions and damage reliability.

Outliers result from several reasons, including the use of inappropriate parameter values, experimental and typographical errors, the implementation of an inadequate mathematical model, the presence of activity cliffs, the flexibility of the binding site, the errors that can arise during QSAR development, such as over-fitting or inappropriately fitting data [66], In addition to other reasons mentioned in the research of *Dearden et al* [67].

4.2.2.3.1 Types of outliers

Outliers observed in QSAR studies involve three types:

- **Activity (or the dependent variable Y-direction) outliers:** are the points that deviate from the normal distribution of dependent variable-**y**, for example, point 1 in Figure II.6. This type results in a large sum of squares error. If there are few outliers and no masking effect, strong regression algorithms should easily deal with this problem [64].
- **Leverage (or structural) outliers:** These are points that deviate from the main distribution of the samples in the independent variable X-direction. Since their x-value is outlying, points 2 and 3 Figure II.6 are examples of leverage points. Point 2 is a good leverage point with no significant sum of squares error, but point 3 is a bad leverage point. When the datasets are used to build the model without deleting the leverage values, a negligible variation can result in large fluctuation for this model [64].
- **Model outliers:** are special sorts of outliers that arise in large-scale datasets as a result of the diverse molecular structures used in the study. They reflect a unique relationship between X and Y. Points marked with 4 in Figure II.6 are outliers toward the model that are outliers in both Y and X at the same time [64].

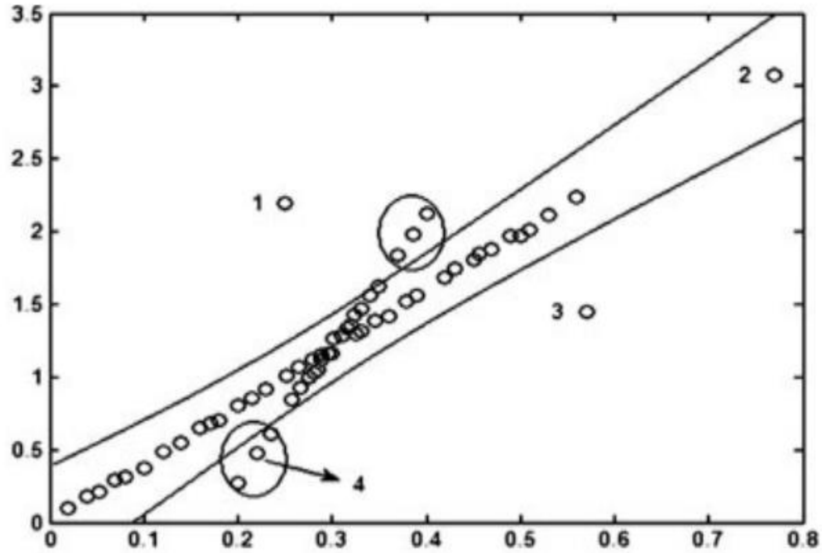


Figure II.6. Different types of outliers in simple regression: (1) Y-outlier, (2) good leverage point, (3) bad leverage point, (4) model outlier [64].

Mostly, the three types of outliers coexist in models. Hence, a robust method for detecting outliers must be able to recognize them exhaustively and concurrently within a multivariate point cloud. Classical identification methods are not always effective since they depend on the sample mean and the covariance matrix. Nowadays, new methods based on new strategies have been developed and proposed to prevent the impotence of classical procedures [64].

4.2.2.3.2 Studentized deleted residual and leverage value

Outlier diagnostics in multiple regression models can be carried out by computing leverage values (h_{ii}) for identifying outlying x-variables and studentized deleted residuals ($\mathrm{r}_{\mathrm{i}}^{*}$), also known as the externally studentized residual, for identifying outlying y-variables.

For a simple linear regression, h_{ii} defined as:

$$h_{ii} = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum_{i'=1}^{n}(x_{i'} - \bar{x})^2} \tag{39}$$

Where, x_i is the ith variable, $\bar{x}$ is the sample mean; $\bar{x} \equiv \frac{1}{n}\sum_{i=1}^{n} x_i$.

The values provided by h_{ii} vary from 1/n to 1. An observation is considered an outlier if its h_{ii} value exceeds the value of the average leverage of all the observations (p+1)/n [57].

r_i^* of the ith observation is determined as follows [68]:

$$r_i^* = \frac{y_i - \hat{y}_i}{s_{(i)}\sqrt{1 - h_{ii}}} \tag{40}$$

Where, y_i is the ith observed response value, and $\hat{y}_i$ is the ith response value predicted by the model, $s_{(i)}$ is the estimated variance, and h_{ii} is the leverage value.

For r_i^*, two critical levels are suggested: ±2 (5% risk) and ±3 (0.27% risk, typical in process control charts) [68].

4.2.3 Model validation

Given the essential role that QSAR analysis has occupied over the past decades in medicinal chemistry and computer-aided molecular design, any QSAR modeling should eventually lead to statistically robust models capable of estimating the potential effects of chemical entities, materials, and nano-materials on human health and ecological systems, as well as, yielding accurate and reliable predictions of biological response. Thereby, the validation step is critical before any prospective applications of the generated model. In this sense, the scientific community has set up two validation strategies to check the statistical fit and predictability of the selected models: internal validation and external validation [69].

4.2.3.1 Internal validation

Internal validation is carried out by utilizing the data points used in model generation. The most common internal methods are Randomization tests, Cross-Validation (CV), and Bootstrapping [42]. Among them, the leave-one-out (LOO) technique from the CV procedure is the most applied in QSAR modeling [70]. This technique is based on randomly removing one data point from the data set in each cycle and using the remaining data points of the training set to create the model. The latter is used to predict the activity of the removed data point. This procedure is continued until all data points have been deleted [71]. The CV outcomes allow suggesting several correlation-based metrics to assess the internal validity of QSARs such as the coefficient of determination (R^2) and similar measures, the adjusted coefficient of determination (R^2_{Adj}) [72], considered in this work.

4.2.3.1.1 R^2 and R^2_{Adj}

The arithmetic formula for R^2 and R^2_{Adj} were previously reported in eqs (34) and (38). The term Q^2 is often used in instead of R^2 when referring to techniques based on sample reuse, such as CV_{LOO} and CV_{LMO}. R^2 is an estimation of the percentage of the variance between the observed response and independent variables explained by the regression equation. Due to its being independent of the number of variables in the data set, R^2_{Adj} was developed as an adjustment of R^2. As previously stated, it is useful when comparing a set of models of different sizes; the model with the highest R^2_{Adj} is ranked the better [72–74]. Statistically, both R^2 and R^2_{Adj} yield values between 0 and 1, with R^2_{Adj} always being less than or equal to R^2. For a significant QSAR model, the value of R^2 must be greater than, or equal to, the minimum recommended value of 0.6 (The closer the value of R^2 to 1, the better the model). Moreover, the number of variables included in a QSAR model is considered appropriate if it achieves the relationship: $(R^2 - R^2_{Adj}) < 0.3$ [71].

4.2.3.1.2 F-statistics

Fisher-statistics (F) is optimal to judge the overall significance level of the regression model. The F-statistic can be easily obtained from equation (41). it is crucial to highlight that the assumption that "the higher the F-value, the better the model" is only valid when comparing F-values with the same degrees of freedom [71,75,76].

$$F = \frac{(n-p-1)}{n} \frac{\sum_{i=1}^{n}(y_{i\,exp} - \bar{y}_{TS})^2}{\sum_{i=1}^{n}(y_{i\,exp} - y_{i\,pred})^2} \quad (41)$$

Where, **n** is the number of molecules in the Training Set (TS), **p** is the number of descriptors in the model. $y_{i\,exp}$ and $y_{i\,pred}$ are the i^{th} observed and predicted values of biological activity, $\bar{y}_{TS}$ is the average value of biological activity for the training set.

Moreover, the F-statistic can be used in hypothesis testing. In this context, the null hypothesis (H_0) states that the developed model does not outperform the average activity value in prediction. The H_1 hypothesis, on the other hand, imposes the opposite. The acceptance or rejection of H_1 at a specific significance level α (typically 0.95 or 0.99) is determined by the F-distribution function. In Brief, H_1 is accepted if the F-ratio value seems to be higher than the corresponding value of the F-distribution function for the considering degrees of freedom. Otherwise, it is rejected [70].

Determining the boundary significance level α between H_1 and H_0 is another alternative to test the predictive power of the QSAR model where, the higher the α value, the better the model. This is obtained by solving the following equation:

$$F_{dfs,\ \alpha} = F \quad (42)$$

Where, $F_{dfs,\ \alpha}$ denotes the F-distribution function with the considered degrees of freedom (dfs). For good models, the value of α is close to 1, so most of the time P-values were used instead [70].

4.2.3.2 External validation

External evaluation is carried out by an independent dataset of test structures not used in model generation. Predictions using an independent external dataset ensure the robustness of the developed QSAR model for the prediction of unidentified chemicals. Its process is based on the comparison between observed and predicted responses [71]. Recently, three effective Q^2-like formulas have been proposed to evaluate the external prediction of developed models: Q^2_{F1}, Q^2_{F2}, and Q^2_{F3}.

4.2.3.2.1 Q^2_{F1}, Q^2_{F2}, and Q^2_{F3}

Q^2_{F1} and Q^2_{F2} metrics are estimated by dividing the predictive sum of squares ($PRESS$) by the sum of squares of the external data set (SS_{EXT}) approximated by the training set response mean ($\bar{y}_{TR}$) and the test set response mean ($\bar{y}_{EXT}$), respectively [77]. Equations below provide the basic expressions:

$$Q_{F1}^2 = 1 - \frac{\sum_{i=1}^{n_{EXT}} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n_{EXT}} (y_i - \bar{y}_{TR})^2} = 1 - \frac{PRESS}{SS_{EXT}(\bar{y}_{TR})} \quad (43)$$

$$Q_{F2}^2 = 1 - \frac{\sum_{i=1}^{n_{EXT}} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n_{EXT}} (y_i - \bar{y}_{EXT})^2} = 1 - \frac{PRESS}{SS_{EXT}(\bar{y}_{EXT})} \quad (44)$$

The characteristics, drawbacks, and limitations of these functions in external evaluation are currently well documented in the literature [72,78–81]. As a result, another alternative function, Q_{F3}^2, has been proposed that is based on the ratio of the mean square error of the external evaluation set over the mean squared deviation of the training set [77]:

$$Q_{F3}^2 = 1 - \frac{\frac{\left[\sum_{i=1}^{n_{EXT}} (\hat{y}_i - y_i)^2\right]}{n_{EXT}}}{\frac{\left[\sum_{i=1}^{n_{TR}} (y_i - \bar{y}_{TR})^2\right]}{n_{TR}}} = 1 - \frac{\frac{PRESS}{n_{EXT}}}{\frac{TSS}{n_{TR}}} = 1 - \frac{msep}{s^2} \quad (45)$$

Where, $PRESS$ denotes the predictive error sum of squares, TSS is the total sum of squares, $msep$ is the mean square error in prediction and s^2 a biased estimate of the training set response variance.

These three metrics yield simple and easily interpretable estimates, where the closer the value to 1, the better [72].

5 Molecular docking

In organism, two chemical elements can interact in many ways, the most common being the protein-protein or protein-small molecule interaction. Overall, molecular docking makes it possible to estimate the intermolecular framework generated between a macromolecule and a small molecule or between two macromolecules, as well as identify the binding modes required for the target regulation. Classically, the notion of molecular recognition of ligand/target complex was analogous to a "lock and key". Now, this terminology has been refined to include the target-flexibility and the mutual adaptation with the ligand to become analogous to a "hand and a glove" (Figure II.7).

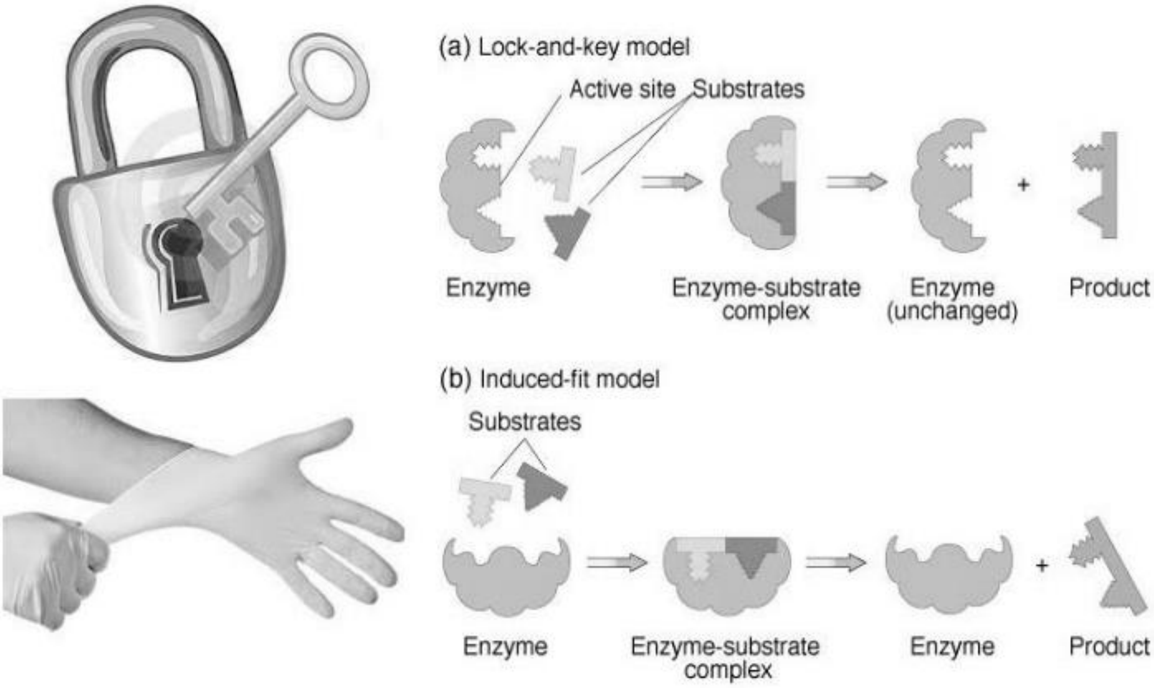


Figure II.7. The classical analogy of ligand/target complex as "lock and key" and the current analogy as "hand and glove" [82].

For reliable docking experiments, a high-resolution X-ray, NMR, or homology-modeled structure with a known or expected binding locus in the target is required. Docking approaches fit a ligand into a binding site by integrating and optimizing factors such as steric, hydrophobic, and electrostatic complementarity, as well as evaluating binding free energy (scoring) [83–85].

5.1 General protocol

Molecular docking is accomplished through two pathways: first, the sampling of all possible ligand poses within the target binding site, then associates each pose with a score value that approximates its free energy landscape. At the ends of simulations, the best binding modes are ranked based on the values of the latter, and thus the most appropriate complexes are selected. Ideally, sampling algorithms should be able to duplicate the experimental binding mode and the scoring function (SF) should also rank it as the best of all generated conformations [86]. Figure II.8 summarizes the general workflow of molecular docking simulation.

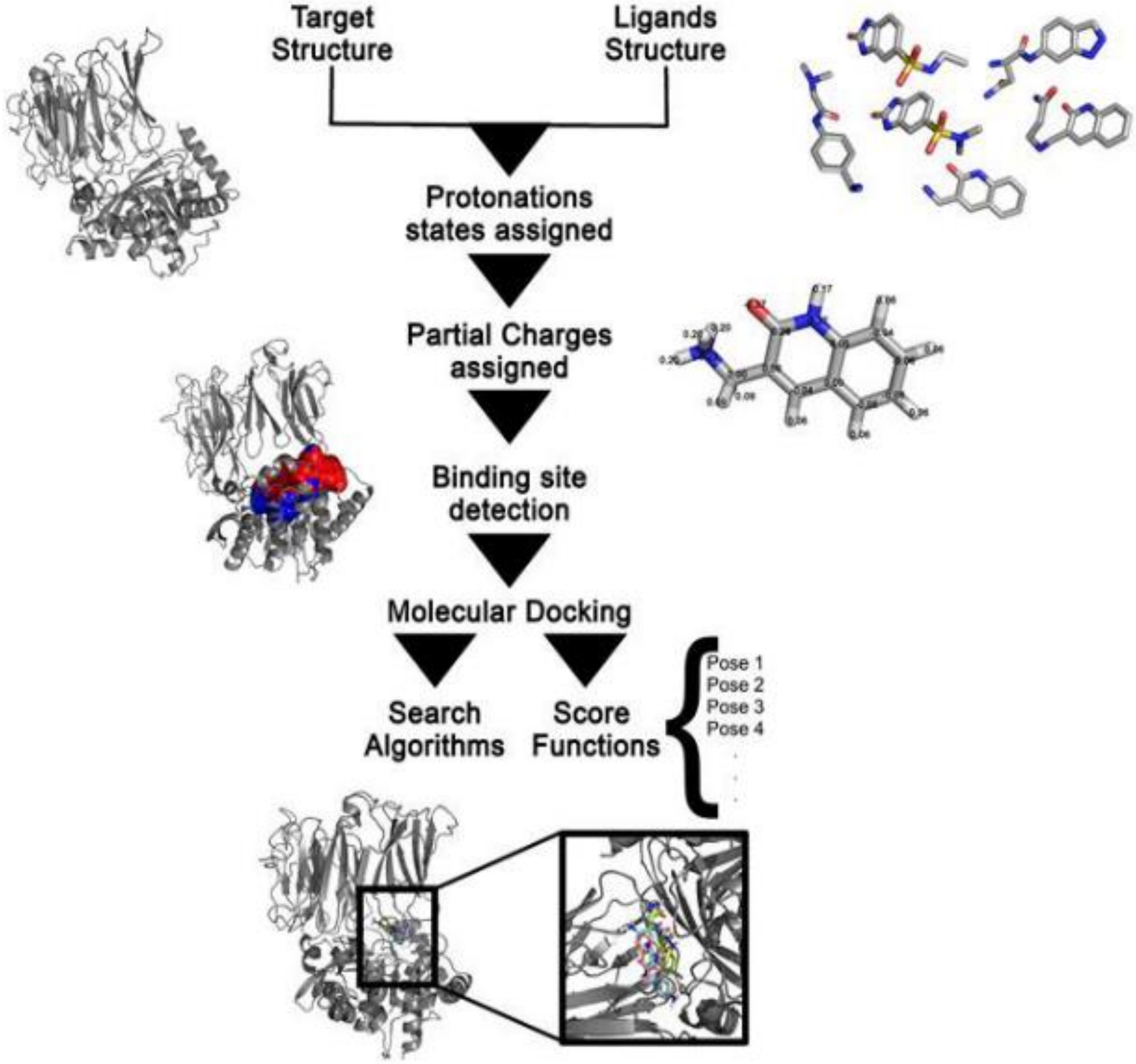


Figure II.8. General workflow of molecular docking simulation [87].

5.1.1 Ligand preparation

Include generating, optimizing, and analyzing its 3D structure. Several accessible drawing software including ChemSketch, ChemDraw, Avogadro, and others, may be used to manually build 3D structures of ligands, either from their 2D structures or from simpler representation schemes, such as SMILES. Besides, they can be directly retrieved from virtual databases like the Cambridge Structural Database (CSD), Available Chemical Directory (ACD), MDL Drug Data Report (MDDR), PubChem, etc. These databases provide a large number of conformations for virtual screening [85,87].

5.1.2 Target preparation

Include minimizing, correcting, and protonating its 3D structure. The structures of targets can be retrieved as PDB files format from several Protein Databases such as the RCSB-Protein Data Bank (https://www.rcsb.org/). Protein databases provide access to 3D atomic coordinates of target conformations characterized by experimental and/or computational prediction methods (such as X-ray crystallography, nuclear magnetic resonance (NMR), infrared spectroscopy, homology modeling, electron density, etc) [85,87].

5.1.3 Binding site detection

Typically, the binding site concerned by the docking simulation is already known and just assigned by the docking software. Nevertheless, when the binding location information is unavailable, the most likely locus can be predicted algorithmically or by implementing a so-called "blind docking". Because the latter covers the whole target surface, it has a high computational cost [85,87].

5.1.4 Docking validation

If the docking procedure regenerated the same binding mode that was created and defined using experimental and/or computational methodologies, it is considered valid. Measuring Root-Mean-Square Deviation (RMSD) is useful for comparing the structural similarity between two superimposed configurations. Its formula is given in eq (46)[88]. Also, RMSD and scoring function values (S_{score}) are used to assess and rank the stability and affinity of ligand/target complexes [85,87].

$$RMSD = \sqrt{\frac{1}{n}\sum_{x=1}^{n} d_x^2} \tag{46}$$

5.2 Types of molecular docking

5.2.1 Rigid docking

In this type, both the ligand and the protein are treated as rigid objects which limiting the search space by taken into account only three translational and three rotational degrees of freedom. The flexibility of ligand might be handled by employing a pre-adjusted list of ligand conformations, or by enabling for atom-atom interaction between both the target and the ligand [86]. This approximation can simulate the "lock-key" binding mechanism and it is used in cases where the number of conformational degrees of freedom is too high to sample, often for protein-protein docking. In general, the target site and ligand are modeled by "hot" points, and the superposition of matching points is assessed [89].

5.2.2 Semi-flexible docking

Since molecular systems are flexible in nature and may alter each other's structure upon interaction, the flexible behavior must be considered in the docking process to enable both ligands and receptors to adapt their conformations and produce the appropriate complex that achieves the lowest energy perfect-fit. However, to avoid the extremely high cost and reduce the computational time, most docking software incorporate semi-flexible docking by assuming that the fixed protein conformation may correspond to the one able to recognize the ligands to be anchored [86]. However, this assumption is not always supported. This docking technique treats ligands as flexible entities while the receptor remains rigid during docking. Also, the search space took into account six translational and six rotational degrees of freedom, as well as sampling all the conformational degrees of freedom of the ligand [90].

5.2.3 Flexible-flexible docking

This docking technique freely anchors the flexible ligand into a flexible receptor considering that the intrinsic kinetics of the protein is strongly correlated to the orientation of the ligand in the binding locus. Implementing receptor flexibility is a significant step forward in the field of molecular docking. Indeed, MD simulation may ensure the modeling of all degrees of freedom in the ligand-receptor combination. Nevertheless, this process suffers from several drawbacks that limit its applications in examining a large chemical library, such as the problem of insufficient sampling and high computing costs [86]. The large size and the complexity of the protein structure limit the possibility to fully include its mobility during the docking process. Thus, studies performed on molecular docking are usually limited to specific residues. Recently, programs that include receptor flexibility, at least in part, have begun to emerge [91]. Accordingly, several strategies for implementing receptor flexibility are available, and they may be categorised based on target conformation as in Figure II.9.

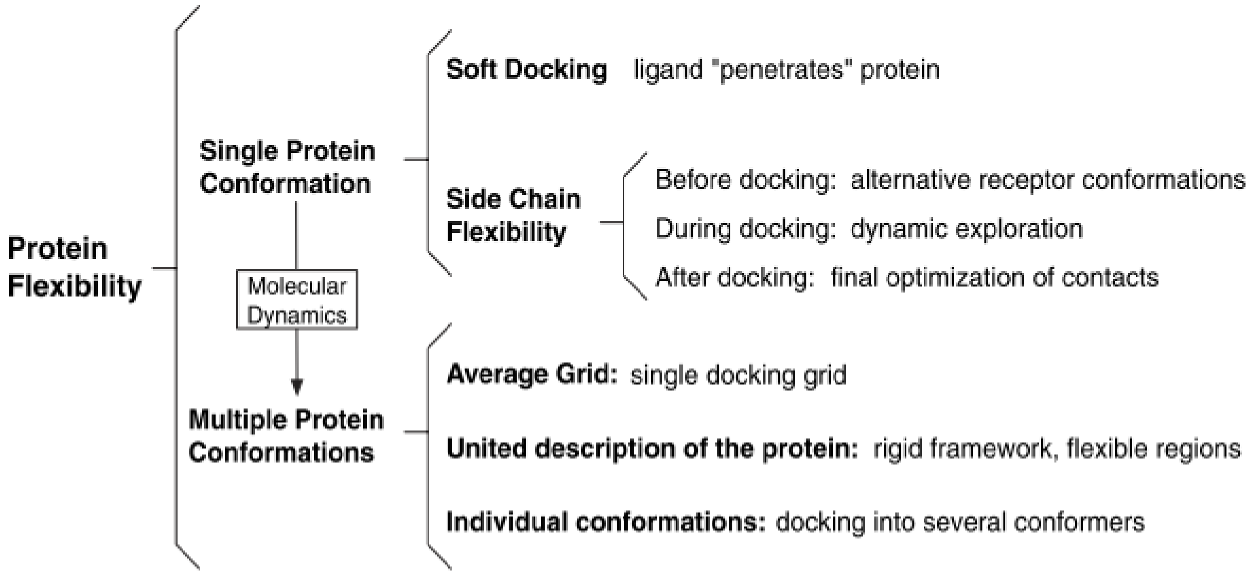


Figure II.9. The described strategies for including receptor flexibility in docking simulation [91].

5.2.3.1 Single target conformation

5.2.3.1.1 Soft docking

Jiang and Kim first described this approach [92], which is based on an implicit and rough evaluation of protein flexibility. Although the receptor is considered a rigid entity, creating a soft interaction may diminish the van der Waals repulsion factor used in force field SFs. As a consequence, allowing the ligand to permeate the surface of the target to some extent and considering for small and localized modifications that may occur in a flexible environment. The advantages of adopting soft docking include that it is easier to execute, inexpensive, and considerably faster than an explicit sampling of multiple receptor conformations. As a disadvantage, this method can only tackle small conformational alterations and may involve unreal poses [86,90].

5.2.3.1.2 Side-chain flexibility

This technique enables the explicit manipulation of the movement of certain residues, particularly those within the target binding site. It offers alternative conformations for protein side chains by using databases of rotamer libraries. The latter contains a set of pre-determined side-chain conformations that are typically collected from the statistical analysis of structural experimental data. Allowing flexibility in the side chains of the residues under-study has the advantage of enabling unrestricted mobility of this protein region, which improves ligand fit. However, because only the side chains of chosen residues are permitted to move, substantial changes in the protein backbone, such as those involved in loop motions, are overlooked [86,90].

5.2.3.2 Multiple protein conformations

The basic idea behind multiple receptor conformations docking is to dock the ligand, sequentially, into an ensemble of rigid receptor conformations rather than one, and then merging the outcomes depending on the approach adopted. For this purpose, experimental or computational techniques can be employed to provide the requisite multiple target conformations [86].

5.2.3.2.1 Combined target grid (average grid)

The docking procedure utilized grids, or pre-calculated two-body potentials, that were often concentrated around the binding site, to quickly and cost-effectively quantify the interaction energy of multiple ligands and binding site conformations. The multiple grids provided by various receptor conformations can be merged into a single global grid using a combination procedure; the simple average-weighted scheme or the differential-weighted scheme. Following that, a single average docking grid representing the target macromolecule was derived. The differential-weighting scheme promotes the participation of particular conformations over others, resulting in better outcomes than a simple average combination of structures [91].

5.2.3.2.2 United description of the target

This technique extracts the rigid conserved parts from the alternative conformations and used them to generate an average rigid structure that fuses with the alternative flexible parts of the ensemble to create a pool of “chimeras” that are used for docking. During the simulation process, ligand structure is constructed progressively within the binding site, and after each new fragment is incorporated, all potential interactions between the partially formed ligand and the alternative target conformations are assessed. Then, the target conformations that best fit the partially grown ligand are saved for further cycles of growth and optimization [90,91].

5.2.3.2.3 Individual conformations

This technique considers the structures of the ensemble as adequate conformations for ligand binding. Thus, multiple docking runs are carried out to assess the interest ligands on all of the target conformations. Additionally, the ensemble of structures can be filtered using a preliminary benchmark that evaluates the performance of different target structures in a cross-docking experiment [90].

5.3 Search algorithms

Search algorithms are crucial tools to docking programs. They are used to explore the free energy landscape that accompanies the best ligand poses at the target-binding site, keeping in view the roto-translational and internal degrees of freedom of the ligand. Theoretically, if the energy function accurately models the thermodynamic compartment of the studied system (the enthalpic and entropic impacts), then the global minimum of the energy landscape corresponds to the native ligand-target binding mode, while the local minima indicate the alternative binding modes. Unfortunately, accounting for the entropic impacts poses some challenges leading current docking approaches to rely only on rough approximations. As a consequence, the global minimum associated with the energy landscape explored by a docking approach does not ensure that it matches the native binding mode [87,93].

5.3.1 Ligand sampling

When discussing sampling algorithms used to adapt ligand flexibility in molecular docking programs, three broad categories can be distinguished: systematic, stochastic, and deterministic search strategies. Moreover, it is worth mentioning that some algorithms use a hybrid technique that incorporates two or all of these three strategies [94,95].

5.3.1.1 Systematic algorithms

During computations, this category investigates all ligand degrees of freedom. It may also be divided into three sub-categories: exhaustive, incremental construction, and conformational ensemble. For example, both FlexX and eHits combine fragment-based techniques with systematic algorithms (incremental construction and graph matching, respectively) [87].

- **In exhaustive searches**, the values of each degree of freedom are systematically explored in a combinatorial way. In addition, all dihedral angles rotate according to a predefined interval of values and a set of preliminary limitations such as geometrical and chemical constraints. Obviously, the more flexible ligands will exhibit a higher number of rotatable bonds, which significantly complicates the optimization process [93].
- **In incremental construction**, the ligand is split from rotatable bonds into smaller fragments. From among these, a base fragment is selected and docked into the target binding site. At the end of the docking process, the ligand is progressively rebuilt by covalently attaching the other fragments to the basic group. This technique is commonly used in *de novo* ligand design, which attempts to discover new entities by connecting the best fragments docked inside the receptor binding site. In addition to the FlexX program, also DOCK 4.0, SLIDE, eHiTS, SKELGEN, ProPose, PatchDock, MacDock, FLOG, Surflex, etc implement the fragmentation approach [83,86,94,96].
- **In conformational ensemble strategy**, a pre-generated set of ligand conformations rigidly docks into the target binding site. This technique is improved to solve the combinatorial explosion problem observed in exhaustive and incremental construction approaches. Examples of docking programs that execute this strategy are FLOG and DOCK 4.0 [93].

5.3.1.2 Stochastic algorithms

At each phase of the docking process, this category randomly adapts all ligand degrees of freedom (translational, rotational, and conformational), resulting in a diverse set of solutions. Subsequently, probabilistic criteria are applied to sort the acceptable and rejected ligand poses. Among the most frequent stochastic algorithms, we mention: Monte Carlo (MC), Evolutionary Algorithms (EAs), Tabu Search (TS) and Swarm Optimisation (SO) [93,97].

- **In MC-based methods**, all ligand degrees of freedom are subjected to random modifications at the binding site, accompanied by an energy minimization for each generated pose. Afterwise, the Boltzmann constant was used as a criterion for pose acceptance or rejection. The Boltzmann constant takes into account the difference in energy before and after the random modifications, as well as the absolute temperature, as factors. The key strength of this method, termed Simulated Annealing, is that it incorporates temperature changes to enhance the probability of attaining the global minimum since heating may allow certain energy barriers to be crossed. DockVision 1.0.3, FDS, GlamDock, ICM, MCDOCK, PRODOCK, QXP, ROSETTALIGAND, RiboDock, Yucca, AutoDock, ICM, QXP, etc are examples of programs based on MC algorithms [83,86,93,96].
- **In EAs-based methods**, the theory of the evolution of biological samples through natural selection is applied to find optimal solutions to specific problems. This is accomplished by selecting the best individuals and then passing them on to the next generation, in addition, by performing random or biased mutations to boost genetic variety and avoid early convergence. These methods fall into three main classes: genetic algorithms, evolutionary programming, and evolution strategies. The main commonalities between these categories are the ability to cross local minima simply and the ability to offer a variety of low-energy solutions [93,94].
- **In TS-based method**, the similarity of newly formed non-lowest energy pose is investigated using the "Tabu-List". The latter function as memory, storing previously discovered low-energy solutions and encouraging the search for new solutions. TS research is an iterative procedure focused on solving optimization challenges. Typically, the pose is only kept if it differs from all of the previously mentioned solutions [93,96].
- **In SO-based method**, the updates made in a current solution are intended to follow the best pose of the population. The SO approach is also called swarm intelligence

(SI), and its basic concept is inspired by the collective behavior of animals such as insects, herds, birds, and fishes, allowing it to explore a vast part in the solution space of the optimized objective function at the same time. This method includes: Genetic Algorithms (GA), Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), Differential Evolution (DE), Artificial Bee Colony (ABC), Glowworm Swarm Optimization (GSO), and Cuckoo Search Algorithm (CSA) [98,99].

5.3.1.3 Deterministic algorithms

In each iteration, deterministic methods use the actual state of the ligand as the basis for generating the next state, such that the new state exhibit an energy value equal to or less than the base state. Thus, the final output is extremely dependent on the initial input structure, since when the starting system settings and parameters are standardized, the final state will always be the same. Examples of this type of algorithm include energy minimization methods and molecular dynamics (MD) simulations [93].

- **In EM method**, the direction associated with the potential energy gradient is applied to explore the energy landscape, guiding the system to approach the local minimum. Often, EM is used in integration with other docking methods, such as with the fragment-based method, to obtain the energy minimization of docking solutions [93].
- **In MD method**, the motions of a system over time are simulated as a function of thermodynamic factors such as temperature and pressure. MD strategies, like EM methods, are performed in integration with other docking strategies, such as with the simulated annealing, to enhance the prediction of CDOCKER algorithm [93].

5.4 Scoring Functions (SFs)

SFs serve as pose selectors, discriminating between the most effective biological binding modes and binders from inactive ligands in the set of poses obtained by the sampling algorithm (Figure II.10). SFs, on the other hand, employ a variety of approximations and simplifications to estimate, rather than calculate, the binding affinity of the target and the ligand [90,100]. Accordingly, a reliable SF should have three essential skills:

- The competence to select the optimal binding mode of a ligand from a set of computationally simulated poses.
- The competence to accurately rank a given set of ligands with known binding modes when attached to the same target.
- The competence to generate binding scores linearly linked to experimentally-measured binding affinities of target-ligand complexes with known 3D structures[101].

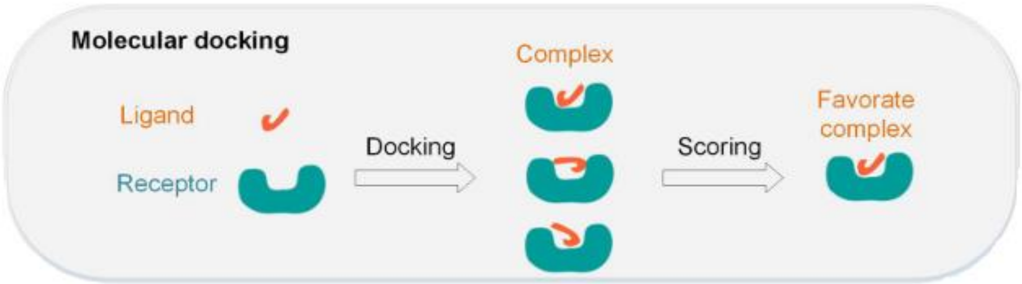


Figure II.10. The SF role as pose selector [102].

Recently, five classes have been developed in the field of SFs: Physics-based, empirical, knowledge-based, Consensus, and Machine-learning-based SFs [102].

5.4.1 Physics-based SFs

Physics-based SFs including the SFs based on force field, solvation models, and quantum mechanics methods (Figure II.11) [102]. Force-field is a classic concept of molecular mechanics that uses a combination of the bonded (intramolecular) and non-bonded (intermolecular) constituents of a system to approximate its potential energy. The docking approach often considers non-bonded elements with the possibility of including ligand-bonded terms, notably the torsional elements [90].

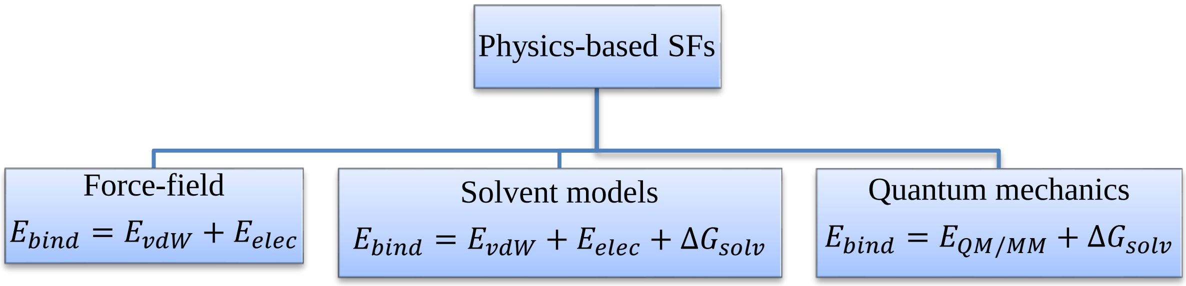


Figure II.11. The description for physics-based SF [102].

The entire sum of the non-bonded interactions indicated by electrostatic and van der Waals energy terms is frequently used to estimate the binding energy in classical force-field-based SFs. Extensions such as hydrogen bonds, solvations, and entropy contributions also can be considered in the force-field-based SFs. The coulomb formula is used to calculate the electrostatic forces. Because such point charge computations have trouble describing the real surroundings of target, a distance-dependent dielectric function is typically applied to tune the effect of charge-charge interactions. Besides, the Lennard-Jones potential function defines the van der Waals terms. The hardness of the potential, which controls the permitted cut-off distances between target and ligand atoms, may be adjusted by using different parameter settings for the Lennard-Jones potential [86].

Software programs differ in their handling of hydrogen bonding, the form of the energy function, and other aspects. Furthermore, to increase the accuracy of the prediction of binding energies, the findings of docking utilizing force-field-based functions may be adjusted by other approaches such as linear interaction energy [103] and free-energy perturbation methods (FEP) [104].

The following are some examples of SFs based on force fields: D-Score, G-Score, GOLD, AutoDock, DOCK, HADDOCK Score, ICM SF, QXP SF, GBVI/WSA, etc [83,90,105].

5.4.2 Empirical SFs

When compared to force-field-based SFs, empirical SFs take into account simpler energy factors, resulting in substantially quicker binding score calculations and reasonable predicted binding energy. To estimate the binding energy of a system, they incorporate many energy components such as van der Waals, electrostatic, hydrogen bond, ionic interaction, desolvation, hydrophobic effect, binding entropy, etc. To achieve the final score, each energy term is multiplied by a coefficient and then added together. These coefficients are derived by regression analysis on an experimental training set fitted with a test set of ligand-target complexes with known binding affinities [83,100]. Accordingly, these SFs have the drawback of being reliant on the molecular data sets utilized to perform regression analysis and fitting. This frequently results in different weighting factors for the individual terms. As a result, the difficulties of merging terms from separate SFs into a new SF have been recorded [105].

The following are some examples of SFs that belong to this category: LUDI, GlideScore, ChemScore, PlantsChemplp, SCORE, RankScore, LigScore, HINT, F-Score, Fresno, X-Score [83,86,90,105].

5.4.3 Knowledge-based SFs

Knowledge-based SFs are intended to replicate experimental structures instead of binding energies. In essence, they are founded on the hypothesis that the more favorable the interaction, the more it will occur. To attain this purpose, they apply statistical analysis at crystal 3D-structures of ligand-target complexes dataset to determine the frequencies of interatomic interaction, and, or distances between the ligand and target. These frequency distributions are then used to derive paired atom-type potentials. After that, the score is computed by prioritizing favorable connections and eliminating repulsive interactions between ligand-target atoms within a pre-defined cutoff [83,105] . Figure II.12 illustrates the general pathway of knowledge-based SFs.

Knowledge-based functions allow for easy scanning of large molecular databases. They can also model certain unusual interactions that are frequently overlooked in empirical SFs, such as sulphur-aromatic or π-cation. Another advantage is that the training sets focus only on structural details and are independent of experimental binding affinity, which removes any uncertainty in the binding affinity induced by the experimental environment. Therefore, knowledge-based SFs are better suited for predicting binding poses than binding affinities [102,105]. Nevertheless, the key challenges are the accuracy of estimating the reference state and the under-representation of interactions with halogens and metals in the restricted training sets [100].

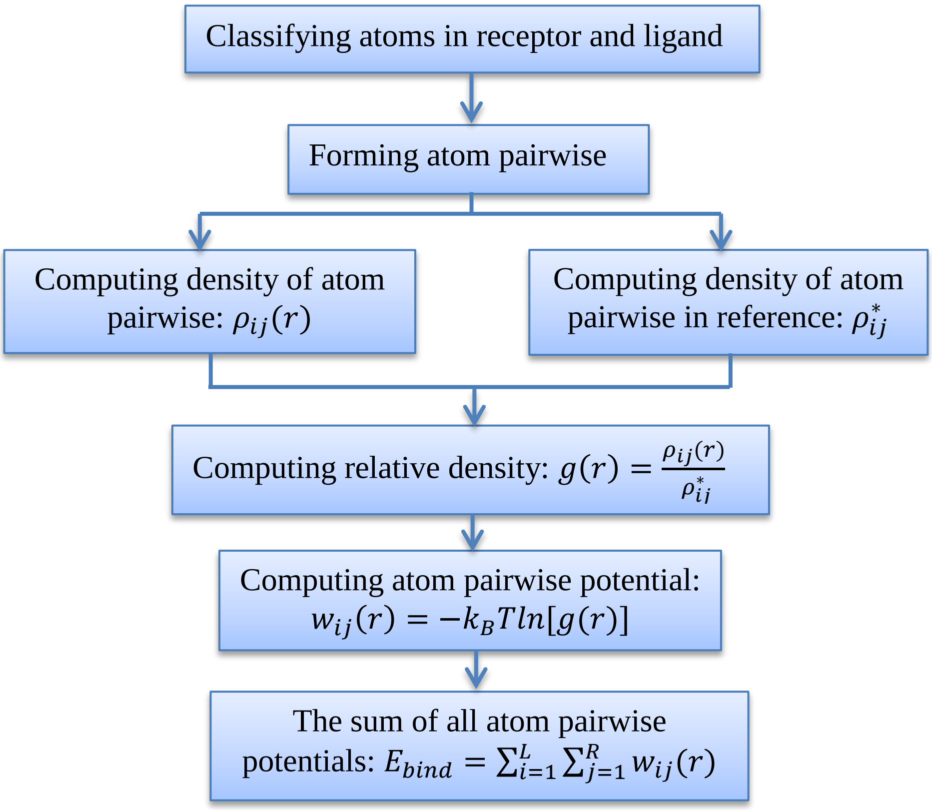


Figure II.12. The general pathway of knowledge-based SFs. $\boldsymbol{\rho_{ij}(r)}$: number density of the target-ligand atom pair i_j at distance r . $\boldsymbol{\rho_{ij}^{*}}$: pair density in a reference state. g(r): relative number density of atom pairwise i_j at distance r. $\boldsymbol{k_B}$: Boltzmann constant. T: absolute temperature [102].

Here are some examples of SFs that fall within this category: DrugScore, GOLD/ASP, PMF, SMoG, Bleep, MScore, ITScore/SE [86,90,105].

5.4.4 Consensus scoring

To minimize the inherent constraints of the above-mentioned classes of SFs, the combination of more than one SF is a recent trend in this field. "Consensus scoring" is the term given to this combinatorial approach [106]. The consensus scoring concept is based on integrating information from multiple scores to balance errors in single score and enhance the chance of detecting the exact ligand conformation. Generally, a suggested ligand pose may be approved if it scores well in a variety of scoring schemes [107].

X-CSCORE, which combines GOLD-like, DOCK-like, ChemScore, PMF, and FlexX SFs, is an example of consensus scoring implementation [86,105].

5.4.5 Machine-learning-based SFs

Machine-learning-based SFs were carried out using several machine-learning algorithms, including support vector machine, random forest, neural network, deep-learning, and others (Figure II.13). Because machine-learning-based SFs are dependent on the information in the training dataset, they are often used for rescoring and rarely integrated directly into docking software, despite outperforming classical SFs. When the ligand and target are docked using classical docking techniques, and then the anchored structure is rescored using machine learning SFs, the accuracy improves [102].

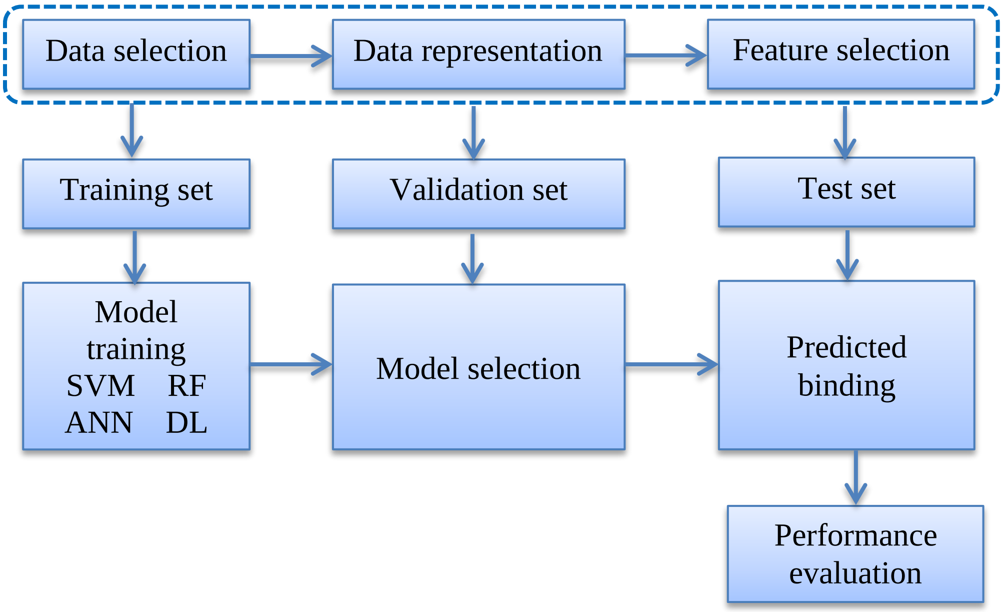


Figure II.13. Workflow of training machine-learning-based SFs [102].

6 Molecular dynamics (MDs)

MD is a computational tool that follows the dynamic behavior of molecular systems at different time scales. It explicitly treats structural flexibility and entropic effects of all objects under simulation (ligand, protein, waters, etc). To help in the selection of a successful candidate, MDs simulations are used in drug development to explore the extrinsic surface and bulk properties of different forms of pharmacologically active compounds. It also featured a precise assessment of numerous ligand characteristics in different solvents, such as the binding modes, kinetics, stabilities, densities, conductivities, dipole moments, and thermodynamic factors including interactions energies and entropies [108].

MD is useful not only for investigating the molecular properties measured experimentally but also in refining most conformations identified by X-ray or NMR methods. Thus, in the field of MD simulations, the collaboration between computational and experimental approaches has a long history. Theoretical methods are contributing to the interpretation and analysis of experimental data, whereas, experimental methods are critical for validating and developing computational approaches and processes [91].

6.1 MD simulation before/during the docking process

Before performing molecular docking, MD simulations can be generated to accomplish a variety of goals, including:

- ✓ Optimizing the target structure while taking flexibility into consideration.
- ✓ Refine the docked complexes.
- ✓ Incorporate solvent influences and account for induced fit.
- ✓ Determine the binding free energies.
- ✓ Offer a precise classification of the potential ligands.

Also, in recent improvements, during the docking process it-self to accurately select the binding site and appropriately attach the ligand [91].

7 Combined molecular docking/ MD simulations

The strategy of combining molecular docking and MDs simulation in a single protocol is a logical approach to improve the drug-design process. The cost of computations can be balanced by integrating fast and low-cost docking simulations with more accurate and expensive MD operations to predict the most trustworthy target/ligand complexes. The power of this combination derives from their complementing skills and deficiencies [91].

Docking techniques can rapidly examine the extensive conformational space of ligands, allowing large libraries of compounds to be analyzed at a relatively low cost. However, it suffers from poor receptor flexibility that hinders its conformation adjustment during complexation. As well as, the lack of a unique and broadly applicable SF, this is essential for establishing a credible classification of the resulting complexes. Besides, the main drawbacks of applying MD simulation are that it is large time-consuming and the system can get trapped in local minima. However, the opportunity it offers to approach both ligands and receptors as flexible entities, allowing for an induced-fit between the anchored ligand and its target-binding site, cannot be overlooked. Also, through MD simulation, the effect of explicit water molecules can be studied directly, and very accurate binding free energies can be obtained [90,91].

Bibliography

[1] P.K. Chattaraj, Chemical reactivity theory: a density functional view, 2009. https://doi.org/10.1201/9781420065442.

[2] P. Geerlings, F. De Proft, W. Langenaeker, Conceptual density functional theory, Chem. Rev. 103 (2003) 1793–1873. https://doi.org/10.1021/cr990029p.

[3] S. LaPointe, D. Weaver, A Review of Density Functional Theory Quantum Mechanics as Applied to Pharmaceutically Relevant Systems, Curr. Comput. Aided-Drug Des. 3 (2007) 290–296. https://doi.org/10.2174/157340907782799390.

[4] H. Tandon, T. Chakraborty, V. Suhag, A Concise Review on the Significance of QSAR in Drug Design, Chem. Biomol. Eng. 4 (2019) 45–51. https://doi.org/10.11648/j.cbe.20190404.11.

[5] R.O. Jones, Density functional theory: Its origins, rise to prominence, and future, Rev. Mod. Phys. 87 (2015) 897–923. https://doi.org/10.1103/RevModPhys.87.897.

[6] G.-F. Yang, X. Huang, Development of Quantitative Structure-Activity Relationships and Its Application in Rational Drug Design, Curr. Pharm. Des. 12 (2006) 4601–4611. https://doi.org/10.2174/138161206779010431.

[7] P. Geerlings, E. Chamorro, P.K. Chattaraj, F. De Proft, J.L. Gázquez, S. Liu, C. Morell, A. Toro-Labbé, A. Vela, P. Ayers, Conceptual density functional theory: status, prospects, issues, Theor. Chem. Acc. 139 (2020) 1–18. https://doi.org/10.1007/s00214-020-2546-7.

[8] S. Kenouche, A. Belkadi, R. Djebaili, N. Melkemi, High regioselectivity in the amination reaction of isoquinolinequinone derivatives using conceptual DFT and NCI analysis, J. Mol. Graph. Model. 104 (2021) 107828. https://doi.org/10.1016/j.jmgm.2020.107828.

[9] P.K. Chattaraj, ed., Chemical Reactivity Theory, A Density Functional View, 1st Editio, 2009. https://doi.org/https://doi.org/10.1201/9781420065442.

[10] W.Y. R.G. Parr, Density-Functional Theory of Atoms and Molecules, Oxford University press, New York, 1989.

[11] R.T. Sanderson, Partial charges on atoms in organic compounds, Science (80-.). 121 (1955) 207–208. https://doi.org/https://doi.org/10.1126/science.121.3137.207.

[12] R.P. Iczkowski, J.L. Margrave, Electronegativity, J. Am. Chem. Soc. 83 (1961) 3547–3551. https://doi.org/10.1021/ja01478a001.

[13] R.G. Parr, R.G. Pearson, Absolute Hardness: Companion Parameter to Absolute Electronegativity, J. Am. Chem. Soc. 105 (1983) 7512–7516. https://doi.org/10.1021/ja00364a005.

[14] N. Sukumar, A Matter of Density: Exploring the Electron Density Concept in the Chemical, Biological, and Materials Sciences, John Wiley & Sons, 2012.

https://doi.org/10.1002/9781118431740.

[15] M. V. Putz, D.M.P. Mingos, eds., applications of Density Functional Theory to Chemical Reactivity, 1st ed., Springer, Berlin, Heidelberg, 2013. https://doi.org/https://doi.org/10.1007/978-3-642-32753-7.

[16] R.G. Parr, L. V. Szentpály, S. Liu, Electrophilicity index, J. Am. Chem. Soc. 121 (1999) 1922–1924. https://doi.org/10.1021/ja983494x.

[17] E. Chamorro, P.K. Chattaraj, P. Fuentealba, Variation of the Electrophilicity Index along the Reaction Path, J. Phys. Chem. A. 107 (2003) 7068–7072. https://doi.org/10.1021/jp035435y.

[18] P. Jaramillo, L.R. Domingo, E. Chamorro, P. Pérez, A further exploration of a nucleophilicity index based on the gas-phase ionization potentials, J. Mol. Struct. THEOCHEM. 865 (2008) 68–72. https://doi.org/10.1016/j.theochem.2008.06.022.

[19] L.R. Domingo, P. Pérez, Global and local reactivity indices for electrophilic/nucleophilic free radicals, Org. Biomol. Chem. 11 (2013) 4350–4358. https://doi.org/10.1039/c3ob40337h.

[20] L.R. Domingo, M.J. Aurell, P. Pérez, R. Contreras, Quantitative characterization of the global electrophilicity power of common diene/dienophile pairs in Diels-Alder reactions, Tetrahedron. 58 (2002) 4417–4423. https://doi.org/10.1016/S0040-4020(02)00410-6.

[21] J. Leszczynski, ed., Handbook of Computational Chemistry, 1st ed., Springer, Dordrecht, 2012. https://doi.org/https://doi.org/10.1007/978-94-007-0711-5.

[22] C. Morell, A. Grand, A. Toro-Labbé, Theoretical support for using the Δf(r) descriptor, Chem. Phys. Lett. 425 (2006) 342–346. https://doi.org/10.1016/j.cplett.2006.05.003.

[23] C. Cárdenas, N. Rabi, P.W. Ayers, C. Morell, P. Jaramillo, P. Fuentealba, Chemical reactivity descriptors for ambiphilic reagents: Dual descriptor, local hypersoftness, and electrostatic potential, J. Phys. Chem. A. 113 (2009) 8660–8667. https://doi.org/10.1021/jp902792n.

[24] W. Yang, W.J. Mortier, The Use of Global and Local Molecular Parameters for the Analysis of the Gas-Phase Basicity of Amines, J. Am. Chem. Soc. 108 (1986) 5708–5711. https://doi.org/10.1021/ja00279a008.

[25] L.H. Mendoza-Huizar, G. Salgado-Morán, R. Ramirez-Tagle, D. Glossman-Mitnik, A theoretical quantum study of the intramolecular interactions and chemical reactivity of polymorphs A and B of famotidine in the gas, DMSO, and aqueous phases, Comput. Theor. Chem. 1075 (2016) 54–62. https://doi.org/10.1016/j.comptc.2015.11.007.

[26] R. Flores-Moreno, Symmetry conservation in fukui functions, J. Chem. Theory Comput. 6 (2010) 48–54. https://doi.org/10.1021/ct9002527.

[27] C. Morell, A. Grand, A. Toro-Labbé, New dual descriptor for chemical reactivity, J. Phys. Chem. A. 109 (2005) 205–212. https://doi.org/10.1021/jp046577a.

[28] R. Pino-Rios, D. Inostroza, G. Cárdenas-Jirón, W. Tiznado, Orbital-Weighted Dual Descriptor for the Study of Local Reactivity of Systems with (Quasi-) Degenerate States, J. Phys. Chem. A. 123 (2019) 10556–10562. https://doi.org/10.1021/acs.jpca.9b07516.

[29] Jorge Ignacio Martínez-Araya, An intermediate level of approximation for computing the dual descriptor, J. Mol. Model. 19 (2013) 2811–2820. https://doi.org/10.1007/s00894-012-1599-5.

[30] J. Martínez, Local reactivity descriptors from degenerate frontier molecular orbitals, Chem. Phys. Lett. 478 (2009) 310–322. https://doi.org/10.1016/j.cplett.2009.07.086.

[31] J.I. Martínez-Araya, A generalized operational formula based on total electronic densities to obtain 3D pictures of the dual descriptor to reveal nucleophilic and electrophilic sites accurately on closed-shell molecules, J. Comput. Chem. 37 (2016) 2279–2303. https://doi.org/10.1002/jcc.24453.

[32] J.I. Martínez, J.L. Moncada, J.M. Larenas, The dual descriptor to measure local reactivity on Buckminster fullerenes: An analysis within the framework of conceptual DFT, J. Mol. Model. 16 (2010) 1825–1832. https://doi.org/10.1007/s00894-009-0638-3.

[33] C. Crdenas, P.W. Ayers, A. Cedillo, Reactivity indicators for degenerate states in the density-functional theoretic chemical reactivity theory, J. Chem. Phys. 134 (2011). https://doi.org/10.1063/1.3585610.

[34] J.S. Murray, T. Brinck, P. Lane, K. Paulsen, P. Politzer, Statistically-based interaction indices derived from molecular surface electrostatic potentials: a general interaction properties function (GIPF), J. Mol. Struct. THEOCHEM. 307 (1994) 55–64. https://doi.org/10.1016/0166-1280(94)80117-7.

[35] J.S. Murray, P. Politzer, The electrostatic potential: An overview, Wiley Interdiscip. Rev. Comput. Mol. Sci. 1 (2011) 153–163. https://doi.org/10.1002/wcms.19.

[36] T. Lu, F. Chen, Quantitative analysis of molecular surface based on improved Marching Tetrahedra algorithm, J. Mol. Graph. Model. 38 (2012) 314–323. https://doi.org/10.1016/j.jmgm.2012.07.004.

[37] C. Cárdenas, N. Rabi, P.W. Ayers, C. Morell, P. Jaramillo, P. Fuentealba, Chemical reactivity descriptors for ambiphilic reagents: Dual descriptor, local hypersoftness, and electrostatic potential, J. Phys. Chem. A. 113 (2009) 8660–8667. https://doi.org/10.1021/jp902792n.

[38] R.F.W. Bader, M.T. Carroll, J.R. Cheeseman, C. Chang, Properties of Atoms in Molecules: Atomic Volumes, J. Am. Chem. Soc. 109 (1987) 7968–7979. https://doi.org/10.1021/ja00260a006.

[39] P. Politzer, J.S. Murray, The fundamental nature and role of the electrostatic potential in atoms and molecules, Theor. Chem. Acc. 108 (2002) 134–142. https://doi.org/10.1007/s00214-002-0363-9.

[40] C. Hansch, T. Fujita, ρ-σ-π Analysis. A Method for the Correlation of Biological Activity and Chemical Structure, J. Am. Chem. Soc. 86 (1964) 1616–1626. https://doi.org/10.1021/ja01062a035.

[41] J. Verma, V. Khedkar, E. Coutinho, 3D-QSAR in Drug Design - A Review, Curr. Top. Med. Chem. 10 (2010) 95–115. https://doi.org/10.2174/156802610790232260.

[42] K. Roy, On some aspects of validation of predictive quantitative structure-activity relationship models, Expert Opin. Drug Discov. 2 (2007) 1567–1577. https://doi.org/10.1517/17460441.2.12.1567.

[43] B.J. Neves, R.C. Braga, C.C. Melo-Filho, J.T. Moreira-Filho, E.N. Muratov, C.H. Andrade, QSAR-based virtual screening: Advances and applications in drug discovery, Front. Pharmacol. 9 (2018) 1–7. https://doi.org/10.3389/fphar.2018.01275.

[44] K.Z. Myint, X.Q. Xie, Recent advances in fragment-based QSAR and multi-dimensional QSAR methods, Int. J. Mol. Sci. 11 (2010) 3846–3866. https://doi.org/10.3390/ijms11103846.

[45] R.D. Cramer, D.E. Patterson, J.D. Bunce, Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins, J. Am. Chem. Soc. 110 (1988) 5959–5967. https://doi.org/10.1021/ja00226a005.

[46] A. Golbraikh, X.S. Wang, H. Zhu, A. Tropsha, Predictive QSAR Modeling: Methods and Applications in Drug Discovery and Chemical Risk Assessment, in: Leszczynski J. (Ed.), Handb. Comput. Chem., Springer, Dordrecht, 2016: pp. 1–48. https://doi.org/10.1007/978-94-007-6169-8_37-3.

[47] U. Muhammad, A. Uzairu, D. Ebuka Arthur, Review on: quantitative structure activity relationship (QSAR) modeling, J. Anal. Pharm. Res. 7 (2018) 1–9. https://doi.org/10.15406/japlr.2018.07.00232.

[48] R. Perkins, H. Fang, W. Tong, W.J. Welsh, Quantitative structure-activity relationship methods: Perspectives on drug discovery and toxicology, Environ. Toxicol. Chem. 22 (2003) 1666–1679. https://doi.org/10.1897/01-171.

[49] D. Fourches, E. Muratov, A. Tropsha, Trust, but Verify II: A Practical Guide to Chemogenomics Data Curation, J. Chem. Inf. Model. 56 (2016) 1243–1252. https://doi.org/10.1021/acs.jcim.6b00129.

[50] D. Fourches, E. Muratov, A. Tropsha, Trust, but verify: On the importance of chemical structure curation in cheminformatics and QSAR modeling research, J. Chem. Inf. Model. 50 (2010) 1189–1204. https://doi.org/10.1021/ci100176x.

[51] Danishuddin, A.U. Khan, Descriptors and their selection methods in QSAR analysis: paradigm for drug design, Drug Discov. Today. 21 (2016) 1291–1302. https://doi.org/10.1016/j.drudis.2016.06.013.

[52] N.S. Sethi, A Review on Computational Methods in Developing Quantitative Structure-Activity Relationship (QSAR) ABSTRACT : Introducton, (2012).

[53] J.O. Rawlings, S.G. Pantula, D.A. Dickey, Applied Regression Analysis : A Research

Tool, Second Edi, NY: Springer New York., New York, 1998.

[54] S. De Jong, SIMPLS: an alternative approach to partial least squares regression, Chemom. Intell. Lab. Syst. 18 (1993) 251–263. https://doi.org/10.1016/0169-7439(93)85002-X.

[55] J.O. Rawlings, S.G. Pantula, D.A. Dickey, Applied Regression Analysis : A Research Tool, second edi, NY : Springer New York, New York, 1998. https://doi.org/https://doi.org/10.1007/b98890.

[56] G.M. Furnival, R.W. Wilson, Regressions by leaps and bounds, Technometrics. 42 (2000) 69–79. https://doi.org/10.1080/00401706.2000.10485982.

[57] G. James, D. Witten, T. Hastie, Robert Tibshirani, An Introduction to Statistical Learning - with Applications in R, springer, New York, 2013.

[58] P. Ruengvirayudh, G.P. Brooks, Comparing Stepwise Regression Models to the Best-Subsets Models, or, the Art of Stepwise, Gen. Linear Model J. 42 (2016) 1–14.

[59] K. Wang, Z. Chen, Stepwise Regression and All Possible Subsets Regression in Education, Electron. Int. J. Educ. Arts, Sci. 2 (2016) 60–81.

[60] M.Z.I. Chowdhury, T.C. Turin, Variable selection strategies and its importance in clinical prediction modelling, Fam. Med. Community Heal. 8 (2020) 1–7. https://doi.org/10.1136/fmch-2019-000262.

[61] R.B. BENDEL, A.A. AFIFI, Comparison of Stopping Rules in Forward “Stepwise” Regression, J. Am. Stat. Assoc. 72 (1977) 46–53. https://doi.org/DOI: 10.1080/01621459.1977.10479905.

[62] G. R. Pasha, SELECTION OF VARIABLES IN MULTIPLE REGRESSION USING STEPWISE REGRESSION, J. Res. 13 (2002) 119–127.

[63] A.T. McCray, J. McNames, D. Abercrombie, Locating disturbances in semiconductor manufacturing with stepwise regression, IEEE Trans. Semicond. Manuf. 18 (2005) 458–468. https://doi.org/10.1109/TSM.2005.852118.

[64] D.-S. CAO, Y.-Z. LIANG, Q.-S. XU, H.-D. LI, X. CHEN, A New Strategy of Outlier Detection for QSAR/QSPR, J. Comput. Chem. 31 (2010) 529–602. https://doi.org/doi: 10.1002/jcc.21351.

[65] R.P. Verma, C. Hansch, An approach toward the problem of outliers in QSAR, Bioorganic Med. Chem. 13 (2005) 4597–4621. https://doi.org/10.1016/j.bmc.2005.05.002.

[66] K.H. Kim, Outliers in SAR and QSAR: 3. Importance of considering the role of water molecules in protein–ligand interactions and quantitative structure–activity relationship studies, J. Comput. Aided. Mol. Des. 35 (2021) 371–396. https://doi.org/10.1007/s10822-021-00377-7.

[67] J.C. Dearden, M.T.D. Cronin, K.L.E. Kaiser, How not to develop a quantitative structure-activity or structure-property relationship (QSAR/QSPR), SAR QSAR

Environ. Res. 20 (2009) 241–266. https://doi.org/10.1080/10629360902949567.

[68] G. Coenders, M. Saez, Collinearity, Heteroscedasticity and Outlier Diagnostics in Regression. Do They Always Offer What They Claim?, in: Anuška Ferligoj and Andrej Mrvar (Ed.), New Approaches Appl. Stat., 2000: pp. 79–94.

[69] A. Cherkasov, E.N. Muratov, D. Fourches, A. Varnek, I.I. Baskin, M. Cronin, J. Dearden, P. Gramatica, Y.C. Martin, R. Todeschini, V. Consonni, V.E. Kuz'Min, R. Cramer, R. Benigni, C. Yang, J. Rathman, L. Terfloth, J. Gasteiger, A. Richard, A. Tropsha, QSAR modeling: Where have you been? Where are you going to?, J. Med. Chem. 57 (2014) 4977–5010. https://doi.org/10.1021/jm4004285.

[70] A. Golbraikh, A. Tropsha, Beware of q2!, in: J. Mol. Graph. Model., 2002: pp. 269–276. https://doi.org/10.1016/S1093-3263(01)00123-1.

[71] R. Veerasamy, H. Rajak, A. Jain, S. Sivadasan, C.P. Varghese, R.K. Agrawal, Validation of QSAR Models - Strategies and Importance, Int. J. Drug Des. Disocovery. 2 (2011) 511–519.

[72] V. Consonni, R. Todeschini, D. Ballabio, F. Grisoni, On the Misleading Use of QF32 for QSAR Model Comparison, Mol. Inform. 38 (2019) 2–6. https://doi.org/10.1002/minf.201800029.

[73] O. Renaud, M.P. Victoria-Feser, A robust coefficient of determination for regression, J. Stat. Plan. Inference. 140 (2010) 1852–1862. https://doi.org/10.1016/j.jspi.2010.01.008.

[74] O. Harel, The estimation of R2 and adjusted R2 in incomplete data sets using multiple imputation, J. Appl. Stat. 36 (2009) 1109–1118. https://doi.org/10.1080/02664760802553000.

[75] M. Abdullahi, S.E. Adeniji, D.E. Arthur, S. Musa, Quantitative structure-activity relationship (QSAR) modelling study of some novel carboxamide series as new anti-tubercular agents, Bull. Natl. Res. Cent. 44 (2020) 1–13. https://doi.org/10.1186/s42269-020-00389-7.

[76] Y. Traoré, M.G.-R. Koné, O. Ouattara, N. Ziao, QSAR APPROACH TO ESTIMATING THE ANALGESIC ACTIVITY OF A SERIES OF TRI-SUBSTITUTED PYRIMIDINE DERIVATIVES, SDRP J. Comput. Chem. Mol. Model. 2 (2018) 221–234.

[77] V. Consonni, D. Ballabio, R. Todeschini, Comments on the definition of the Q2 parameter for QSAR validation, J. Chem. Inf. Model. 49 (2009) 1669–1678. https://doi.org/10.1021/ci900115y.

[78] V. Consonni, D. Ballabio, R. Todeschini, Evaluation of model predictive ability by external validation techniques, J. Chemom. 24 (2010) 194–201. https://doi.org/10.1002/cem.1290.

[79] G. Schüürmann, R.U. Ebert, J. Chen, B. Wang, R. Kühne, External Validation and Prediction Employing the Predictive Squared Correlation Coefficient-Test Set Activity

Mean vs Training Set Activity Mean, J. Chem. Inf. Model. 48 (2008) 2140–2145. https://doi.org/10.1021/ci800253u.

[80] L.M. Shi, H. Fang, W. Tong, J. Wu, R. Perkins, R.M. Blair, W.S. Branham, S.L. Dial, C.L. Moland, D.M. Sheehan, QSAR Models Using a Large Diverse Set of Estrogens, J. Chem. Inf. Comput. Sci. 41 (2001) 186–195. https://doi.org/10.1021/ci000066d.

[81] D.M. Hawkins, The Problem of Overfitting, J. Chem. Inf. Comput. Sci. 44 (2004) 1–12. https://doi.org/10.1021/ci0342472.

[82] M. Mukhopadhyay, a Brief Survey on Bio Inspired Optimization Algorithms for Molecular Docking, Int. J. Adv. Eng. Technol. 7 (2014) 868–878. https://doi.org/10.7323/ijaet/v7_iss3.

[83] A. Sethi, K. Joshi, K. Sasikala, M. Alvala, Molecular Docking in Modern Drug Discovery: Principles and Recent Applications, in: Drug Discov. Dev. - New Adv., 2020: pp. 1–21. https://doi.org/10.5772/intechopen.85991.

[84] D.J. Diller, K.M. Merz, High throughput docking for library design and library prioritization, Proteins Struct. Funct. Genet. 43 (2001) 113–124. https://doi.org/10.1002/1097-0134(20010501)43:2<113::AID-PROT1023>3.0.CO;2-T.

[85] A. Stefaniu, Introductory Chapter: Molecular Docking and Molecular Dynamics Techniques to Achieve Rational Drug Design, in: Mol. Docking Mol. Dyn., 2019: pp. 1–5. https://doi.org/10.5772/intechopen.84200.

[86] X.-Y. Meng, H.-X. Zhang, M. Mezei, M. Cui, Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery., Curr. Comput. Aid. Dru. Des. 7 (2011) 146–157. https://doi.org/10.2174/157340911795677602.

[87] P.H.M. Torres, A.C.R. Sodero, P. Jofily, F.P. Silva-Jr, Key topics in molecular docking for drug design, Int. J. Mol. Sci. 20 (2019) 1–29. https://doi.org/10.3390/ijms20184574.

[88] S.P. Leelananda, S. Lindert, Computational methods in drug discovery, Beilstein J. Org. Chem. 12 (2016) 2694–2718. https://doi.org/10.3762/bjoc.12.267.

[89] R.D. Taylor, P.J.Jewsbury, J.W. Essex1, A review of protein-small molecule docking methods, J. OfComputer-Aided Mol. Des. 16 (2002) 151–166. https://doi.org/doi: 10.1023/A:1020155510718.

[90] V. Salmaso, S. Moro, Bridging molecular docking to molecular dynamics in exploring ligand-protein recognition process: An overview, Front. Pharmacol. 9 (2018) 1–16. https://doi.org/10.3389/fphar.2018.00923.

[91] H. Alonso, A.A. Bliznyuk, J.E. Gready, Combining docking and molecular dynamic simulations in drug design, Med. Res. Rev. 26 (2006) 531–568. https://doi.org/10.1002/med.20067.

[92] F. Jiang, S.H. Kim, “Soft docking”: Matching of molecular surface cubes, J. Mol. Biol. 219 (1991) 79–102. https://doi.org/10.1016/0022-2836(91)90859-5.

[93] I.A. Guedes, C.S. de Magalhães, L.E. Dardenne, Receptor–ligand molecular docking, Biophys. Rev. 6 (2014) 75–87. https://doi.org/DOI 10.1007/s12551-013-0130-2.

[94] N. Brooijmans, I.D. Kuntz, Molecular recognition and docking algorithms, Annu. Rev. Biophys. Biomol. Struct. 32 (2003) 335–373. https://doi.org/10.1146/annurev.biophys.32.110601.142532.

[95] N.S. Pagadala, K. Syed, J. Tuszynski, Software for molecular docking : a review, Biophys. Rev. (2017). https://doi.org/10.1007/s12551-016-0247-1.

[96] R. Dias, W.F. de A. Jr, Molecular Docking Algorithms, Curr. Drug Targets. 9 (2008) 1040–1047. https://doi.org/10.2174/138945008786949432.

[97] S.Y. Huang, S.Z. Grinter, X. Zou, Scoring functions and their evaluation methods for protein-ligand docking: Recent advances and future directions, Phys. Chem. Chem. Phys. 12 (2010) 12899–12908. https://doi.org/10.1039/c0cp00151a.

[98] M.N. Ab Wahab, S. Nefti-Meziani, A. Atyabi, A comprehensive review of swarm optimization algorithms, PLoS One. 10 (2015) 1–36. https://doi.org/10.1371/journal.pone.0122827.

[99] D. Wang, D. Tan, L. Liu, Particle swarm optimization algorithm: an overview, Soft Comput. 22 (2018) 387–408. https://doi.org/10.1007/s00500-016-2474-6.

[100] F.D. Prieto-Martínez, M. Arciniega, J.L. Medina-Franco, Molecular docking: current advances and challenges, TIP Rev. Espec. En Ciencias Químico-Biológicas. 21 (2018) 1–23. https://doi.org/10.22201/fesz.23958723e.2018.0.143.

[101] H.M. Ashtawy, N.R. Mahapatra, Machine-learning scoring functions for identifying native poses of ligands docked to known and novel proteins, BMC Bioinformatics. 16 (2015) 1–17. https://doi.org/10.1186/1471-2105-16-S6-S3.

[102] J. Li, A. Fu, L. Zhang, An Overview of Scoring Functions Used for Protein–Ligand Interactions in Molecular Docking, Interdiscip. Sci. Comput. Life Sci. 11 (2019) 320–328. https://doi.org/10.1007/s12539-019-00327-w.

[103] J. Michel, M.L. Verdonk, J.W. Essex, Protein-ligand binding affinity predictions by implicit solvent simulations: A tool for lead optimization?, J. Med. Chem. 49 (2006) 7427–7439. https://doi.org/10.1021/jm061021s.

[104] J.M. Briggs, T.J. Marrone, J.A. Mccammon, ComputationalScience New Horizons and Relevance to pharmaceutical Design, 6 (1996) 198–204.

[105] D.B. Kitchen, H. Decornez, J.R. Furr, J. Bajorath, Docking and scoring in virtual screening for drug discovery: Methods and applications, Nat. Rev. Drug Discov. 3 (2004) 935–949. https://doi.org/10.1038/nrd1549.

[106] P.S. Charifson, J.J. Corkery, M.A. Murcko, W.P. Walters, Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins, J. Med. Chem. 42 (1999) 5100–5109. https://doi.org/10.1021/jm990352k.

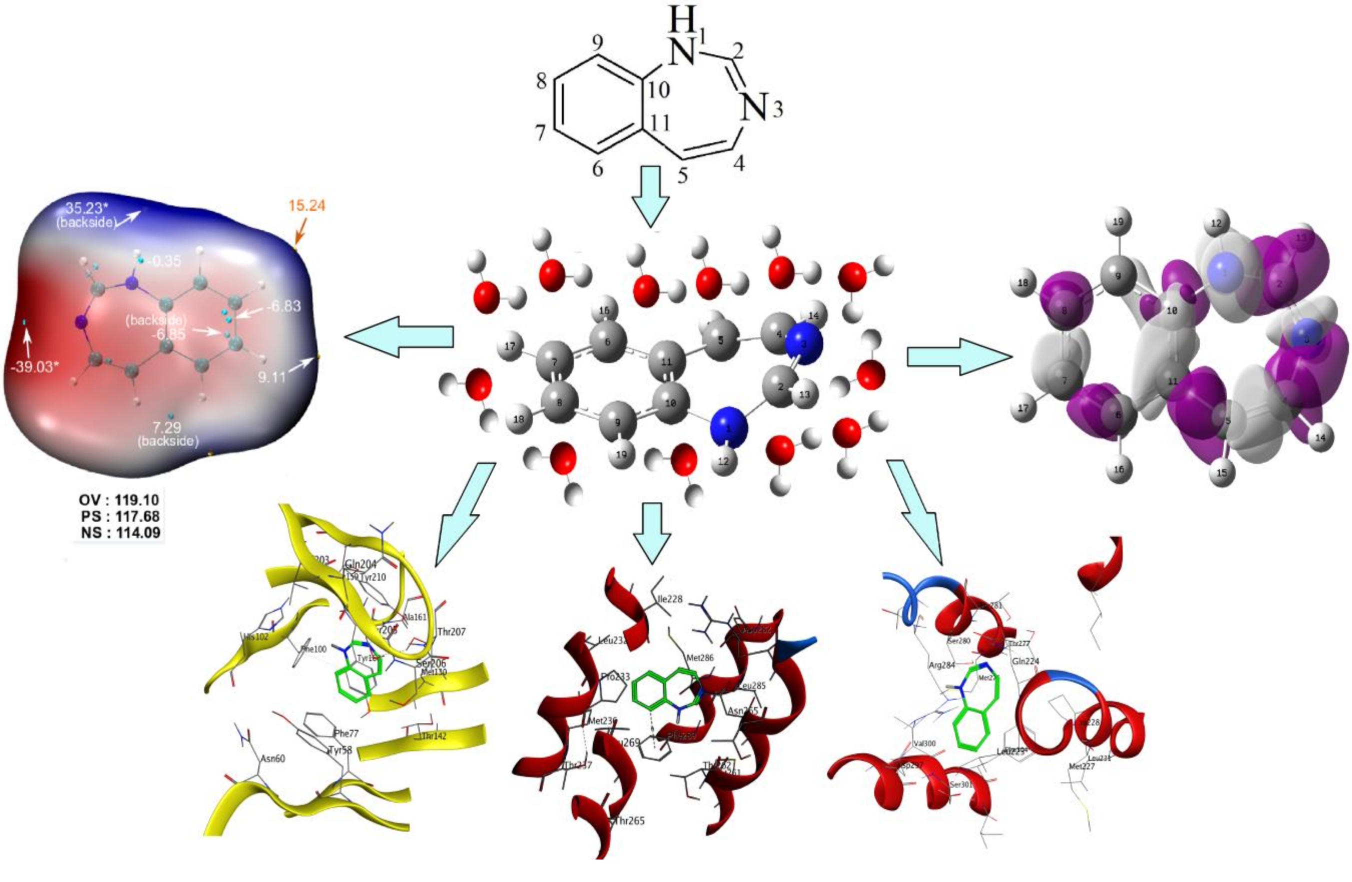
[107] R.D. Clark, A. Strizhev, J.M. Leonard, J.F. Blake, J.B. Matthew, Consensus scoring for ligand/protein interactions, J. Mol. Graph. Model. 20 (2002) 281–295. https://doi.org/10.1016/S1093-3263(01)00125-5.

[108] M. De Vivo, M. Masetti, G. Bottegoni, A. Cavalli, Role of Molecular Dynamics and Related Methods in Drug Discovery, J. Med. Chem. 59 (2016) 4035–4061. https://doi.org/10.1021/acs.jmedchem.5b01684.

Chapter III

Combined Conceptual-DFT, Quantitative MEP Analysis, and Molecular Docking Study of Benzodiazepine Analogs.

Graphical Abstract



1 Introduction

Due to the broad biological applications of the benzodiazepine class, knowledge of structural parameters, electronic proprieties, and chemical reactivity of their basic rings is certainly of great interest and can help in understanding the affinity of those drugs for the specific receptors, thus the contribution in the systematization of their main therapeutic activities by modifying the old molecules or generating new substances [1,2]. Literature has revealed that there are many previous studies [3–6] discussing the synthesis of different benzodiazepine derivatives with various biological activities based on modifications and substitutions of their six basic rings represented in Figure 13 of chapter I. However, to our knowledge, these chemical structures have not been the subject of a theoretical study.

Conceptual DFT-based descriptors and quantitative MEP analysis have emerged as powerful tools for describing and analyzing the chemical reactivity of molecular systems. The proposed global descriptors provide information on the global behavior of chemical species as a whole, and the local descriptors give a profound understanding of the reactivity of a specific atomic site in a molecule during chemical interactions or excitations [7]. Whereas, the analyses of ESP on van der Waals surfaces are usually quantified to provide a good description of how the molecular electrostatic potential distribution interacts with long-extend molecules. More precisely, it gives a prediction of intermolecular non-covalent interactions such as hydrogen bonding, halogen bonding, and π-hole bonding by dissecting the magnitude and positions of minima and maxima ESP on van der Waals surface [8]. Nevertheless, the results obtained from the conceptual-DFT and the quantitative MEP analysis remain in need of experimentation in a biological medium. So, the molecular docking simulation provides a direct opportunity to place the studied molecular systems in a biological environment and follow the interactions that arise between them and the residues of the target binding sites, as well as, estimate theirs best binding modes and binding affinities [9].

In this chapter, a combination method based on conceptual-DFT formalism, quantitative MEP analysis, and molecular docking simulation was applied to investigate the chemical reactivity of six benzodiazepine basic rings. First, a statistical analysis was performed to assess the robustness of atomic charges to the basis sets. Global and local DFT-derived reactivity descriptors were determined and discussed to explain the global and local reactivity of the six studied analogs. Moreover, the quantitative MEP analysis on van der Waals surface was carried out to examine the long-range intermolecular interactions. Finally, a molecular docking procedure was conducted to predict the binding affinities of the issued molecules and estimate the binding poses into four binding sites, three of which were recently discovered, located in the $GABA_A$ receptor.

2 Material and methods

2.1 Statistical analysis

Before computing the condensed-to-atom reactivity descriptors; we should preliminarily determine the condensation method through the population analysis scheme. There are several schemes to achieve this objective. However, there is no consensus on which scheme is the most ideal one to study the reactive sites, through the condensed version of reactivity descriptors.

Indeed, it is difficult to choose the best electronic population analysis by referring exclusively to strictly theoretical arguments. In this study, we used a test statistic to select the most efficient method to assess the net atomic charges. The statistical quantity to be calculated for this purpose is defined according to *Lebart et al* [10].

$$t_k = \frac{\bar{q}_{k,p} - \bar{q}_k}{\sqrt{\frac{S_k^2}{n_p}\left(\frac{n-n_p}{n-1}\right)}} \qquad (1)$$

Where, $\bar{q}_{k,p}$ is the mean calculated for the k^{th} atom using the electronic population ‘’p’’, $\bar{q}_k$ is the mean calculated for the k^{th} atom using all electronic population schemes, n_p and S_k are the number of atomic charges in the electronic population ‘’p’’, and the standard deviation of q_k estimated for the k^{th} atom using all the electronic population schemes, respectively. The p-value or critical probability is computed as follows:

$$p.value = 1 - P(|Z| < |t_k|) = 1 - \phi(|t_k|) Z \sim N\ (0,1)$$

Where, $\emptyset(|t_k|)$ is the cumulative distribution function of the standard normal distribution. The p-value is computed under the null hypothesis ($H_0 : \bar{q}_{k,p} - \bar{q}_k = 0$). Indeed, if the p-value is less than or equal to significance level 0.05 ($1 - \emptyset(|t_k| \leq 0.05)$. Then, $\bar{q}_{k,p}$ is significantly different compared to the overall mean $\bar{q}_k$. In such a case, the corresponding charge characterizes the k^{th} atom. On the other hand, the alternative hypothesis is verified if ($H_1 : \bar{q}_{k,p} - \bar{q}_k \neq 0$).

2.2 Computational details

In order to determine the electronic properties of the benzodiazepines under study, the atomic charges were assessed using four electronic population schemes including Mulliken population analysis (MPA) [11], natural population analysis (NPA) [12], electrostatic method (ChelpG) [13], and Hirshfeld population analysis (HPA) [14]. For this purpose, the starting geometry of 3*H*-1,4-bdz was optimized in water using six different basis sets, 6-31G, 6-311G, 6-31$^+$G (d), 6-311$^+$G (d), 6-31^{++}G (d,p), and 6-311^{++}G (d,p) in combination with the four electron population schemes. Latter, the six benzodiazepine analogs (Figure 13, chapter I) were optimized in water as neutral structures at the DFT/Ub3lyp/6-311^{++}G(d,p) level of theory using the PCM solvation model. The molecular optimizations were performed using the Gaussian 09W software [15]. For all stationary points, there is no imaginary frequency at the optimized molecular geometries ensuring that the optimized structures are in the minimum on the potential energy surface. The atomic charges have been assessed using the Hirshfeld electronic population scheme. The singly charged anionic (N+1) and cationic (N-1) states were computed by single-point calculations at the same equilibrium geometry as the original molecule (neutral system).

The Fukui functions and the dual descriptor density mapped surfaces were visualized according to the FMO approximation. The condensed Fukui functions and the dual descriptor are calculated according to the formulas cited in chapter II. The quantitative analysis of MEP on van der Waals surface was conducted using the multifunctional wave function analyzer program Multiwfn 3.7 [16], combined with the Cubegen utility of Gaussian 09W software [15]. ESP-mapped molecular van der Waals surfaces were rendered by VMD 1.9.1 program [17] based on the outputs of Multiwfn.

2.3 Molecular docking protocol

Molecular docking simulation was performed into four benzodiazepine binding sites in the GABA$_A$ receptor using Moe 2014 software package [18]. The electron microscopy structures of Human GABA$_A$ receptor alpha1-beta2-gamma2 subtypes in complex with GABA plus the diazepam (DZP) structures (PDB ID:6X3X, Resolution= 2.92 Å) were downloaded from RCSB Database (http://www.rcsb.org).

Firstly, molecular docking was carried out into the classical benzodiazepine site located between the subunits D and E. In the protein preparation step, all co-crystallized ligands and non-essential subunits (A, B, and C) have been removed from the GABA$_A$ structure to provide sterically free cavities for ligand docking. Then, after structure correction, protonation, and cavity detection, the native co-crystallized DZP structure was re-docked in the selected binding site pocket and the best pose was chosen based on the given root-mean-square deviation (RMSD) values. Finally, the six neutral benzodiazepine structures previously optimized using the DFT method were converted into database files and docked into the DZP binding site pocket.

Likewise, molecular docking was also carried out in the binding sites of the trans-membrane domain located between the subunits (A and B), (C and D), and (A and E).

3 Results and discussion

3.1 Atomic charges

The change in atomic charges optimized using all the basis sets are shown in Appendix A. The performance of each electron population has been quantified by computing the ratio of critical probabilities. This performance indicator measures the robustness of each method to the basis set. The statistical results obtained are collected in Tables III.1 and III.2.

Table III.1. The statistical test results for Mulliken and NBO populations.

Atom	Mulliken					NPA				
	t_k	$\bar{q}_{k,p}$	$\bar{q}_k$	S_k	p-value	t_k	$\bar{q}_{k,p}$	$\bar{q}_k$	S_k	p-value
1 N	1.529	-0.201	-0.372	0.221	0.063	-1.561	-0.463	-0.372	0.221	**0.059**
2 C	-2.870	0.035	0.121	0.087	**0.002**	1.567	0.144	0.121	0.087	**0.059**
3 C	-2.510	-0.379	-0.059	0.368	**0.006**	-2.178	-0.337	-0.059	0.368	**0.015**
4 N	2.431	-0.191	-0.391	0.237	**0.008**	-0.733	-0.451	-0.391	0.237	0.232
5 C	-3.692	-0.348	0.062	0.320	**0.000**	0.329	0.099	0.062	0.320	0.371
6 C	-2.868	-0.340	-0.183	0.157	**0.002**	-0.013	-0.184	-0.183	0.157	0.495
7 C	-1.177	-0.164	-0.130	0.083	0.119	-3.329	-0.225	-0.130	0.083	**0.000**
8 C	-2.659	-0.225	-0.129	0.104	**0.004**	-2.012	-0.202	-0.129	0.104	**0.022**
9 C	-2.470	-0.328	-0.211	0.136	**0.007**	-0.074	-0.215	-0.211	0.136	0.470
10 C	-3.408	-0.364	0.072	0.368	**0.000**	0.494	0.135	0.072	0.368	0.311
11 C	3.982	0.922	0.155	0.555	**0.000**	-1.570	-0.148	0.155	0.555	**0.058**
12 H	2.551	0.202	0.130	0.080	**0.005**	2.939	0.213	0.130	0.080	**0.002**
13 H	2.479	0.218	0.107	0.128	**0.007**	2.797	0.232	0.107	0.128	**0.003**
14 H	2.270	0.205	0.109	0.122	**0.012**	3.003	0.236	0.109	0.122	**0.001**
15 H	2.456	0.197	0.113	0.099	**0.007**	2.947	0.214	0.113	0.099	**0.002**
16 H	1.371	0.190	0.156	0.072	0.085	3.242	0.237	0.156	0.072	**0.001**
17 H	1.544	0.187	0.147	0.075	0.061	3.497	0.238	0.147	0.075	**0.000**
18 H	1.536	0.188	0.149	0.073	0.062	3.474	0.238	0.149	0.073	**0.000**
19 H	1.483	0.194	0.155	0.076	0.069	3.258	0.241	0.155	0.076	**0.001**

$\bar{q}_{k,p}$: Mean calculated for the k^{th} atom using the electronic population p.

$\bar{q}_k$: Mean calculated for the k^{th} atom using all electronic population schemes.

S_k: The standard deviation of q_k estimated for the k^{th} atom using all the electronic population schemes

Out of 19 critical probabilities P-value, the alternative hypothesis $(1-\varnothing(|t_k| \leq 0.05)$ was verified by the following ratios: 79% (HPA), 74% (NPA and ChelpG), and 68% (MPA).

MPA seems to be the lowest-performing method, and this is not surprising as several objections to its results have been discussed in literature [18,19]. The drawback of MPA is that it is based on the one-particle density matrix defined through non-orthogonal atomic orbital basis sets [11]. Therefore, its sensitivity toward the basis sets is very high. This latter is evident from the data provided in Table III.1, where we note that despite correlated Gaussian basis sets were used, the charges estimated using MPA are cover a wide range of values. Therefore, the MPA method is strongly affected by the used basis set.

Table III.2. The statistical test results for ChelpG and Hirshfeld populations.

Atom	**ChelpG**					**Hirshfeld**				
	$\mathbf{t_k}$	$\mathbf{\bar{q}_{k,p}}$	$\mathbf{\bar{q}_k}$	$\mathbf{S_k}$	**p-value**	$\mathbf{t_k}$	$\mathbf{\bar{q}_{k,p}}$	$\mathbf{\bar{q}_k}$	$\mathbf{S_k}$	**p-value**
1 N	-3.758	-0.660	-0.372	0.221	**0.000**	2.698	-0.166	-0.372	0.221	**0.003**
2 C	4.130	0.246	0.121	0.087	**0.000**	-2.006	0.061	0.121	0.087	**0.022**
3 C	4.376	0.499	-0.059	0.368	**0.000**	0.312	-0.019	-0.059	0.368	0.377
4 N	-4.168	-0.734	-0.391	0.237	**0.000**	2.470	-0.188	-0.391	0.237	**0.007**
5 C	3.593	0.461	0.062	0.320	**0.000**	-0.230	0.037	0.062	0.320	0.409
6 C	0.128	-0.177	-0.183	0.157	0.449	2.753	-0.033	-0.183	0.157	**0.003**
7 C	1.558	-0.089	-0.130	0.083	**0.057**	3.079	-0.042	-0.130	0.083	**0.001**
8 C	1.973	-0.058	-0.129	0.104	**0.024**	2.698	-0.032	-0.129	0.104	**0.003**
9 C	-0.875	-0.253	-0.211	0.136	0.191	3.419	-0.050	-0.211	0.136	**0.000**
10 C	3.185	0.479	0.072	0.368	**0.001**	-0.270	0.037	0.072	0.368	0.393
11 C	-1.554	-0.134	0.155	0.555	**0.057**	-0.913	-0.021	0.155	0.555	0.181
12 H	-2.859	0.051	0.130	0.080	**0.002**	-2.632	0.057	0.130	0.080	**0.004**
13 H	-3.876	-0.066	0.107	0.128	**0.000**	-1.577	0.045	0.107	0.128	**0.059**
14 H	-3.813	-0.053	0.109	0.122	**0.000**	-1.590	0.047	0.109	0.122	**0.057**
15 H	-3.559	-0.009	0.113	0.099	**0.000**	-1.844	0.050	0.113	0.099	**0.033**
16 H	-0.620	0.140	0.156	0.072	0.268	-3.994	0.056	0.156	0.072	**0.000**
17 H	-1.504	0.107	0.147	0.075	**0.051**	-3.497	0.056	0.147	0.075	**0.000**
18 H	-1.418	0.113	0.149	0.073	0.078	-3.592	0.058	0.149	0.073	**0.000**
19 H	-0.704	0.136	0.155	0.076	0.241	-4.038	0.049	0.155	0.076	**0.000**

$\bar{q}_{k,p}$: Mean calculated for the k^{th} atom using the electronic population *p*.

$\bar{q}_k$: Mean calculated for the k^{th} atom using all electronic population schemes.

S_k: The standard deviation of q_k estimated for the k^{th} atom using all the electronic population schemes

NPA and ChelpG analysis provide the same performance. Here, the charges resist better than the charges estimated using MPA, since they seem to be independent of the used basis sets. NPA was developed to improve the problems related to the MPA method. Thus, the working base of NPA is the construction of an orthonormal set of natural atomic orbitals (NAOs) covering the non-orthogonal basis orbitals space. The occupancy weighted symmetric orthogonalization (OWSO) procedure was applied to convert the non-orthogonal atomic orbitals into the orthonormal set, where the highest occupancy orbitals are strongly preserved in form, while orbitals with negligible occupancy can be freely deformed to achieve orthogonality [12]. This feature makes NPA insensitive to the basis set quality.

ChelpG is one of the grid-based methods derived from electrostatic potentials. In which, atomic charges are adapted to reproduce the MEP at a certain number of points around the molecular species. The methods based on ESP are limited for small molecular systems with low flexibility [20]. On larger systems, it is difficult to assign charges when the atoms are far from the points where the MEP is computed. This means that the extra-molecular MEP did not accurately determine the values of the internal atomic charges. And therefore, the predicted charges become unrealistic.

The best performance is obtained by using the HPA method. The HPA values given in Table III.2 are the lowest among the considered charges and cover a small range of values sometimes equal to zero. Because the HPA scheme completely neglects the effect of atomic dipole moments.

It's good for us that Hirshfeld is the best charge population to detect the electronic properties for the studied structures since our studies are heavily based on the results of Fukui indices and the dual descriptor. Hence, HPA is the ideal solution to produces a non-negative condensed Fukui function (FF) compared to MPA, NPA, and ChelpG schemes which in some cases generate negative values. The case of negative condensed FF is systematically unacceptable, except in the case where the effects of orbital relaxation are so important that redox-induced electron rearrangement occurs [21–24].

3.2 Geometry and electronic properties

To illustrate the differences between the structural parameters and the electronic properties of BDZ structures, the bond lengths, bond angles, and dihedral angles were collected in appendix B and revealed in Figure III.1. Table III.3 gives the HPA charges assigned to each atom.

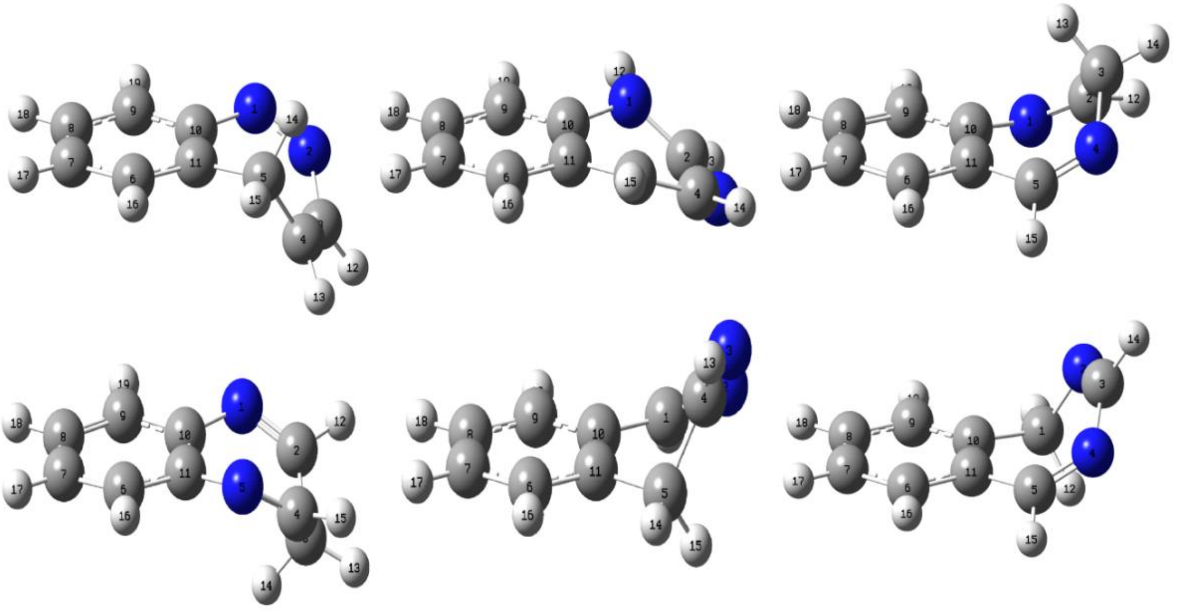


Figure III.1. Structural deformations of studied BDZ.

Phenyl rings present a hexagonal geometry with C-C bond length and bond angles vary from (1.383 A° to 1.423 A°) and from (117.961° to 121.965°), respectively. Accordingly, the binding of the diazepine ring in C_{10} and C_{11} resulted in a slight distortion in the bond lengths and bond angles from the normal values 1.40 A° and 120°. Consequently, the planarity was also affected as indicated by the torsion angles.

The diazepine rings present non-planar structures, the significant deviations out of the phenyl plan are started from the atoms numbered 1 and 5, with dihedral angles 1-10-11-6 and 9-10-11-5 vary from (-179.465° to 178.861°) and from (-48.360° to 65.764°), respectively. These distortions differ from one structure to another depending on the molecular flexibility. The bonds that have the lowest values of lengths are considered the strongest bonds in the structures and a large amount of energy is required to break them, while the weakest bonds need less energy. Here, the lowest values of bond lengths are similar between all the structures (1.27 A°). Unlike, slight differences are appeared between the lengths of weak bonds.

Table III.3. Hirshfeld charges assigned to each atom, expressed in atomic unit (a.u).

Atom	**5*H*-1,2-bdz**	**1*H*-1,3-bdz**	**3*H*-1,4-bdz**	**3*H*-1,5-bdz**	**5*H*-2,3-bdz**	**1*H*-2,4-bdz**
1	-0.095	-0.095	-0.163	-0.170	0.030	-0.019
2	-0.104	0.094	0.064	0.066	-0.156	-0.200
3	-0.026	-0.220	-0.020	-0.052	-0.153	0.069
4	-0.013	-0.020	-0.185	0.066	0.049	-0.179
5	-0.045	-0.067	0.039	-0.170	-0.049	0.071
6	-0.042	-0.041	-0.032	-0.055	-0.041	-0.027
7	-0.032	-0.054	-0.041	-0.045	-0.031	-0.038
8	-0.041	-0.042	-0.031	-0.045	-0.041	-0.026
9	-0.043	-0.058	-0.049	-0.055	-0.032	-0.039
10	0.011	0.042	0.038	0.022	-0.018	0.006
11	0.007	-0.023	-0.019	0.022	0.012	-0.016
12	0.052	0.141	0.055	0.056	0.051	0.038
13	0.061	0.056	0.044	0.061	0.057	0.040
14	0.047	0.038	0.045	0.054	0.057	0.044
15	0.055	0.046	0.048	0.056	0.049	0.055
16	0.053	0.050	0.054	0.043	0.054	0.057
17	0.055	0.049	0.054	0.051	0.055	0.055
18	0.054	0.052	0.056	0.051	0.053	0.056
19	0.046	0.053	0.046	0.043	0.055	0.054

According to Table III.3, if we exclude the results of 1*H*-1,3-bdz we note that, in all the remaining structures, the least positive charges appear over the junction atoms C_{10} or C_{11} and the high negative charges appear over the two nitrogen atoms. In another hand, hydrogens also provide important positive charges especially H_{12} in 1*H*-1,3-bdz which has the highest positive charge between all studied systems, and this due to its direct attachment to the high-electron withdrawing atom N_1.

3.3 Global reactivity

The global reactivity indices derived from the conceptual-DFT (χ, η, S, ω, N) are calculated and discussed to explain the global changes between the six benzodiazepine analogs. Table III.4 summarizes the results.

Table III.4. Values of HOMO-LUMO gap (Δ), global reactivity indices (χ, η, S, ω, N), and dipole moment (DM).

basic rings	E_{HOMO} (eV)	E_{LUMO} (eV)	Δ (eV)	A (eV)	I (eV)	χ (eV)	η (eV)	S (eV^{-1})	ω (eV)	N (eV)	DM (Debye)
5*H*-1,2-bdz	-6.460	-2.503	3.957	2.645	6.402	4.524	1.878	0.266	5.447	2.615	5.388
1*H*-1,3-bdz	-5.751	-1.858	3.893	2.031	5.644	3.837	1.807	0.277	4.076	3.323	3.742
3*H*-1,4-bdz	-6.726	-1.832	4.895	1.992	6.620	4.306	2.314	0.216	4.006	2.348	2.570
3*H*-1,5-bdz	-6.668	-1.519	5.148	1.702	6.563	4.133	2.430	0.206	3.514	2.407	3.313
5*H*-2,3-bdz	-7.004	-1.851	5.154	2.004	6.898	4.451	2.447	0.204	4.048	2.070	5.907
1*H*-2,4-bdz	-6.902	-2.229	4.673	2.371	6.793	4.582	2.211	0.226	4.749	2.172	4.688

$A= E_N - E_{N+1}$; $I= E_{N-1}-E_N$; $\Delta= | E_{HOMO} - E_{LUMO} |$; $\chi= - \mu= 1/2(A+I)$; $\eta= 1/2(I-A)$; $S = 1/2\eta$; $\omega = \mu^2/2\eta$

$N = E_{HOMO} - E_{HOMO\ (TCE)}$ with $E_{HOMO\ (TCE)} = -9.074535$ eV, calculated with DFT (Ub3lyp) /6-311^{++}G (d, p)/ Hirshfeld/PCM model.

The overall hardness classifies the studied systems from the most reactive to the least, as follows: 1*H*-1,3-bdz, 5*H*-1,2-bdz, 1*H*-2,4-bdz, 3*H*-1,4-bdz, 3*H*-1,5-bdz, and 5*H*-2,3-bdz. The HOMO-LUMO energy gap (Δ) confirmed the results of global hardness, where, the high kinetic stability (high Δ) is generally associated with the low chemical reactivity (high η).

The high reactivity of 1*H*-1,3-bdz is accompanied by the highest nucleophilicity (N= 3.323 eV) and the lowest electronegativity (χ =3.837 eV). Whereas, the low reactivity of 5*H*-2,3-bdz is accompanied by the lowest nucleophilicity (N=2.070 eV) and the highest polarization (DM=5.907 Debye). Thus, the intermolecular interactions between 5*H*-2,3-bdz and the neighboring water molecules are of great importance compared to the rest of the studied systems. Using the nucleophilicity and electrophilicity scales [25,26], the 1*H*-1,3-bdz is classified as a strong nucleophile with strong electrophilicity, while the other systems are classified as moderate nucleophiles with strong electrophilicity.

3.4 Local reactivity and quantitative MEP analysis

In this section, the selectivity of atomic sites towards the approach of nucleophilic and electrophilic reactants is investigated using the dual descriptor $f^{(2)}(r)$. The numerical values are depicted in Tables III.5 and III.6. The numerical values of Table III.6 are mapped in Figure III.2 (for $f^+(r)$ and $f^-(r)$ mapped surfaces see Appendix C) through the use of the FMO approximation. The structures are arranged and numbered in the order given in Figure 13 of chapter I. Regions of positive signs are depicted by purple color, and regions of negative signs are in white (Isovalue MO = 0.0015 a.u, density = 0.00040 a.u).

Table III.5. Values of condensed Fukui Functions $f^+(r)$ and $f^-(r)$ evaluated in the terms of anionic and cationic spin-densities, respectively.

atom	**5*H*-1,2-bdz**		**1*H*-1,3-bdz**		**3*H*-1,4-bdz**		**3*H*-1,5-bdz**		**5*H*-2,3-bdz**		**1*H*-2,4-bdz**	
	$f^+(r)$	$f^-(r)$	$f^+(r)$	$f^-(r)$	$f^+(r)$	$f^-(r)$	$f^+(r)$	$f^-(r)$	$f^+(r)$	$f^-(r)$	$f^+(r)$	$f^-(r)$
1	0.278	0.346	0.028	0.206	0.038	0.237	0.114	0.118	0.185	0.044	0.010	0.040
2	0.274	0.335	0.235	0.025	0.199	0.058	0.133	0.073	0.188	0.253	0.123	0.382
3	0.007	0.079	0.048	0.115	0.023	0.033	0.032	0.016	0.032	0.260	0.051	0.013
4	0.170	0.022	0.157	0.125	0.214	0.123	0.133	0.073	0.110	0.009	0.159	0.214
5	0.014	0.005	0.118	0.138	0.070	0.015	0.114	0.118	0.011	0.009	0.257	-0.003
6	0.000	0.006	0.046	0.004	0.022	0.072	0.119	-0.012	0.027	0.067	0.092	-0.007
7	0.076	0.030	0.023	0.065	0.055	0.167	0.053	0.123	0.144	0.137	-0.008	0.093
8	-0.009	0.016	0.104	0.070	0.098	-0.023	0.053	0.123	-0.007	-0.009	0.124	0.019
9	0.065	-0.001	-0.006	0.005	-0.007	0.142	0.119	-0.012	0.114	0.106	0.000	0.024
10	0.018	0.102	0.098	0.110	0.120	0.095	0.004	0.176	0.061	0.104	0.085	0.089
11	0.061	0.044	0.077	0.106	0.106	0.039	0.004	0.176	0.071	0.002	0.047	0.078
12	0.008	0.001	0.004	0.007	0.020	0.011	0.026	0.004	0.018	0.005	0.004	0.016
13	0.016	0.011	0.023	0.005	0.006	0.006	0.010	0.001	0.010	-0.002	0.001	0.010
14	0.005	-0.001	0.017	0.007	0.004	0.005	0.026	0.007	0.002	0.001	0.012	0.006
15	0.002	0.000	0.010	0.008	0.011	0.003	0.026	0.004	0.003	-0.001	0.022	0.020
16	0.001	0.001	0.005	-0.001	0.004	0.003	0.011	-0.001	0.003	0.003	0.009	-0.001
17	0.008	0.002	0.003	0.004	0.006	0.009	0.006	0.006	0.016	0.007	0.000	0.006
18	-0.001	0.001	0.012	0.004	0.011	-0.002	0.006	0.006	0.000	-0.001	0.012	0.001
19	0.006	0.002	0.000	-0.001	0.002	0.007	0.011	-0.001	0.013	0.004	0.001	0.001

Our results show that the delocalization of nitrogen along the diazepine cycle leads to significant differences in local reactivity. As a result, the ability of each atom to receive nucleophilic or electrophilic attack varies from structure to another and this may explain the wide biodiversity of the benzodiazepine family.

Table III.6. Values of dual descriptor $\boldsymbol{f^{(2)}(r)}$ evaluated in the term of spin-density, expressed in atomic units (a.u).

atoms	**5*H*-1,2-bdz**	**1*H*-1,3-bdz**	**3*H*-1,4-bdz**	**3*H*-1,5-bdz**	**5*H*-2,3-bdz**	**1*H*-2,4-bdz**
1	-0.068	-0.177	-0.199	-0.004	0.141	-0.030
2	-0.061	0.209	0.141	0.060	-0.064	-0.259
3	-0.071	-0.067	-0.010	0.016	-0.228	0.038
4	0.148	0.032	0.091	0.060	0.100	-0.055
5	0.009	-0.020	0.055	-0.004	0.002	0.260
6	-0.006	0.043	-0.050	0.131	-0.040	0.099
7	0.047	-0.042	-0.112	-0.070	0.007	-0.101

Table III.6. Continued

8	-0.026	0.034	0.121	-0.070	0.002	0.104
9	0.066	-0.012	-0.150	0.131	0.008	-0.024
10	-0.084	-0.011	0.025	-0.173	-0.043	-0.004
11	0.017	-0.029	0.066	-0.173	0.069	-0.031
12	0.007	-0.003	0.008	0.022	0.013	-0.012
13	0.005	0.018	0.000	0.009	0.012	-0.008
14	0.007	0.010	-0.001	0.019	0.000	0.006
15	0.002	0.001	0.008	0.022	0.004	0.003
16	0.000	0.006	0.001	0.012	0.000	0.010
17	0.006	-0.001	-0.003	0.000	0.009	-0.006
18	-0.002	0.008	0.013	0.000	0.001	0.011
19	0.004	0.001	-0.005	0.012	0.008	0.000

The preferred sites for nucleophilic (electrophilic) reagents are easily determined by the highest positive (negative) value of $f^{(2)}(r)$. The symmetrical behavior of 3*H*-1,5-bdz caused it to have two identical preferred electrophilic sites are C_6 and C_9, and two identical preferred nucleophilic sites C_{10} and C_{11}. In contrast, the $f^{(2)}(r)$ values around the nitrogen atoms N_1 and N_5 are close to zero (0.004), signifying that they are neither nucleophiles nor electrophiles. Thus, the possibility of attacks on these sites is relatively negligible. The same notes for hydrogens, since most of them apparent inactive towards both attacks, and this is the only observation that can be generalized to the six studied structures.

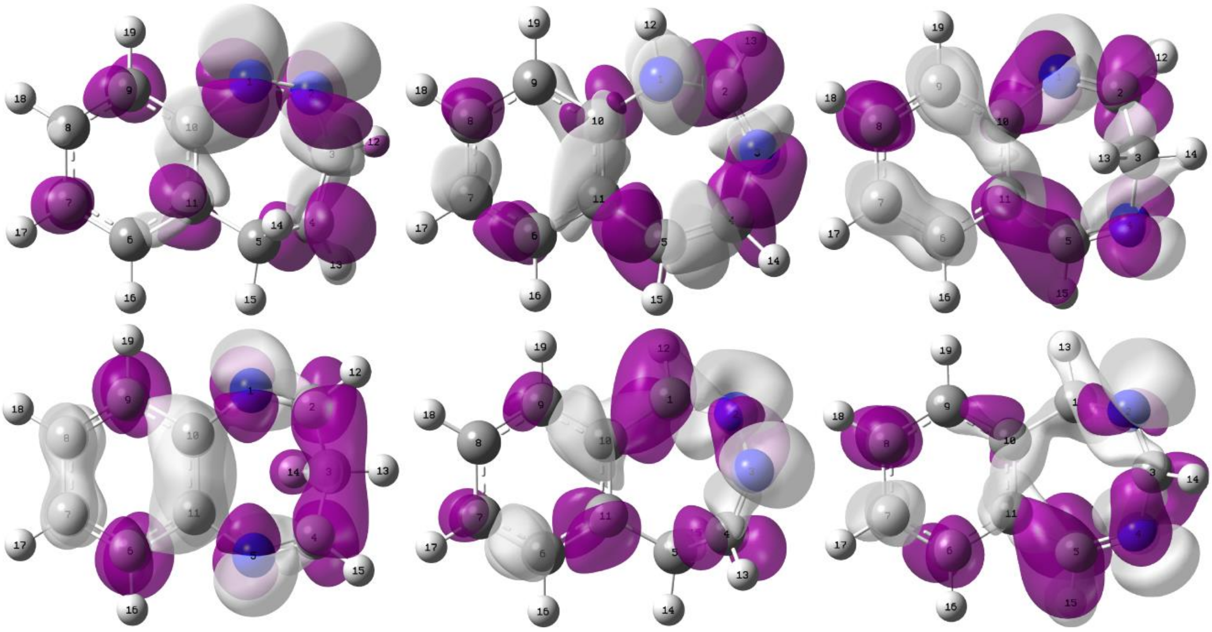


Figure III.2. Dual descriptor 3D-mapped surfaces, mapped through (FMO) approximation, Isovalue MO = 0.0015 a.u, purple color: regions of nucleophilic attacks, white color: regions of electrophilic attacks.

Recently, drugs based on 3*H*-1,4-bdz structure are the wide prescribed among all the benzodiazepines. The QSAR researches [27,28] suggested that the presence of an electron-donating group in position 1, carbonyl oxygen (C=O) in position 2, phenyl group in position 5, and an electron-withdrawing group in position 7 are necessary for increasing their affinity toward the $GABA_A$ receptor. Here, we will attempt to discuss the possibility of attaching these substituents in the required positions based on the results obtained from the dual descriptor.

The values of dual descriptors suggest that N_1 and C_2 are the first privileged sites for electrophilic and nucleophilic attacks, respectively. This makes it easy to fix an electron-donating group in position 1 and a carbonyl oxygen group in position 2, using the electrophilic and nucleophilic attacks. On the other hand, the positive $f^{(2)}(r)$ value in C_5 is relatively small. Therefore, the attachment of a phenyl group in this position is possible but not much preferred since the nucleophilic attack will be directed mainly towards position 2. Finally, our results suggest that C_7 is the third preferred nucleophilic site. Accordingly, the direct attachment of an electron-withdrawing group in this position is not recommended.

To predict the intermolecular interaction between the studied structures and the distant reagents; the quantitative MEP analysis was performed and discussed. The repulsive and attractive electrostatic interactions are long-range in comparison to the charge-transfer effects characterized by the Fukui functions and the derived reactivity indices [29]. MEP in the vicinity of a molecule is defined as the energy required bringing a unit test positive charge from infinity to the point **r**. Statistically-based molecular descriptors derived from molecular surface electrostatic potentials studied here have been defined by *Politzer et al* [30].

ESP-mapped van der Waals surfaces are depicted in Figure III.3. The different colors in plots representing the different values of ESP at the surface, the color code is in the range between -39.51 kcal/mol (deepest red), and 35.23 kcal/mol (deepest blue). Areas with positive ESP are electron-deficient sites and therefore subject to nucleophilic attacks, and the negative areas are electron-rich sites reacting with electrophilic reagents. The positive and negative values with a star are the global maximum and minimum ESP vdW-surface. These values indicate the most privileged sites for distant nucleophiles and electrophiles.

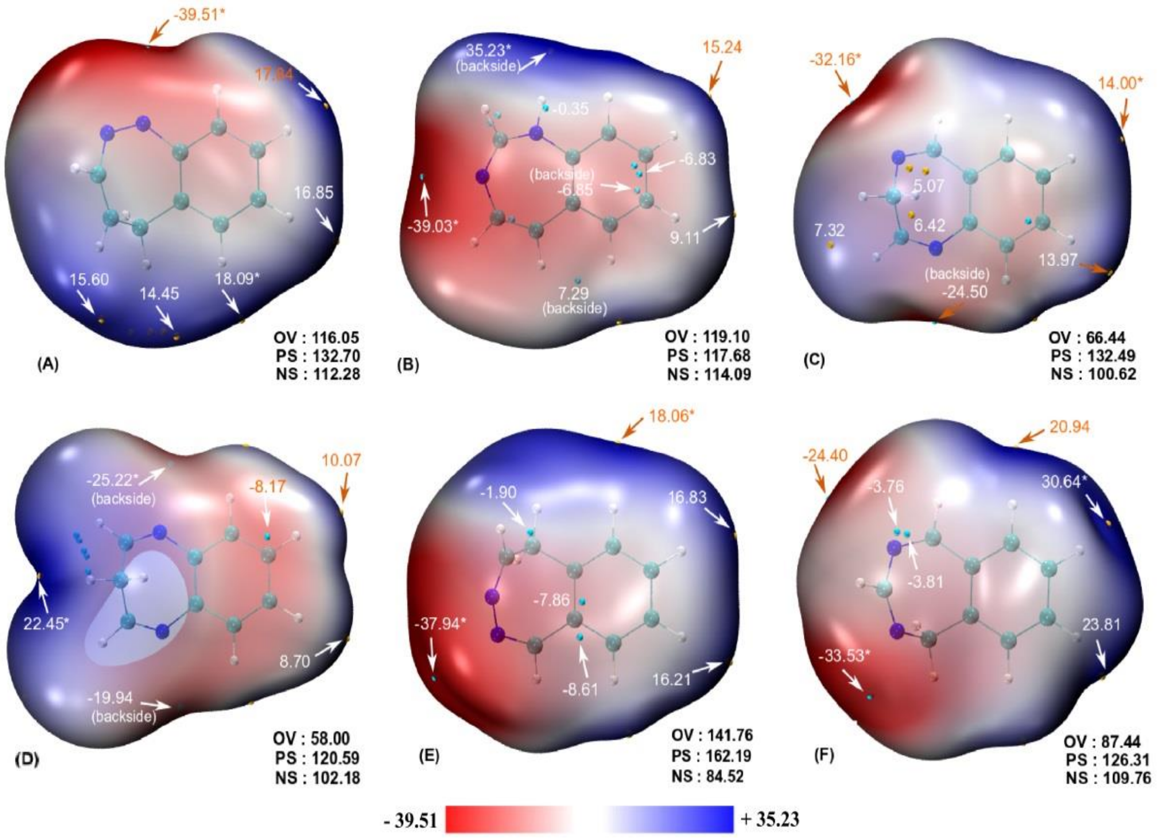


Figure III.3. ESP-mapped van der Waals surfaces (kcal/mol) using a color scale ranging fromred (negative ESP), through white (neutral ESP) to blue (positive ESP). All the iso-surface maps were rendered by VMD software based on the surface analysis result of Multiwfn program. The grid spacings were set to 0.2 *Bohr* and the van der Waals surface denotes the iso-surface of $\rho = 0.001\ a.u.$ Values with a star indicate global extremums. The bold numbers in the bottom right-hand corner are the overall ESP variance (OV), positive surface area (PS) and negative surface area (NS) whose unit are $[\text{Kcal/mol}]^2$, $(\text{A°})^2$, respectively.

For more accuracy, the ESP vdW-surfaces for each atom (V_k) are given in Table III.7. The atoms not mentioned in the table are having ESP values close to zero and are therefore not susceptible to any attack. The overall variance (OV) reflects the tendency of molecules to interact electrostatically with long-rang reagents. The highest OV value is 141.760 $(\text{kcal\textbackslash mol})^2$. This implies that the strong tendency for long-rang interactions is attributed to 5*H*-2,3-bdz, followed by 1*H*-1,3- bdz, 5*H*-1,2- bdz, 1*H*-2,4- bdz, 3*H*-1,4- bdz, and 3*H*-1,5- bdz.

Table III.7. Atomic ESP van der Waals surfaces (V_k), expressed in (kcal/mol).

atoms	5*H*-1,2-bdz	1*H*-1,3-bdz	3*H*-1,4-bdz	3*H*-1,5-bdz	5*H*-2,3-bdz	1*H*-2,4-bdz
1	-24.521	-5.471	-15.684	-15.797	-10.523	/
2	-21.236	-5.983	0.500	1.197	-24.795	-20.295
3	-4.861	-23.371	/	/	-24.869	-10.395
4	-1.742	-12.025	-18.542	1.106	-5.574	-14.107

Table III.7. Continued

5	/	-11.252	-1.598	-13.036	/	-1.452
6	-1.478	-7.462	-2.717	-9.884	0.146	0.083
7	-2.676	-5.653	-2.704	-9.454	-0.434	0.773
8	-4.810	-4.966	-4.173	-10.085	-1.254	-0.839
9	-7.166	-3.255	-6.106	-11.339	-2.449	-2.801
10	-6.868	-1.122	-8.442	/	-2.575	-2.929
11	-3.751	-10.793	-6.505	/	-0.099	-1.615
12	-2.064	17.254	-0.376	3.137	-2.639	-3.858
13	9.534	2.958	2.197	14.755	1.724	-3.097
14	0.044	-7.989	0.332	3.965	10.771	-8.757
15	10.592	-2.952	-3.859	3.207	2.196	1.157
16	10.494	1.219	6.267	-5.388	10.895	9.614
17	8.452	3.004	6.725	1.187	9.257	10.365
18	5.367	6.752	6.838	1.744	8.338	9.972
19	-6.724	13.767	-1.697	-5.765	5.252	4.796

If we compare this reactivity order with that given by the global hardness discussed in the global reactivity section, it is clear that both are completely identical except in the case of 5*H*-2,3-bdz where the global hardness orders it as the least reactive structure. The distribution of negative surfaces area (NS) on 5*H*-2,3-bdz and 1*H*-1,3-bdz occupy 34.3% and 49.2% of the total surfaces, respectively. This makes 5*H*-2,3-bdz have the least nucleophilic surface and 1*H*-1,3-bdz have the most nucleophilic surface among the studied systems. This result confirming the results of the global nucleophilicity indices (N) given in Table III.4. Simultaneously, the positive and negative surfaces area in 1*H*-1,3-bdz have rather close values, especially when compared to the other studied systems where the negative surfaces area (NS) are significantly lower than the positive surfaces area (PS). This confirmed the classification of 1*H*-1,3-bdz as strong nucleophiles with strong electrophilicity, and the remains as moderate nucleophiles with strong electrophilicity.

According to Figure III.3, each structure is surrounded by positive and negative ESP surfaces that appeared over the hydrogen atoms and the two nitrogen atoms, respectively. The highest local positive ESP values on the surfaces are appearing around: (H_{12} in 1*H*-1,3-bdz), (H_{13} in 3*H*-1,5-bdz), (H_{15} in 5*H*-1,2-bdz), (H_{16} in 5*H*-2,3-bdz), and (H_{17} in 3*H*-1,4-bdz and 1*H*-2,4-bdz), Indicating the primary favorable sites for distant nucleophiles. simultaneously, the highest local negative ESP values on the surfaces are distributed around the lone pairs of (N_1 in 5*H*-1,2-bdz and 3*H*-1,5-bdz), (N_2 in 1*H*-2,4-bdz), (N_3 in 1*H*-1,3-bdz and 5*H*-2,3-bdz), and (N_4 in 3*H*-1,4-bdz), Thus are the primary preferred sites for distant electrophiles. Also, most carbons are having negative ESP regions but less than that of the nitrogen, this negativity due to the π-electron cloud delocalization over the benzene and diazepine rings. Therefore, the possibility of electrophilic attacks in these areas is also acceptable.

3.5 Molecular docking simulation

The molecular docking simulation provides a direct opportunity to place the six studied structures in a biological environment and follow the interactions that arise between them and the residues of the binding sites. Moe docking was performed in four distinct benzodiazepine binding sites: the classical site at the ECD α-γ interface, the two TMD sites at the α-β interfaces, and the TMD site at the β-γ interface (see Figure 11, chapter I). Firstly, the co-crystallized ligand diazepam (DZP) has been re-docked into the four binding sites and the best poses were chosen according to the RMSD values and visualized in Figure III.4. Then, the six studied ligands have been docked into each binding site, and the best poses were selected according to the lowest energy score values.

Results such as the energy score, types, and distances of interactions between the studied ligands and the residues of the binding sites were collected in Tables III.8 to III.11. The docking scoring is required to quantitatively estimate the binding affinity between the target macromolecule and ligands; the best poses are those that have an energy score closer to that of the co-crystallized ligand. In the four docking results, Diazepam (DZP) gives lower energy scores than the docked ligands, the Cl substitution in position 7 forms an average bond type H-Donor with His D102 and Asp E297 in sites (a) and (d), respectively. The π-electron cloud delocalization over the phenyl ring attached in position 5 forms a withdrawing group electrostatically attracting the hydrogen of Pro D233 in the site (b). While the interactions between the basic structure and the binding site residues are appearing only in sites (a) and (c), between the N_4 and the π-electron cloud of the phenyl ring in Phe E77 and Phe A289, respectively.

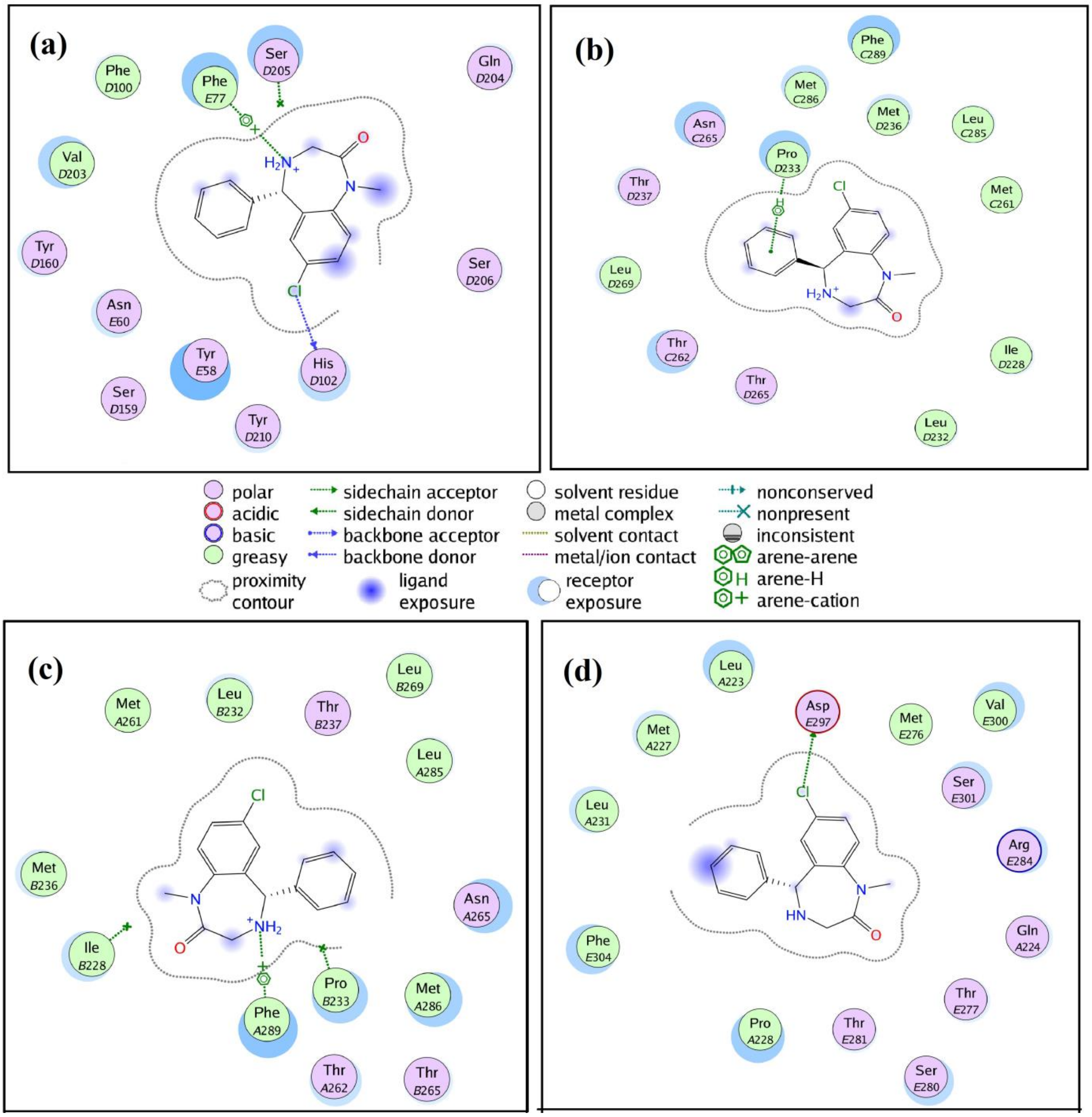


Figure III.4. The best binding poses for the re-docked co-crystallized ligand (Diazepam) in the classical site at the ECD α-γ interface (a), the two TMD sites at the α-β interfaces (b) and (c), and the TMD site at β-γ interface (d). RMSD = 0.279 A°, 0.357 A°, 0.487 A°, and 0.314 A°, respectively.

The score of binding free energy of the six docked ligands is between (-5.021 and -5.235) Kcal/mol, (-4.889 and -5.054) Kcal/mol, (-4.624 and -4.951) Kcal/mol, and (-4.479 and -4.644) Kcal/mol for the sites (a), (b), (c), and (d), respectively. As it clear, the six ligands rather have affinities close to each other and simultaneously far from that of DZP. Despite this convergence in affinities, each ligand has a different mode of interaction with the target binding sites, and this due to the difference in ESP generated by each structure (see Figure III.3). This latter considered being the primary responsibility in the orientation of intermolecular interactions.

Table III.8. S-score, bond interactions, bond distances, and bond energy for the co-crystallized Ligand (Diazepam) and the six ligands docked in the classical BDZ site at the extracellular domain (ECD) α-γ interface (site (a)).

	S-score (Kcal/mol)	Atom of Ligand	Involved receptor atoms	Involved receptor residues	Type of interactions	Distances (A°)	Bond Energy (kcal/mol)
			Co-crystallized Ligand				
rmsd = 0.279895216	-7.00339508	CL 36	O	HIS 102 (D)	H-donor	3.38	-1.7
		N15 23	6-ring	PHE 77 (E)	cation-pi	3.97	-2.2
			Complexe GABA$_A$-ligand				
5*H*-1,2-bdz	-5.23495197	/	/	/	/	/	/
1*H*-1,3-bdz	-5.0283432	N 3	N	ALA 161 (D)	H-acceptor	3.52	-2.1
3*H*-1,4-bdz	-5.0801158	/	/	/	/	/	/
3*H*-1,5-bdz	-5.18940687	/	/	/	/	/	/
5*H*-2,3-bdz	-5.02146149	N 3	OG1	THR 207 (D)	H-acceptor	3.06	-1.0
1*H*-2,4-bdz	-5.04629469	/	/	/	/	/	/

Table III.9. S-score, bond interactions, bond distances, and bond energy for the co-crystallized ligand (Diazepam) and the six ligands docked in the transmembrane domain (TMD) site at the α-β interface (site (b)).

	S-score (Kcal/mol)	Atom of Ligand	Involved receptor atoms	Involved receptor residues	Type of interactions	Distances (A°)	Bond Energy (kcal/mol)
			Co-crystallized Ligand				
rmsd = 0.357037723	-6.26063156	6-ring	CB	PRO 233 (D)	pi-H	3.67	-0.9
			Complexe GABA$_A$-ligand				
5*H*-1,2-bdz	-5.05429792	/	/	/	/	/	/
1*H*-1,3-bdz	-4.9483633	C 9	6-ring	PHE 289 (C)	H-pi	3.80	-0.6
3*H*-1,4-bdz	-4.88927412	6-ring	6-ring	PHE 289 (C)	pi-pi	3.98	-0.0
3*H*-1,5-bdz	-4.91819811	/	/	/	/	/	/
5*H*-2,3-bdz	-4.90998173	6-ring	6-ring	PHE 289 (C)	pi-pi	3.97	-0.0
1*H*-2,4-bdz	-4.89463186	6-ring	6-ring	PHE 289 (C)	pi-pi	3.94	-0.0

Table III.10. S-score, bond interactions, bond distances, and bond energy for the co-crystallized ligand (Diazepam) and the six ligands docked in the transmembrane domain (TMD) site at the α-β interface (site (c)).

	S-score (Kcal/mol)	Atom of Ligand	Involved receptor atoms	Involved receptor residues	Type of interactions	Distances (A°)	Bond Energy (kcal/mol)
			Co-crystallized Ligand				
rmsd = 0.487408489	-6.15904188	N15 23	6-ring	PHE 289 (A)	cation-pi	4.22	-0.7
			Complexe GABA$_A$-ligand				
5*H*-1,2-bdz	-4.95116711	/	/	/	/	/	/
1*H*-1,3-bdz	-4.92590475	7-ring	CD	ARG 269 (A)	pi-H	4.73	-0.9
3*H*-1,4-bdz	-4.69323301	/	/	/	/	/	/
3*H*-1,5-bdz	-4.62413931	/	/	/	/	/	/
5*H*-2,3-bdz	-4.64065981	/	/	/	/	/	/
1*H*-2,4-bdz	-4.79273367	/	/	/	/	/	/

Table III.11. S-score, bond interactions, bond distances, and bond energy for the co-crystallized ligand (Diazepam) and the six ligands docked in the transmembrane domain (TMD) site at the β-γ interface (site (d)).

	S-score (Kcal/mol)	**Atom of Ligand**	**Involved receptor atoms**	**Involved receptor residues**	**Type of interactions**	**Distances (A°)**	**Bond Energy (kcal/mol)**
			Co-crystallized Ligand				
rmsd = 0.314423084	-6.44459772	CL 35	OD1	ASP 297 (E)	H-donor	3.48	-0.7
			Complexe GABA$_A$-ligand				
5*H*-1,2-bdz	-4.54392052	/	/	/	/	/	/
1*H*-1,3-bdz	-4.54438877	7-ring	CB	SER 280 (E)	pi-H	3.86	-1.0
3*H*-1,4-bdz	-4.64406681	/	/	/	/	/	/
3*H*-1,5-bdz	-4.63208199	/	/	/	/	/	/
5*H*-2,3-bdz	-4.47868633	N 3	NE2	GLN 224 (A)	H-acceptor	2.92	-1.3
1*H*-2,4-bdz	-4.5871954	/	/	/	/		/

Based on the atomic ESP over the whole molecular surfaces of 1*H*-1,3-bdz given in Table III.7, N_3 and H_1 would be expected to be the most acceptable sites for positive and negative distant entities, respectively. The binding poses of 1*H*-1,3-bdz mapped in Figure III.5 indicate the formation of an average interaction type H-acceptor between N_3 and one of the two hydrogens of the amine group in Ala D161 of the site (a). While site (b) shows the existence of weak interaction type H-pi between H_{19} and the π-electron cloud of the phenyl ring of Phe C289. This latter can be explained by observing the orientation of ligand in the binding site, where Phe C289 seems to be further from the diazepine ring and closer to the phenyl, so, logically, Phe C289 would be interacting with the most electron-deficient site in the Phenyl ring, which is H_{19}. Sites (c) and (d) also indicate the existence of weak interactions type H-pi between the π-electron cloud of the diazepine ring and hydrogen of Arg A269 and Ser E280. This interaction does not observe in the rest studied ligands. This may be due to the strong nucleophilicity of this structure (NS occupies 49.2% of the total surface, N=3.323 eV).

Similarly, the 5*H*-2,3-bdz binding poses mapped in Figure III.6 indicate the formation of two strong interactions type H-acceptor between the most electron rich-site (N_3) and the two residues: Thr D207 of the site (a) and Gln A224 of the site (d). Whereas in the site (b) (see also Figure III.7), the phenyl rings of 5*H*-2,3-bdz, 3*H*-1,4-bdz, and 1*H*-2,4-bdz established an arene-arene interaction with the phenyl ring of Phe C289.

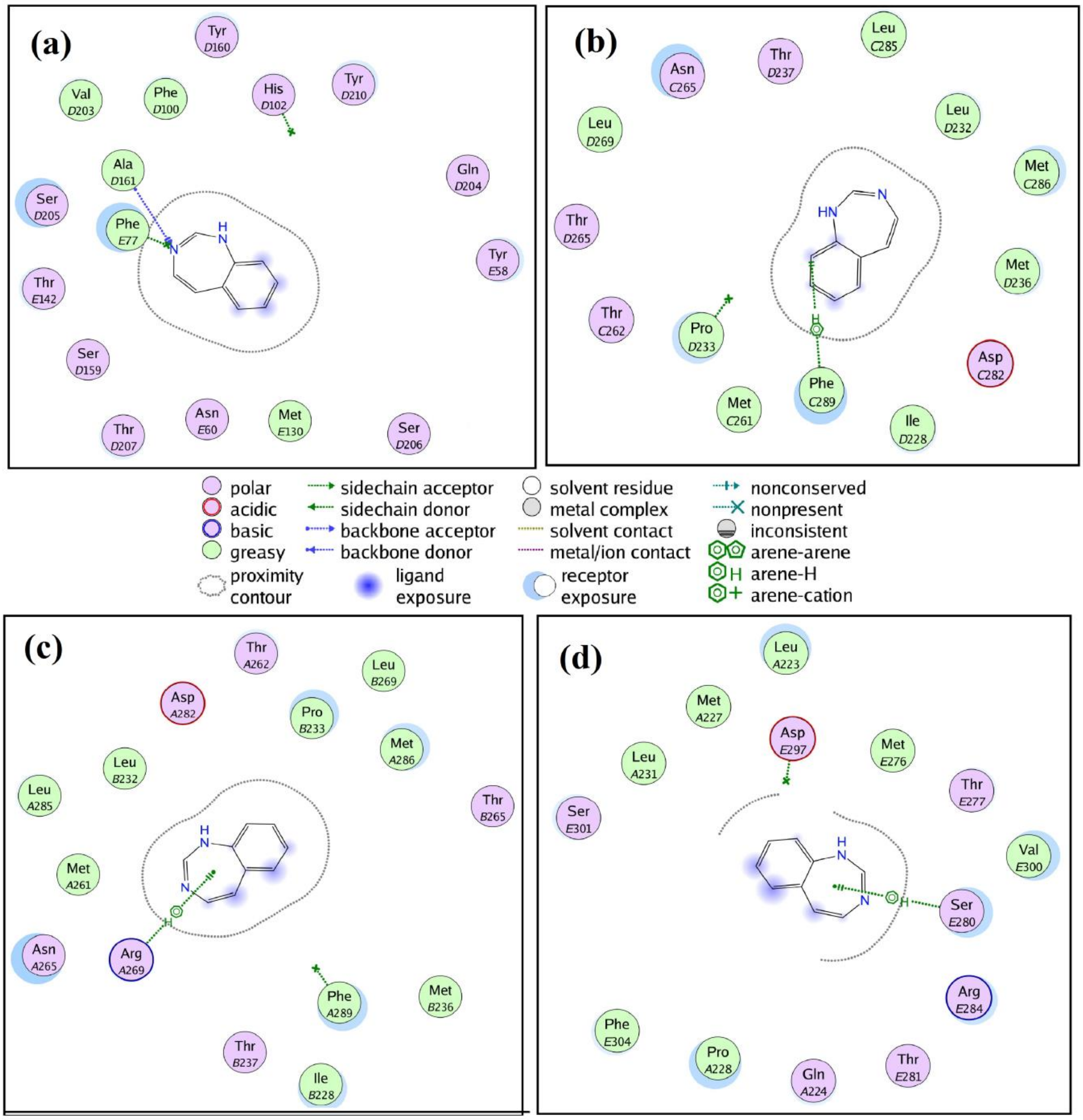


Figure III.5. The best binding poses for 1*H*-1,3-bdz ligand in the classical site at the ECD α-γ interface (a), the two TMD sites at the α-βinterfaces (b) and (c), and the TMD site at β-γ interface (d).

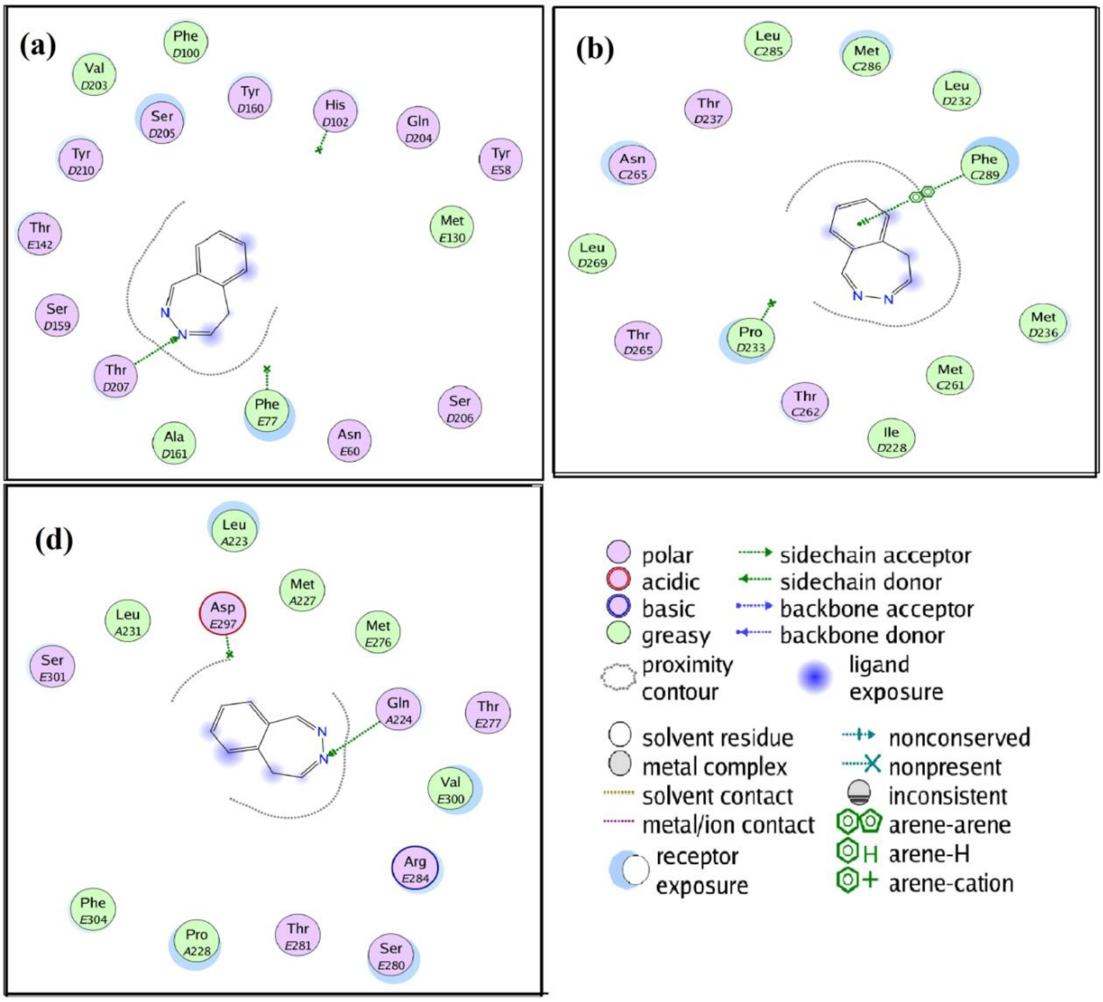


Figure III.6. The best binding poses for 5*H*-2,3-bdz ligand in the classical site at the ECD α-γ interface (a), the TMD site at the α-β interfaces (b), and the TMD site at β-γ interface (d).

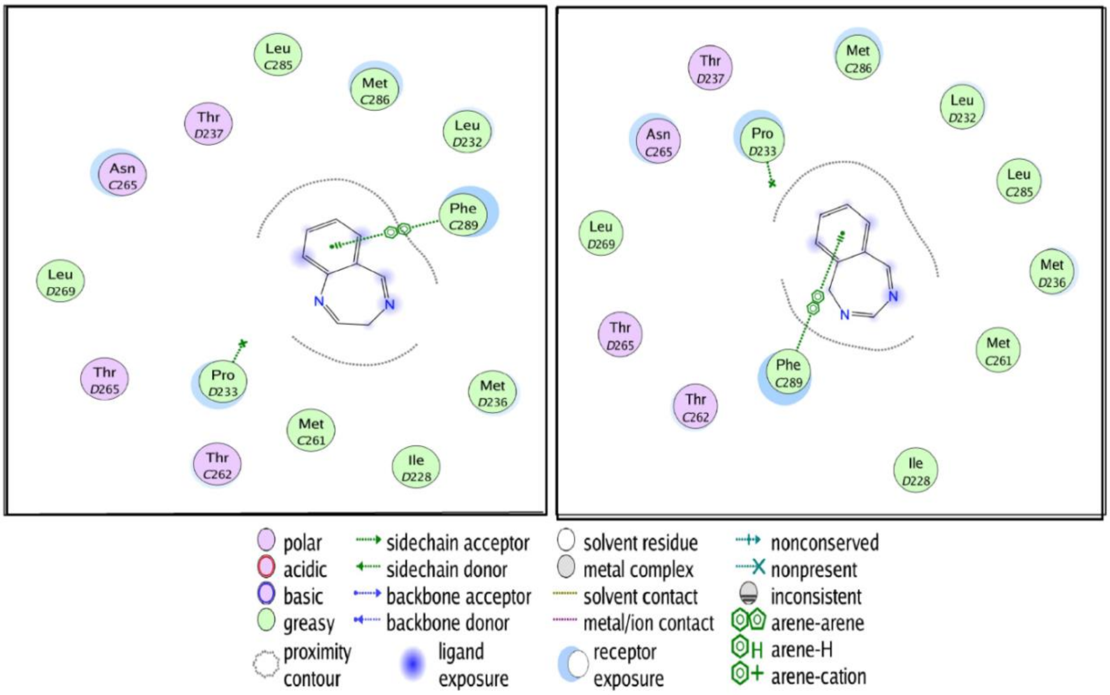


Figure III.7. The best binding poses for 3*H*-1.4-bdz and 1*H*-2,4-bdz ligands in the TMD site at the α-β interfaces (b).

4 Conclusion

In this chapter, a combined approach based on conceptual-DFT theory and molecular docking simulations were performed to investigate the chemical reactivity of six Benzodiazepine analogs. Chemical reactivity descriptors derived from the conceptual-DFT were determined and discussed to explain the global and local reactivity of the six studied analogs. Also, long-range interactions were studied using the quantitative analyses of the molecular electrostatic potential on van der Waals surface to identify the nucleophilic and electrophilic sites. A statistical analysis determined the robustness of each population method toward the used basis sets, the density-based method Hirshfeld presents the best performance and therefore exhibits high robustness against the basis set quality. Both global and local conceptual DFT-based descriptors were performed to provide information about global and local reactivity. The dual descriptor led to a fairly good prediction of favorable electrophilic and nucleophilic sites, making it possible to easily direct the reagents to the desired positions during the creation of new derivatives as mentioned for the 3*H*-1,4-bdz structure. The quantitative ESP analysis exhibited that the ability of each atomic site to interact electrostatically with the long-distance nucleophilic and electrophilic entities differed from structure to another. As well, the molecular docking results indicated that the six studied structures have converging binding affinities and despite this, each ligand has a different binding mode with the target sites. The interpretation of the results was given based on the quantitative analyses of ESP on van der Waals surface.

Finally, this work is an expansion to theoretical researches related to the study of heterocyclic compounds of pharmaceutical interest using DFT theory in combination with molecular docking simulation. The results obtained in this study allow deepening the understanding, from a theoretical point of view, of the chemical reactivity of six benzodiazepine basic structures. This work will certainly stimulate similar studies on a larger range of organic reactions.

Bibliography

[1] D.J. Sanger, J. Benavides, G. Perrault, E. Morel, C. Cohen, D. Joly, B. Zivkovic, Recent developments in the behavioral pharmacology of benzodiazepine (omega) receptors: evidence for the functional significance of receptor subtypes., Neurosci. Biobehav. Rev. 18 (1994) 355–72. https://doi.org/0149-7634(94)90049-3 [pii].

[2] J.R. Atack, The benzodiazepine binding site of GABA(A) receptors as a target for the development of novel anxiolytics., Expert Opin. Investig. Drugs. 14 (2005) 601–18. https://doi.org/10.1517/13543784.14.5.601.

[3] G.A. Archer, L.H. Sternbach, The chemistry of benzodiazepines, Chem. Rev. 68 (1968) 747–784. https://doi.org/10.1021/cr60256a004.

[4] R.I. Fryer, Bicyclic Diazepines: Diazepines with an Additional Ring, 2009.

[5] P. Aastha, K. Navneet, A. Anshu, S. Pratima, K. Dharma, 1 , 5 Benzodiazepines : Overview of Properties and Synthetic Aspects, Res. J. Chem. Sci. 3 (2013) 90–103. https://doi.org/10.1002/CHIN.201422238.

[6] A.G. Meyer, A.C. Bissemberx, C.J.T. Hyland, Charlotte C. Williams, M. Szabo, S.-A.G. Abel, M.J. Bird, I.K. Hyland, H. Pham, Seven-Membered Rings, in: Prog. Heterocycl. Chem., Elsevier Ltd, 2018: pp. 493–550. https://doi.org/10.1016/B978-0-08-102788-2.00016-7.

[7] N. Sukumar, A Matter of Density: Exploring the Electron Density Concept in the Chemical, Biological, and Materials Sciences, John Wiley & Sons, 2012. https://doi.org/10.1002/9781118431740.

[8] R.G. Pearson, Recent advances in the concept of hard and soft acids and bases, J. Chem. Educ. 64 (1987) 561. https://doi.org/10.1021/ed064p561.

[9] R.G. Parr, R.G. Pearson, Absolute Hardness: Companion Parameter to Absolute Electronegativity, J. Am. Chem. Soc. 105 (1983) 7512–7516. https://doi.org/10.1021/ja00364a005.

[10] L. Lebart, A. Morineau, M. Piron, Statistique exploratoire multidimensionnelle, Paris: Dunod, 1995.

[11] R.S. Mulliken, Electronic population analysis on LCAO-MO molecular wave functions. I, J. Chem. Phys. 23 (1955) 1833–1840. https://doi.org/10.1063/1.1740588.

[12] A.E. Reed, R.B. Weinstock, F. Weinhold, Natural population analysis, J. Chem. Phys. 83 (1985) 735–746. https://doi.org/10.1063/1.449486.

[13] L.E. Chirlian, M.M. Francl, Atomic charges derived from electrostatic potentials: A detailed study, J. Comput. Chem. 8 (1987) 894–905. https://doi.org/10.1002/jcc.540080616.

[14] F.L. Hirshfeld, Bonded-atom fragments for describing molecular charge densities, Theor. Chim. Acta. 44 (1977) 129–138. https://doi.org/10.1007/BF00549096.

[15] M. J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, G.A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B.G. Janesko, R. Gomperts, B. Mennucci, H.P. Hratchian, J. V. Ortiz, A.F. Izmaylov, J.L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V.G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J.A. Montgomery, Jr., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, J.M. Millam, M. Klene, C. Adamo, R. Ca, D.J. Fox, Gaussian 09W, (2010).

[16] T. Lu, F. Chen, Quantitative analysis of molecular surface based on improved Marching Tetrahedra algorithm, J. Mol. Graph. Model. 38 (2012) 314–323. https://doi.org/10.1016/j.jmgm.2012.07.004.

[17] W. Humphrey, A. Dalke, K. Schulten, visual molecular dynamics.pdf, J. Mol. Graph. 14 (1996) 33–38.

[18] H.P. Lüthi, J.H. Ammeter, J. Almlöf, K. Faegri, How well does the Hartree-Fock model predict equilibrium geometries of transition metal complexes? Large-scale LCAO-SCF studies on ferrocene and decamethylferrocene, J. Chem. Phys. 77 (1982) 2002–2009. https://doi.org/10.1063/1.444053.

[19] J.B. Collins, A. Streitwieser, Integrated spatial electron populations in molecules: Application to simple molecules, J. Comput. Chem. 1 (1980) 81–87. https://doi.org/10.1002/jcc.540010111.

[20] C.A. Reynolds, J.W. Essex, W.G. Richards, Atomic Charges for Variable Molecular Conformations, J. Am. Chem. Soc. 114 (1992) 9075–9079. https://doi.org/10.1021/ja00049a045.

[21] R.K. Roy, S. Pal, K. Hirao, On non-negativity of Fukui function indices, J. Chem. Phys. 110 (1999) 8236–8245. https://doi.org/10.1063/1.478792.

[22] R.K. Roy, K. Hirao, S. Pal, On non-negativity of Fukui function indices . II, J. Chem. Phys. 113 (2000) 1372–1379. https://doi.org/10.1063/1.481927.

[23] R.K. Roy, K. Hirao, S. Krishnamurty, S. Pal, Mulliken population analysis based evaluation of condensed Fukui function indices using fractional molecular charge, J. Chem. Phys. 115 (2001) 2901–2907. https://doi.org/10.1063/1.1386699.

[24] C. Cárdenas, N. Rabi, P.W. Ayers, C. Morell, P. Jaramillo, P. Fuentealba, Chemical reactivity descriptors for ambiphilic reagents: Dual descriptor, local hypersoftness, and electrostatic potential, J. Phys. Chem. A. 113 (2009) 8660–8667. https://doi.org/10.1021/jp902792n.

[25] P. Jaramillo, L.R. Domingo, E. Chamorro, P. Pérez, A further exploration of a nucleophilicity index based on the gas-phase ionization potentials, J. Mol. Struct. THEOCHEM. 865 (2008) 68–72. https://doi.org/10.1016/j.theochem.2008.06.022.

[26] L.R. Domingo, P. Pérez, Global and local reactivity indices for electrophilic/nucleophilic free radicals, Org. Biomol. Chem. 11 (2013) 4350–4358. https://doi.org/10.1039/c3ob40337h.

[27] A. Micheli, A. Sperduti, A. Starita, A.M. Bianucci, Analysis of the Internal Representations Developed by Neural Networks for Structures Applied to Quantitative Structure-Activity Relationship Studies of Benzodiazepines, J. Chem. Inf. Comput. Sci. 41 (2001) 202–218. https://doi.org/10.1021/ci9903399.

[28] D.J. Maddalena, G.A.R. Johnston, Prediction of Receptor Properties and Binding Affinity of Ligands to Benzodiazepine/GABAA Receptors Using Artificial Neural Networks, J. Med. Chem. 38 (1995) 715–724. https://doi.org/10.1021/jm00004a017.

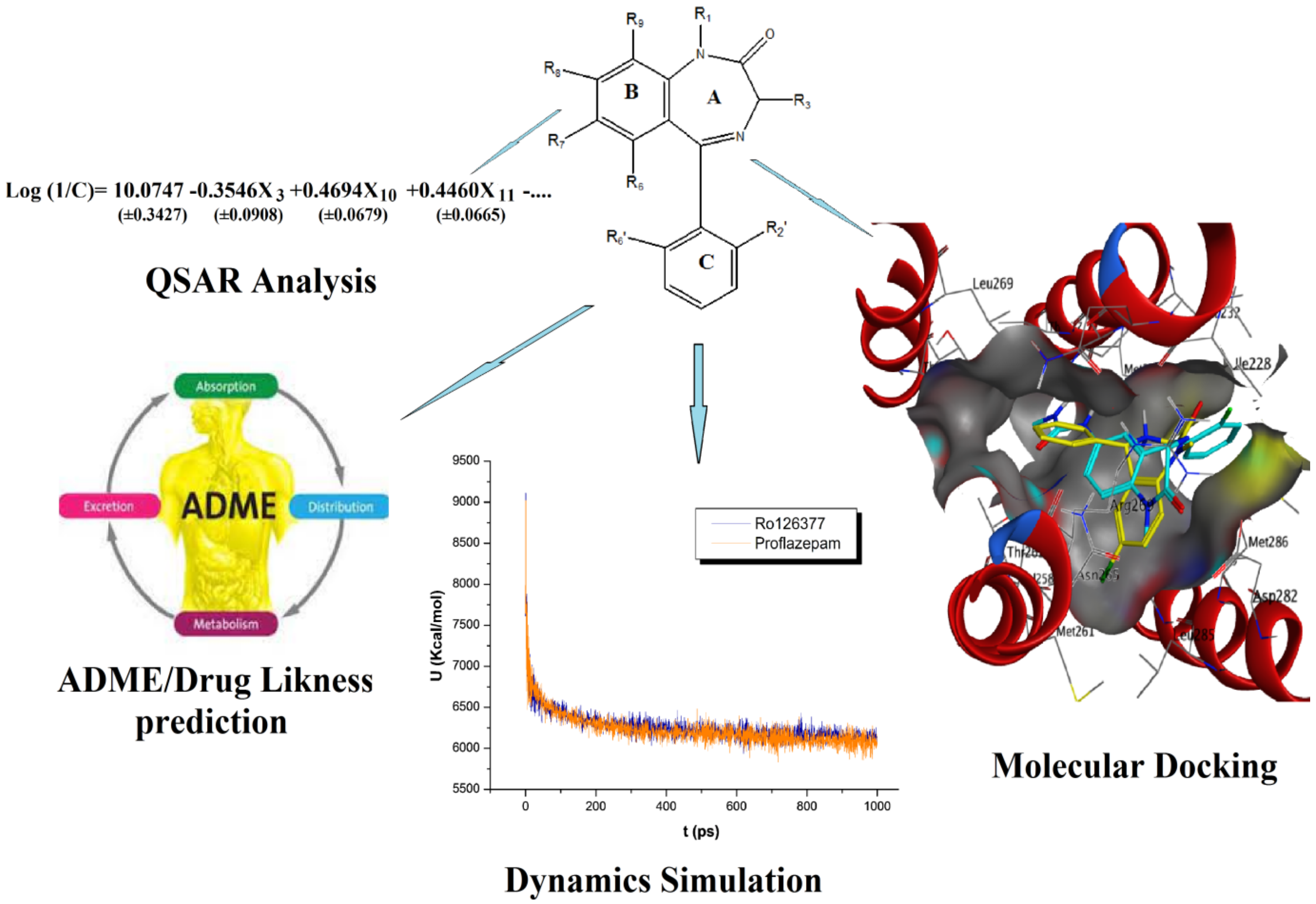
[29] J.I. Martínez, J.L. Moncada, J.M. Larenas, The dual descriptor to measure local reactivity on Buckminster fullerenes: An analysis within the framework of conceptual DFT, J. Mol. Model. 16 (2010) 1825–1832. https://doi.org/10.1007/s00894-009-0638-3.

[30] J.S. Murray, T. Brinck, P. Lane, K. Paulsen, P. Politzer, Statistically-based interaction indices derived from molecular surface electrostatic potentials: a general interaction properties function (GIPF), J. Mol. Struct. THEOCHEM. 307 (1994) 55–64. https://doi.org/10.1016/0166-1280(94)80117-7.

Chapter IV

Investigation of $[^{3}H]$diazepam derivatives as allosteric modulators of $GABA_A$ receptor $\alpha_1\beta_2\gamma_2$ subtypes: combination of molecular docking, molecular dynamics simulation, pharmacokinetics/drug-likeness prediction, and QSAR analysis.

Graphical abstract



1 Introduction

The strategy of combining molecular docking, molecular dynamics simulation, and QSAR analysis has emerged as a practical tool in the process of drug development through computational techniques. Its main advantage lies in improving the success rate of drug screening in less time and at a lower cost [1].

Molecular docking simulations are designed to determine the best ligand/target binding mode that generates the biological response. In practice, it allows the screening of large libraries of compounds by implementing fast and inexpensive docking algorithms. Mostly, after molecular docking simulations, the best-docked complexes are subject to stability investigation through molecular dynamics simulation. Monitoring the dynamic profile of complexes over a certain time range provides the advantage of detecting various internal motions and conformational changes that occur in the binding site. Hence, validate the docking protocols. Otherwise, MD simulations can be generated before performing molecular docking for several objectives such as optimizing the target structure and ensuring its flexibility, quantifying the ligands/target free binding energies,…etc. As well as, during the docking process, to accurately detect the binding locus and properly dock ligands [2].

Quantitative structure-activity relationships (QSARs) are mainly devoted to predict the activities of new chemical entities from knowledge of their chemical structures. QSAR models quantitatively correlate the physicochemical and biological properties of compounds with their biological responses [3]. The success of any QSAR model depends on many factors, including the selection of practical statistical techniques for model development and validation strategies. The quality of models is evaluated through an internal validation process usually based on the use of the cross-validation method (CV). Whereas, the predictive power can be estimated using independent test data that was not involved in model generation [4]. In the literature, considerable researchers have sought to study the BDZ by performing QSAR analysis. Against this background, fifty-seven compounds from the dataset of interest in our study were examined by *D.J. Maddalena et al* [5] using the back-propagation artificial neural network method (ANN) and multilinear regression analysis (MLR). The two-layer ANN model gave an excellent correlation between the binding affinities of the 38 compounds of the training set (R_t= 0.941) and the 10 selected input variables: (π_7, F_7, MR_1, $MR_{2'}$, R_1, F_2, MR_6, μ_1, δ_{p8}, and δ_{m3}). Where, F, R, μ, δ_p, and δ_m denote polar constant, resonance constant, dipole moment, hammett-para constant, and hammett-meta constant, respectively. The predictive power was tested with an external set of 19 compounds and an optimal cross-validation coefficient (R_{cv}=0.910) was found. Later, *S. S.*

So et al [6] re-examined the data provided by *D.J. Maddalena et al* with an improved version of the genetic neural network (GNN). The top-ranking variables selected by GNN shared only four variables with the ANN selection (π_7, F_7, MR_1, and $MR_{2'}$). Therefore, an accurate comparison and a further discussion were made between the results of the two statistical methods, as well as, the three highly predictive models (T6-2 # 1-3) were combined with the optimal functional groups proposed by *D.J. Maddalena* to be placed in positions 1, 7, and 2' and used to design 20 new BDZ derivatives with their predicted activities.

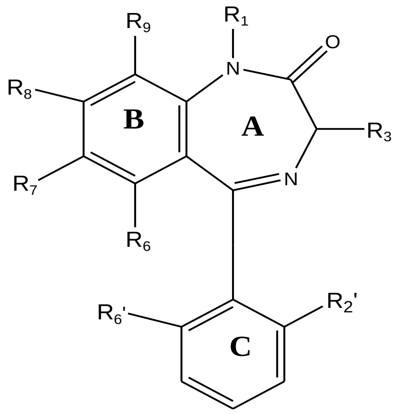
This chapter aims to highlight the binding mechanism by which a data set of classical benzodiazepines allosterically modulates $GABA_A$ receptor $\alpha_1\beta_2\gamma_2$ subtypes, from inducing neuronal inhibition at lower doses to the anesthetic effect at higher doses. In addition to the well-known ECD binding interface, our study was further expanded to include the three TMD interfaces that were recently identified by *J.J. Kim et al* [7]. In the first step, a data set of [^{3}H]diazepam derivatives was subjected to fast screening through the molecular docking approach. Subsequently, molecular dynamics simulation and pharmacokinetics/drug-likeness evaluations were performed to refine the best-docked complexes. Finally, an improved version of PLS regression was implemented to quantify structural features that contribute to improving the response of $GABA_A/\alpha_1\beta_2\gamma_2$ receptor to benzodiazepine drugs. Throughout this paper, the results have been interpreted in light of the combination of the cited approaches.

2 Material and methods

2.1 Biological data

The results in vitro for the 50% inhibition of the binding of [^{3}H]diazepam to homogenates of rat brain cell membranes by BDZs expressed as log (1/C) reported earlier in the review of *D.H. Litinat et al* [8] (Table IV.1) were investigated to perform a molecular docking simulation and to predict QSAR model using PLS analysis. According to *A. Micheli et al* [9], the good structural diversity allows this dataset to be optimal for undergoing QSAR analysis.

Table IV.1. The classical BDZ data set under study [8].



N°	Nam	R_1	R_3/ R_5/R_6	R_7	R_8/ R_9	$R_{2'}$	$R_{6'}$	Log (1/C)$_{obs}$
1	**Ro05-4318/ Ro05-3418**	CH_3	R_3,R_5,R_6=H	NH_2	R_8,R_9=H	H	H	**6.34**
2	**Ro05-3072**	H	R_3,R_5,R_6=H	NH_2	R_8,R_9=H	H	H	**6.41**
3	**Ro05-4528**	CH_3	R_3,R_5,R_6=H	CN	R_8,R_9=H	H	H	**6.42**
4	**Ro05-2921**	H	R_3,R_5,R_6=H	H	R_8,R_9=H	H	H	**6.45**
5	**Ro20-7736**	CH_3	R_3,R_5,R_6=H	NHOH	R_8,R_9=H	F	H	**7.02**
6	**Ro05-4619**	H	R_3,R_5,R_6=H	NH_2	R_8,R_9=H	Cl	H	**7.12**
7	**Ro20-5397**	H	R_3,R_5,R_6=H	CHO	R_8,R_9=H	H	H	**7.37**
8	**Ro05-3061**	H	R_3,R_5,R_6=H	F	R_8,R_9=H	H	H	**7.40**
9	**Ro20-2533**	H	R_3,R_5,R_6=H	C_2H_5	R_8,R_9=H	H	H	**7.44**
10	**Ro20-2541**	CH_3	R_3,R_5,R_6=H	CN	R_8,R_9=H	F	H	**7.52**
11	**Ro20-5747**	H	R_3,R_5,R_6=H	$CHCH_2$	R_8,R_9=H	H	H	**7.62**
12	**Ro05-4336**	H	R_3,R_5,R_6=H	H	R_8,R_9=H	F	H	**7.68**
13	**Ro20-3053**	H	R_3,R_5,R_6=H	$COCH_3$	R_8,R_9=H	F	H	**7.74**
14	**Triflunordazepam**	H	R_3,R_5,R_6=H	CF_3	R_8,R_9=H	H	H	**7.89**
15	**Diazepam**	CH_3	R_3,R_5,R_6=H	Cl	R_8,R_9=H	H	H	**8.09**
16	**Ro07-5220**	CH_3	R_3,R_5,R_6=H	Cl	R_8,R_9=H	Cl	Cl	**8.26**
17	**Ro14-3074**	H	R_3,R_5,R_6=H	N_3	R_8,R_9=H	F	H	**8.27**
18	**Flunitrazepam**	CH_3	R_3,R_5,R_6=H	NO_2	R_8,R_9=H	F	H	**8.42**
19	**Ro05-3590**	H	R_3,R_5,R_6=H	NO_2	R_8,R_9=H	CF_3	H	**8.45**
20	**Norflurazepam**	H	R_3,R_5,R_6=H	Cl	R_8,R_9=H	F	H	**8.7**
21	**Delorazepam**	H	R_3,R_5,R_6=H	Cl	R_8,R_9=H	Cl	H	**8.74**
22	**Clonazepam**	H	R_3,R_5,R_6=H	NO_2	R_8,R_9=H	Cl	H	**8.74**
23	**Fonazepam**	H	R_3,R_5,R_6=H	NO_2	R_8,R_9=H	F	H	**8.82**
24	**Ro05-6822**	CH_3	R_3,R_5,R_6=H	F	R_8,R_9=H	F	H	**8.29**
25	**Ro05-4865**	CH_3	R_3,R_5,R_6=H	F	R_8,R_9=H	H	H	**7.77**
26	**Ro05-6820**	H	R_3,R_5,R_6=H	F	R_8,R_9=H	F	H	**8.13**
27	**Nordazepam**	H	R_3,R_5,R_6=H	Cl	R_8,R_9=H	H	H	**8.03**
28	**Ro07-3953**	H	R_3,R_5,R_6=H	Cl	R_8,R_9=H	F	F	**8.79**
29	**Difludiazepam**	CH_3	R_3,R_5,R_6=H	Cl	R_8,R_9=H	F	F	**8.39**
30	**Ro07-5193**	H	R_3,R_5,R_6=H	Cl	R_8,R_9=H	Cl	F	**8.52**
31	**Ro22-3294**	H	R_3,R_5,R_6=H	Cl	R_8,R_9=H	Cl	Cl	**8.15**
32	**Nitrazepam**	H	R_3,R_5,R_6=H	NO_2	R_8,R_9=H	H	H	**7.99**
33	**Methylclonazepam**	CH_3	R_3,R_5,R_6=H	NO_2	R_8,R_9=H	Cl	H	**8.66**
34	**7-Aminoflunitrazepam**	CH_3	R_3,R_5,R_6=H	NH_2	R_8,R_9=H	F	H	**7.19**
35	**Ro12-6377**	CH_3	R_3,R_5,R_6=H	$NHCONHCH_3$	R_8,R_9=H	F	H	**6.34**
36	**Halazepam**	CH_2CF_3	R_3,R_5,R_6=H	Cl	R_8,R_9=H	H	H	**7.04**

Table IV.1. Continued

37	**Pinazepam**	CH_2CCH	$R_3,R_5,R_6=H$	Cl	$R_8,R_9=H$	H	H	**7.03**
38	**Prazepam**	$CH_2C_3H_5$	$R_3,R_5,R_6=H$	Cl	$R_8,R_9=H$	H	H	**6.96**
39	**Motrazepam**	CH_2OCH_3	$R_3,R_5,R_6=H$	NO_2	$R_8,R_9=H$	H	H	**6.37**
40	**Ro20-1310**	$C(CH_3)_3$	$R_3,R_5,R_6=H$	Cl	$R_8,R_9=H$	H	H	**6.21**
41	**Ro07-2750**	CH_2CH_2OH	$R_3,R_5,R_6=H$	Cl	$R_8,R_9=H$	F	H	**7.61**
42	**Ro08-9013**	$(CH_2)_2OCH_2CONH_2$	$R_3,R_5,R_6=H$	Cl	$R_8,R_9=H$	F	H	**7.37**
43	**Proflazepam**	$CH_2CHOHCH_2OH$	$R_3,R_5,R_6=H$	Cl	$R_8,R_9=H$	F	H	**6.85**
44	**Ro22-4683**	$C(CH_3)_3$	$R_3,R_5,R_6=H$	NO_2	$R_8,R_9=H$	Cl	H	**6.52**
45	**Ro11-4878**	H	$R_3=(s)CH_3$ $R_5,R_6=H$	Cl	$R_8,R_9=H$	F	H	**8.46**
46	**Meclonazepam**	H	$R_3=(s)CH_3$ $R_5,R_6=H$	NO_2	$R_8,R_9=H$	Cl	H	**8.92**
47	**Ro11-6896**	CH_3	$R_3=(s)CH_3$ $R_5,R_6=H$	NO_2	$R_8,R_9=H$	F	H	**8.15**
48	**L48**	CH_3	$R_3=(rac)CH_3$ $R_5,R_6=H$	Cl	$R_8,R_9=H$	H	H	**7.31**
49	**Temazepam**	CH_3	$R_3=(rac)OH$ $R_5,R_6=H$	Cl	$R_8,R_9=H$	H	H	**7.79**
50	**L50**	CH_3	$R_3=(rac)Cl$ $R_5,R_6=H$	Cl	$R_8,R_9=H$	F	H	**8.27**
51	**L51**	H	$R_3,R_5=H$ $R_6=CH_3$	CH_3	$R_8,R_9=H$	H	H	**6.77**
52	**Ro07-4419**	H	$R_3,R_5,R_6=H$	H	$R_8,R_9=H$	F	F	**7.72**
53	**Ro05-4520**	CH_3	$R_3,R_5,R_6=H$	H	$R_8,R_9=H$	F	H	**7.85**
54	**Ro05-4608**	CH_3	$R_3,R_5,R_6=H$	H	$R_8,R_9=H$	Cl	H	**8.42**
55	**Ro05-3546**	H	$R_3,R_5=H$ $R_6=Cl$	H	$R_8,R_9=H$	H	H	**6.49**
56	**Ro13-0699**	CH_3	$R_3,R_5=H$ $R_6=Cl$	H	$R_8,R_9=H$	F	H	**6.82**
57	**Ro07-6198**	H	$R_3,R_5,R_6=H$	H	$R_8=Cl$ $R_9=H$	F	F	**7.55**
58	**Ro20-8895**	H	$R_3,R_5,R_6=H$	H	$R_8=CH_3$ $R_9=H$	F	H	**7.72**
59	**Ro13-0593**	CH_3	$R_3,R_5,R_6=H$	H	$R_8=H$ $R_9=Cl$	F	H	**7.14**
60	**L60**	CH_3	$R_3,R_5=H$ $R_6=Cl$	H	$R_8=Cl$ $R_9=H$	F	H	**6.52**
61	**Ro22-6762**	CH_3	$R_3,R_5,R_6=H$	Cl	$R_8=Cl$ $R_9=H$	H	H	**7.40**
62	**Ro20-8065**	H	$R_3,R_5,R_6=H$	Cl	$R_8=Cl$ $R_9=H$	F	H	**8.44**
63	**Ro20-8552**	H	$R_3,R_5,R_6=H$	CH_3	$R_8=Cl$ $R_9=H$	F	H	**7.85**
64	**L64**	H	$R_3,R_5,R_6=H$	Cl	$R_8=H$ $R_9=Cl$	H	H	**7.43**
65	**L65**	H	$R_3,R_5,R_6=H$	Cl	$R_8=H$ $R_9=CH_3$	H	H	**7.28**

2.2 Molecular descriptors generation

First, a gradient norm limit of 0.1 kcal/(Å mol) was chosen for the pre-optimization of sixty-five BDZ derivatives, using the molecular mechanics force field (MM^+) method included in HyperChem package version 8.08 [10]. Then, the molecular geometries were optimized at the DFT/Ub3lyp/6-311^{++}G(d,p) level of theory using the Gaussian 09W software [11]. For all stationary points, there is no imaginary frequency at the optimized molecular geometries ensuring that the optimized structures are at the minimum on the potential energy surface. The atomic charges q_{N1}, q_{C3}, q_{N4}, q_{C6}, q_{C7}, q_{C8}, q_{C9}, $q_{C2'}$, $q_{C6'}$ and the dipole moment (DM) have been assessed using the ChelpG electronic population scheme [12]. The number of hydrogen-donors (HD), number of hydrogen-acceptors (HA), Molecular lipophilicity Log (P), and molar refractivity (M_R) have been computed using MarvinSketch version 20.21 [13]. The flexible torsions F_T have been computed using Molegro Virtual Docker version 5.5 [14]. Finally, the hydrophobic constants in positions 7 (π_{C7}) and 2' ($\pi_{C2'}$) have been extracted from the literature [13].

2.3 QSAR analysis

2.3.1 PLS regression

The chemical, physical, topological and quantum properties are necessarily correlated for a given molecule. This is a reflection of the innate properties of the system and additional data collected in the same way will show the same collinearity [14]. Indeed, PLS regression is a useful method for multivariate data containing correlated molecular descriptors. This method based on dimension reduction technique, builds orthogonal components, often called factors or latent variables, as linear combinations of the original predictor variables [15]. PLS constructs these components while considering the observed response values, leading to a parsimonious model with reliable predictive power [16].

In this work, the number of components used in PLS is chosen by 5-fold cross-validation method [17]. The dataset is randomly divided into training dataset (80%) and testing dataset (20%). Training sets were used for model development and test sets for model external validation. Before conducting PLS analysis, each response variable is scaled to unit variance by dividing it by its standard deviation, and the molecular descriptors are centered by subtracting the average value and scaled to unit variance. First, variable selection by Stepwise regression method is used to identify the best subset of molecular descriptors. This is a combination of backward and forward selection [15]. The objective is to use the minimum number of descriptors to develop a good predictive model. Thus, we must select the good subsets of descriptors. However, it should be noted that the subset of molecular predictors that do the best at meeting well-defined objective criteria can be highly variables depending on precisely which observations are included in the training set. In addition, the best training model does not necessarily guarantee a better quality of prediction. This depends on the training and test sets obtained from the original dataset. For this reason, we conducted a statistical simulation for which 10 000 splits were performed resulting in 10 000 training and test sets. For each simulation, regression diagnostics for detecting possible outliers was carried out by computing leverage values (h_{ii}) for identifying outlying x-variables and studentized deleted residuals (r_i^*) for identifying outlying y-variables. Note that once outliers have been detected the model is regenerated excluding the outlying observations from the dataset. Afterwise, the best model is selected on each training set resulting in 10 000 best training models following the Bayesian Information Criterion (BIC) [18]. This criterion is chosen because it penalizes larger models more heavily and will tend to select a smaller subset of descriptors in comparison to other criteria [19]. The best choice of descriptors will balance fit with model size. Subsequently, among these 10 000 best models, we sought to select the best molecular descriptors according to the highest probability of their occurrence.

To verify a model's predictive ability, the developed QSAR model is quantified using the coefficient of determination (R^2) [20], the adjusted coefficient of determination (R^2_{adj}) [21], and the fisher-statistics (F) [22]. This latter is computed to judge the overall significance of the regression model. The external predictive ability of the developed QSAR model is determined by computing the leave-one-out cross-validation coefficient (Q^2_{loo}) [22] and the predictive squared correlation coefficient (Q^2_{F3}) [23–25]. The external validation ensures the predictability of the developed QSAR model for the prediction of untested molecules [4].

2.4 Molecular docking protocol

The electron microscopy structure of the human $GABA_A$ receptor $\alpha_1\beta_2\gamma_2$ subtypes in complex with GABA plus the DZP structures (PDB ID:6X3X, Resolution= 2.92 Å) was downloaded from RCSB Database (http://www.rcsb.org). The downloaded PDB file contains nine chains: five chains denote the subunits α_1 (B and D), β_2 (A and C), and γ_2 (E), and four chains denote the Fab-chains (named from I to K).

MOE 2014.0901 software package (Molecular Operating Environment (MOE), 2014) was used to prepare the four benzodiazepine binding sites: the classical site at the ECD D^+/E^- chains interface and the three TMD sites at the A^+/B^-, C^+/D^-, and E^+/A^- chains interfaces. For each binding site, all co-crystallized ligands and non-essential subunits were removed from the $\alpha_1\beta_2\gamma_2$-DZP complex. Then, after structure correction, protonation at neutral medium (PH=7), and cavity detection, the native co-crystallized DZP structure was re-docked into the selected binding site pocket. The method was validated by giving the best binding pose which has a low RMSD value (root-mean-square deviation).

The BDZ structures previously optimized using the DFT method were converted into database input files and docked one by one into the four DZP binding pockets using the semi-flexible docking process of MOE 2014.0901 package [28]. During the process, the conformation of the receptor was fixed, while ligands remained flexible. Here, the best binding poses were selected according to the lowest energy score values, registered in the PDB file, and visualized using BIOVIA Discovery Studio visualizer v20.1.0.19295 package [27].

2.5 Molecular dynamics protocol

The best-docked ligand/$\alpha_1\beta_2\gamma_2$ complexes were subjected to stability tests using molecular dynamics simulation (MD). MD simulation was implemented using the "compute" option included in the MOE 2014.0901 software [26]. First, the selected complexes were prepared by deleting the DZP co-crystallized structure, fixing hydrogens, and fixing charges. Afterwise, the parameters of the "dynamics" tool were adjusted to execute the combination Nosé-Poincaré-Andersen (NPA) algorithm/Merck molecular force field (MMFF94x) [28,29], with enabling bonding, van der Waals, electrostatics, and restraints. The protocols were settled for an equilibrium period of 100 ps followed by a production period of 900 ps, at a constant temperature of 310 K. Finally, the variations in potential energies U (Kcal/mol) as a function of time t (ps) are retained and plotted using Origin 6.0 software [32].

2.6 Pharmacokinetics/drug-likeness prediction

In silico estimation of pharmacokinetic properties and prediction of drug-likeness were carried out by using the free web tool SwissADME [33]. Our study is based on the prediction of the following pharmacokinetic parameters: gastrointestinal absorption (GI), P-glycoprotein (P-gp) substrate, blood brain barrier (BBB) penetration, and cytochrome enzyme (CYP) inhibition. Indeed, out of 57 human CYP450 enzymes the CYP1A2, CYP2C9, CYP2C19, CYP2D6, CYP3A4, and CYP2E1 metabolize 90 % of drugs [30]. In addition, the drug-likeness prediction is based on several rules such as: Lipinski, Ghose, Veber, Egan, and Muegge.

3 Results and discussion

3.1 Molecular docking simulation

Orientations, interactions, and binding affinities of 65 positive allosteric modulators of GABA$_A$Rs (BDZ dataset, Table IV.1) were investigated in four distinct BDZ binding sites: the classical site at the ECD α_1^+/γ_2^- interface, the TMD sites at the two $\beta_2^+(A)/\alpha_1^-(B)$ and $\beta_2^+(C)/\alpha_1^-(D)$ interfaces and the TMD site at the γ_2^+/β_2^- interface (Figure 11, chapter I). Residues involved in each active pocket were detected using the site finder wizard implemented in the MOE 2014.0901 [26] (Table IV.2).

Table IV.2. Residues involved in each active pocket.

Binding site		Residues involved in active pockets
ECD	**α_1^+/γ_2^- interface (Classical Binding Site)**	**1:** (PHE100 PHE101 HIS102 ASN103 GLU138 PRO140 PRO154 LYS156 SER159 TYR160 ALA161 VAL203 GLN204 SER205 SER206 THR207 TYR210)
		2: (ASP56 MET57 TYR58 ASN60 SER61 ASP75 PHE77 ALA79 MET130 THR142 ARG144 SER186 GLU189 ASP192 SER195)
TMD	**$\beta_2^+(A)/\alpha_1^-(B)$ interface**	**1:** (ILE255 VAL258 LEU259 MET261 THR262 ASN265 THR266 ARG269 GLU270 ASP282 LEU285 MET286 PHE289 VAL290)
		2: (ILE228 GLN229 LEU232 PRO233 MET236 THR237 LEU240 PHE258 THR261 THR262 THR265 LEU269 SER272)
	$\beta_2^+(C)/\alpha_1^-(D)$ interface	**1:** (MET261 THR262 ASN265 THR266 ARG269 ASP282 LEU285 MET286 PHE289 VAL290)
		2: (VAL227 ILE228 LEU232 PRO233 MET236 THR237 THR265 LEU269)
	γ_2^+/β_2^- interface	**1:** (TYR220 PHE221 LEU223 GLN224 THR225 MET227 PRO228 LEU231 ILE232 THR263 ILE264 HIS267 LEU268 THR271 LEU272)
		2: (MET276 THR277 SER280 THR281 ALA283 ARG284 LYS285 LYS289 MET296 ASP297 VAL300 SER301 PHE304 ILE305)

During the docking process, the conformation of residues remains unchanged, while ligands are altered. The best binding modes of re-docked DZPs were selected based on the given root-mean-square deviation values (RMSD). Generally, docking protocols that are able to generate the same co-crystallized binding modes with an RMSD value of less than 1.5 or 2 A° (depending on ligand size) or even better, less than 1 A° are considered validated [31,32]. Whereas, the best binding modes of docked ligands were selected according to the given S$_{score}$ values. The binding free energy score (S$_{score}$) is a quantitative estimate of the most stable binding pose between the target macromolecule and ligand. Among the generated poses, the best ones are those with the most negative energy score values.

The energy score values of the 65 ligands docked in the α_1^+/γ_2^-, $\beta_2^+(A)/\alpha_1^-(B)$, $\beta_2^+(C)/\alpha_1^-(D)$, and γ_2^+/β_2^- interfaces are between (-8.013 and -6.046) Kcal/mol, (-7.409 and -

5.566) Kcal/mol, (-7.455 and -5.463) Kcal/mol, and (-7.546 and -5.425) Kcal/mol, respectively (Appendix D). As it evident, the studied ligands have rather scattered affinities around the reference values: -7.003 Kcal/mol, -6.159 Kcal/mol, -6.261 Kcal/mol, and -6.347 Kcal/mol, respectively.

Surveying the first five ligands having the lowest energy scores (Table IV.3), the lowest S_{score} value in the three α_1^+/γ_2^-, $\beta_2^+(A)/\alpha_1^-(B)$, and γ_2^+/β_2^- sites was assigned to Ro12-6377. Also, Ro12-6377 exhibits the second-lowest S_{score} value in the TMD $\beta_2^+(C)/\alpha_1^-(D)$ interface. Therefore, out of the 65 studied ligands, Ro12-6377 being the ligand that exhibited the highest predicted affinity towards the four sites, simultaneously. By similar reasoning, the second highest predicted affinity was rated to proflazepam.

Table IV.3. The first five ligands having the highest binding affinity for the four binding interfaces.

ECD α_1^+/γ_2^- interface		**TMD $\beta_2^+(A)/\alpha_1^-(B)$ interface**		**TMD $\beta_2^+(C)/\alpha_1^-(D)$ interface**		**TMD γ_2^+/β_2^- interface**	
Ligand	**S_{score}**	**Ligand**	**S_{score}**	**Ligand**	**S_{score}**	**Ligand**	**S_{score}**
Ro12-6377	-8,013	Ro12-6377	-7,409	Ro08-9013	-7,455	Ro12-6377	-7,546
Meclonazepam	-7,856	Proflazepam	-7,170	Ro12-6377	-7,265	Ro08-9013	-7,070
Methylclonazepam	-7,831	Ro07-2750	-7,152	Proflazepam	-7,077	Pinazepam	-6,990
Proflazepam	-7,807	Pinazepam	-7,146	Pinazepam	-6,848	Proflazepam	-6,972
Ro11-6896	-7,785	Motrazepam	-7,087	Meclonazepam	-6,774	Ro07-2750	-6,918

By comparing the affinities of Ro12-6377 toward the four sites, the α_1^+/γ_2^- site was defined to be the principal target for Ro12-6377. So, when a high dose of Ro12-6377 is administered, it mainly acts to bind at the ECD α_1^+/γ_2^- interface as it is the high-affinity binding site. Then, acts to bind at the TMD γ_2^+/β_2^- and the two β_2^+/α_1^- interfaces as they are the second and third high-affinity binding sites, respectively. On the other hand, proflazepam tends to fill the two β_2^+/α_1^- interfaces before moving to occupy the γ_2^+/β_2^- site. Recently, from a database of 7922 compounds, Proflazepam was selected among the top-100 docked ligands at the binding pocket of SARS-CoV-2 Main Protease crystallized in Holo-form [33].

Likewise, the affinities of the remaining 63 ligands (Appendix D) towards the four binding sites were compared. Unexpectedly, our results are inconsistent with previous findings indicating that the classical site is always the main target of all classical benzodiazepines. As we can see, $\beta_2^+(A)/\alpha_1^-(B)$ and γ_2^+/β_2^- are expected to be, respectively, the main binding sites for (Ro05-4336, Ro07-2750, Ro05-3546) and (L_{51}, Ro13-0699). Moreover, in some cases, both ECD and TMD sites share the same binding affinity towards the bound ligand, as is evident for Ro05-2921, Ro20-5397, Pinazepam, Ro07-4419, Ro05-

3546, and Ro20-8552, which makes it difficult to distinguish, accurately, the main target for the binding. Consequently, after the discovery of the three TMD sites, it became necessary to expand previous findings indicating that the main target of classical benzodiazepines is always located at the ECD α_1^+/γ_2^- interface.

The binding modes of both ligands in each of the four binding interfaces are located at the same level as the co-crystallized DZP (Appendices E and F). However, their docking orientations are markedly not equivalents, with the exception of the binding orientation of proflazepam in the γ_2^+/β_2^- interface where it shares an almost perfect superimposition to the DZP-bound structure (Appendix F (d)). The favorable drug binding pose is determined by the distribution of polar and non-polar regions along the surface of the ligand and its complementary target binding site. Thus, while the non-polar regions create hydrophobic interactions mainly contributing to the binding affinity of the drug toward the biological target, polar Regions create electrostatic points contributing to modulating the drug binding kinetics (specificity and orientation) [34]. In order to estimate all possible interactions, the docking-outputs generated by MOE software were converted into (.pdb) files and visualized with the default parameters of BIOVIA DS visualizer v20.1.0.19295 package [27]. As can be seen, the binding interactions for both ligands with the four target-sites residues exhibit the formation of four types of interactions, most of which are of type hydrogen bonds and hydrophobic interactions. The standard values of distances and energy cutoffs for considering the formation of hydrogen bonds with specific target residues are categorized into three sub-types: strong bonds (2.2-2.5 Å, E: 14-40 kcal/mol), moderate bonds (2.5-3.2 Å, E: 4-15 kcal/mol), and weak bonds (3.2-4.0 Å, E >4 kcal/mol) [35]. Generally, strong bonds in targeted-ligand interactions are undesirables as they mostly tend to have a covalent character. This latter, hinders the process of drug-liberation from its receptor, thus increasing the risk of drug toxicity. Nevertheless, new insights on how to address the pharmacological advantages/potential risks balance of covalent drugs have been emerged and discussed in the literature [36,37]. Referencing to this latter, several standard values of distance cutoffs to consider the formations of hydrophobic interactions have been found. *C. Janiak* [38] suggested that the optimum range is between 3.3 - 3.8 A°. While other researchers have suggested a relatively higher range [39–41].

Owing to the lack of previous experimental and theoretical studies on Ro12-6377 and proflazepam, we will attempt to suggest mechanisms for how they modulate $\alpha_1\beta_2\gamma_2$Rs signaling by binding at the ECD and the TMD interfaces, based on the results obtained from molecular docking and molecular dynamics simulation.

3.1.1 Allosteric modulation of the classical binding site

The bulky structure of Ro12-6377, compared to DZP, enabled the $NHCONHCH_3$ group attached at C_7 and both phenyls (B) and (C) to penetrate deeper into the binding site. Whereas, the diazepine ring exerts its effect by interacting with the residues located in front of the pocket (Figure IV.1 (a), Figure IV.2 (a), Table IV.4). The methyl group attached at N_1 forms two hydrophobic interactions type Alkyl-Alkyl and Pi-Alkyl with α_1Val203 and γ_2Tyr58, respectively. Concurrently, the π-electron clouds of γ_2Tyr58 and γ_2Phe77 were involved in two Pi-Pi stacked interactions with the π-electron cloud of phenyl (B). The linear backbone skeleton of $NHCONHCH_3$ group formed an Alkyl-Alkyl interaction with the side chain of γ_2Ala79, received two moderate H-bonds from α_1Ser206 side chain, and gave two moderate H-bonds to the main carbonyl of γ_2Phe77. The four moderate H-bonds lead to the formation of two intermolecular pentameric rings that contribute to the stability of Ro12-6377 at the binding site. Also, in addition to the strong H-donor bond received from α_1His102, an intramolecular interaction was observed between the fluorine atom bound at $C_{2'}$ and C_2. This interaction could be explained by the strong withdrawing property of the oxygen that made C_2 a more electrophilic center able to receive the nucleophilic attacks. The pendant phenyl (C) is involved deeper into the DZP pocket where is delineated by the hydrophobic side chains of γ_2Phe77, α_1Phe100, α_1His102, α_1Ter160, and α_1Ter210. Here, a hydrophobic interaction type Pi-Pi T-shaped was observed between the π-electron cloud of ring (C) and the surrounded side chain of α_1Phe100.

Table IV.4. Detailed binding interactions resulting from the molecular docking of co-crystallized DZP, Ro12-6377, and proflazepam (PLZ) at the classical site.

ECD α_1^+/γ_2^- interface								
Ligand Name	**Posit** [a]	**category**	**Type of interactions**	**From**	**From chemistry**	**To**	**To chemistry**	**Dist** [b] **(A°)**
DZP	2	Hydrogen Bond	Conventional H-Bond	D:SER206:H	H-Donor	D:DZP404:O1	H-Acceptor	2,68
	4			D:DZP404:H151		D:SER205:OG		2,55
	2		C-H Bond	D:SER205:HA		D:DZP404:O1		2,65
	1			D:DZP404:H202		D:GLN204:O		2,94
	4	Hydrogen Bond; Electrostatic	Pi-Cation; Pi-Donor H-Bond	D:DZP404:H152	Positive; H-Donor	E:PHE77	Pi-Orbitals; Pi-Orbitals	3,03
	Ring C	Hydrophobic	Pi-Pi Stacked	D:TYR160	Pi-Orbitals	D:DZP404	Pi-Orbitals	5,24
				D:TYR210		D:DZP404		4,22
				E:PHE77		D:DZP404		5,49
			Pi-Pi T-shaped	D:PHE100		D:DZP404		5,36
	1		Alkyl	D:DZP404:C20	Alkyl	D:VAL203	Alkyl	5,21
	7		Pi-Alkyl	D:HIS102	Pi-Orbitals	D:DZP404:CL		4,32
	1			E:TYR58		D:DZP404:C20		4,37
	7			E:PHE77		D:DZP404:CL		5,12
	Ring B			D:DZP404		D:VAL203		5,24
Ro12-6377	7	Hydrogen Bond	Conventional H-Bond	D:SER206:HG	H-Donor	:*0:O	H-Acceptor	2,57
				:*0:H		E:PHE77:O		3,06
	2’		C-H Bond	D:HIS102:HE1		:*0:F		2,47
	7			D:SER206:HB3		:*0:O		2,70
				:*0:H		E:PHE77:O		2,98
	Internal [c]	Halogen	Halogen (Fluorine)	:*0:C	Halogen Acceptor	:*0:F	Halogen	3,66
	Ring B	Hydrophobic	Pi-Pi Stacked	E:TYR58	Pi-Orbitals	:*0	Pi-Orbitals	3,93
				E:PHE77		:*0		5.00
	Ring C		Pi-Pi T-shaped	D:PHE100		:*0		5,23
	7		Alkyl	E:ALA79	Alkyl	:*0:C	Alkyl	3,55
	1			:*0:C		D:VAL203		4,55
			Pi-Alkyl	E:TYR58	Pi-Orbitals	:*0:C		4,88
PLZ	1	Hydrogen Bond	Conventional H-Bond	D:THR207:HG1	H-Donor	:*0:O	H-Acceptor	2,10
				D:TYR210:HH		:*0:O		2,85
	4		C-H Bond	D:HIS102:HE1		:*0:N		2,57
	Ring B	Hydrophobic	Pi-Pi Stacked	E:TYR58	Pi-Orbitals	:*0	Pi-Orbitals	4,73
				E:PHE77		:*0		4,55
	7		Pi-Alkyl	E:TYR58		:*0:CL	Alkyl	3,43

a: position of interaction, **b**: distance of interaction, **c**: internal interaction, ***0:** ligand structure

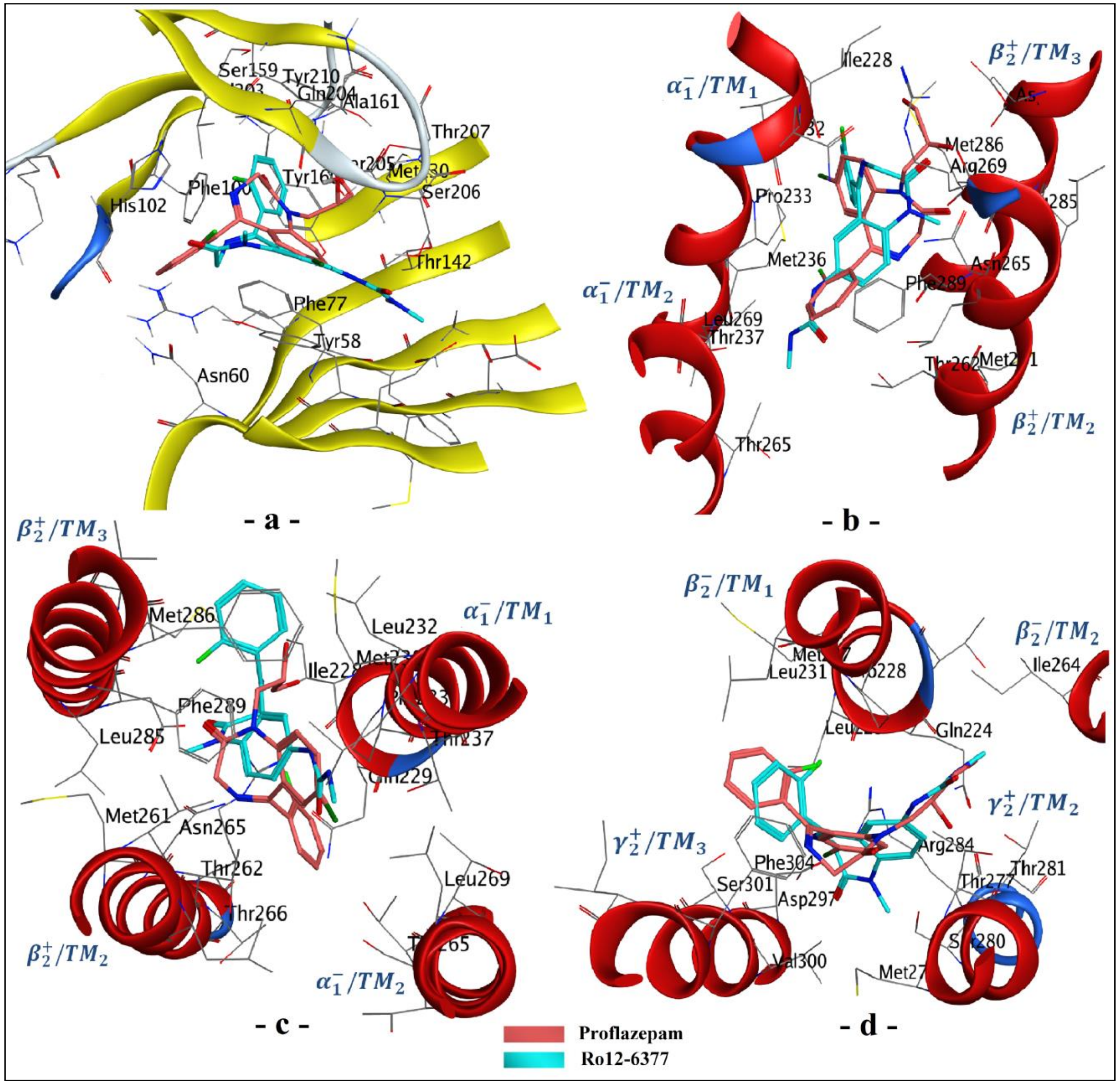


Figure IV.2. Binding modes resulting from the molecular docking of Ro12-6377 and proflazepam at the interfaces of: **a**, ECD α_1^+/γ_2^-. **b**, TMD β_2^+(A)/α_1^-(B). **c**, TMD β_2^+(C)/α_1^-(D), and d, TMD γ_2^+/β_2^-.

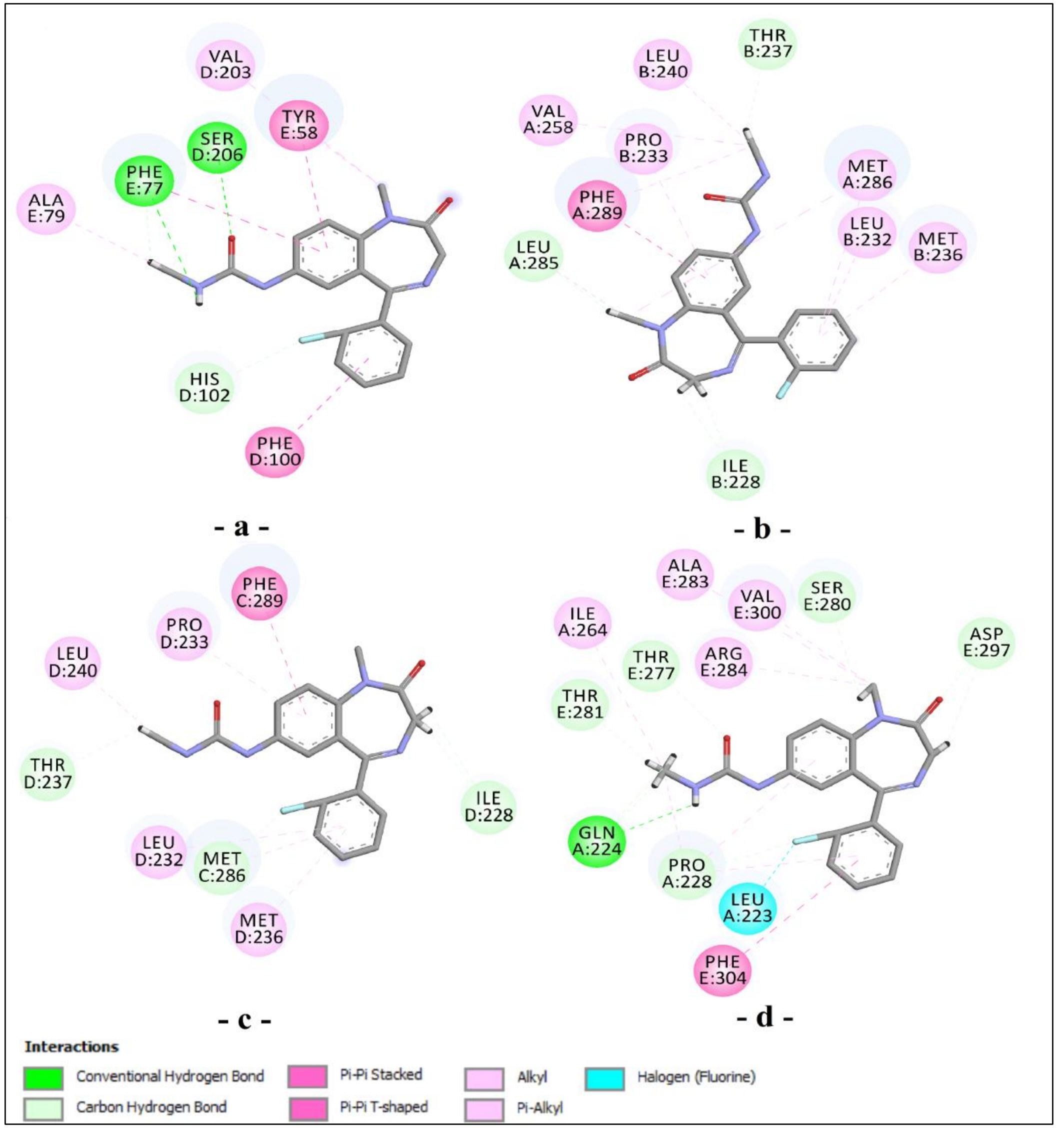


Figure IV.1. Binding interactions resulting from the molecular docking of Ro12-6377 at the interfaces of: **a**, ECD α_1^+/γ_2^-. **b**, TMD β_2^+(A)/α_1^-(B).**c**, TMD β_2^+(C)/α_1^-(D), and **d**, TMD γ_2^+/β_2^-.

Compared with Ro12-6377 and DZP, proflazepam penetrates less within the binding site (Figure IV.1 (a), Figure IV.3 (a), Table IV.4). Phenyl (C) is oriented outside the binding pocket, which explains the absence of any interaction on this pendant ring. Rings (A) and (B) seem to selectively accept rather than accept and donate bonds to the residues of subunits. The α_1 subunit acts through three H-bonds; two are oriented from the hydroxyl groups at the side chains of Thr207 and Tyr210 toward the first hydroxyl group of the $CH_2CHOHCH_2OH$ substitute, and one is oriented from the side chain of His102 toward N_4. The γ_2 subunit prefers to act hydrophobically through two types of interactions: Pi-Pi Stacked and Pi-Alkyl. The Pi-Pi Stacked is created between the π-electron cloud of phenyl (B) and the Tyr58 and Phe77 side chains, while the Pi-Alkyl interaction links the chlorine atom at C_7 to the side chain of Tyr58.

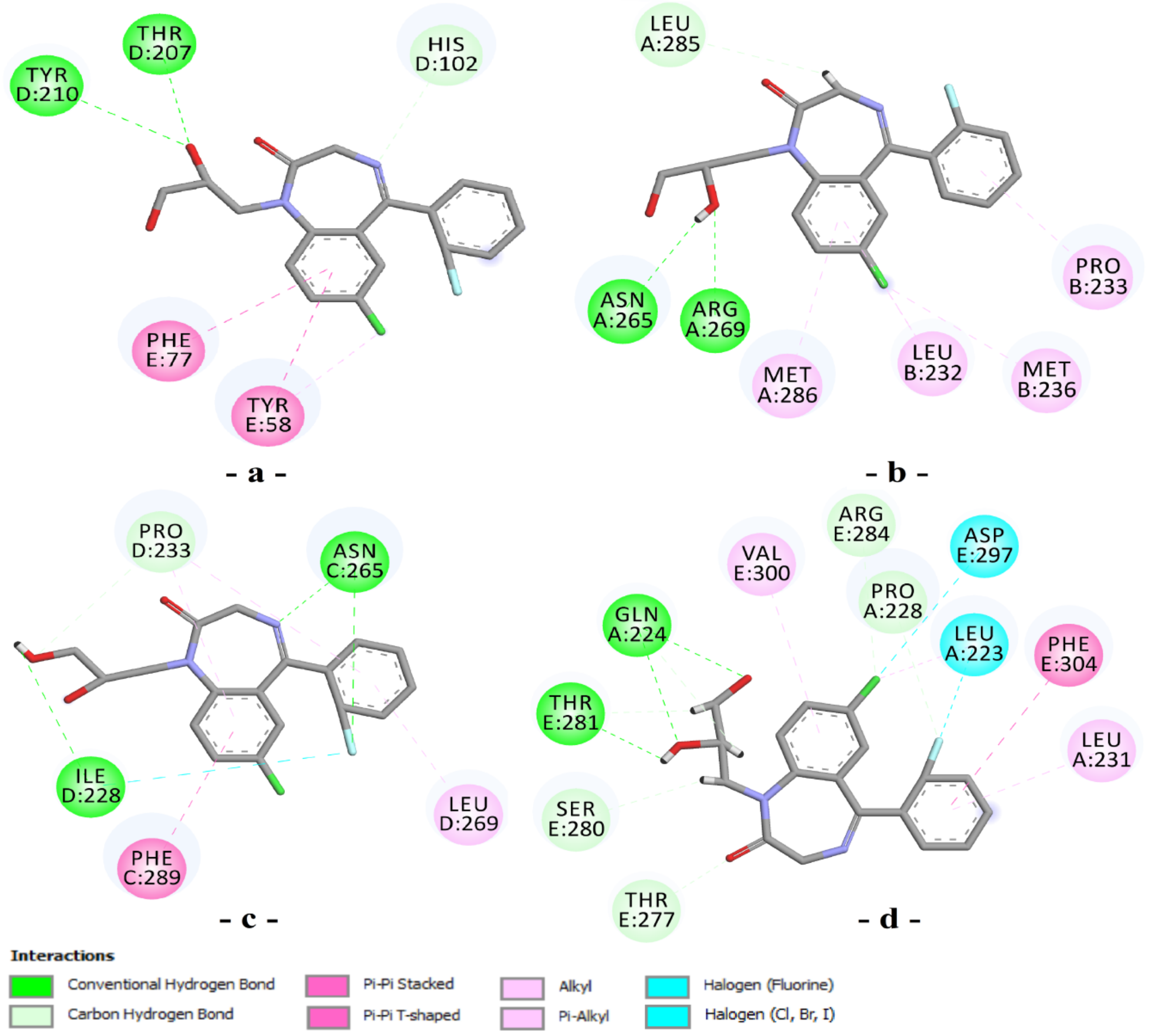


Figure IV.3. Binding interactions resulting from the molecular docking of proflazepam at the interfaces of: **a**, ECD α_1^+/γ_2^-. **b**, TMD $\beta_2^+(A)/\alpha_1^-(B)$. **c**, TMD $\beta_2^+(C)/\alpha_1^-(D)$, and **d**, TMD γ_2^+/β_2^-.

The binding modes of Ro12-6377 and proflazepam elucidated here agree with the previous finding indicating the importance of His102 in the recognition of classical benzodiazepines. With the exception of the His102Cys mutant in the α_5 subunit, $GABA_A$Rs that contain αHis102 mutation to any other residue suffer from total insensitivity to the DZP and its analogs [42]. The topological organization of the $\alpha_4\beta\gamma_2$ and $\alpha_6\beta\gamma_2$ receptors reveals the presence of a natural substitution of His102 by the Arg residue, which leads to steric problems affecting the binding of classical BDZ at their correspondent binding locus. This could explain the selectivity of classical BDZ towards $\alpha\beta\gamma_2$Rs containing the α_1, α_2, α_3, and α_5 subunits rather than those containing the α_4 and α_6 subunits [43,44]. Else, the γ_2Phe77Tyr mutant affects less the binding affinity of DZP but more strongly reduces that of its analogs containing the chlorine substitutes at the pendant phenyl (C). This finding was explained by the difference in flexibility between the two pendant phenyls since the presence of the chlorine atoms possibly caused unfavorable steric clashes with the side chain of the tyrosine residue. Furthermore, several other mutation findings were surveyed in detail in the previous researches [45–47].

3.1.2 Allosteric modulation of the three TMD binding sites

Despite the common topological organization between the combinations β_2(A)/α_1(B) and β_2(C)/α_1(D), the two TMD orthosteric pockets inserted at $\beta_2^+(\mathrm{A})/\alpha_1^-(\mathrm{B})$ and $\beta_2^+(\mathrm{C})/\alpha_1^-(\mathrm{D})$ interfaces are not qualitatively identical and thus may not be functionally equivalent. The binding pocket at $\beta_2^+(\mathrm{A})/\alpha_1^-(\mathrm{B})$ interface was rated as the largest, and estimated to have a higher affinity for Ro12-6377 and Proflazepam than the binding pocket at $\beta_2^+(\mathrm{C})/\alpha_1^-(\mathrm{D})$ interface.

The binding modes of Ro12-6377 for both TMD β_2^+/α_1^- interfaces are equivalents: the $NHCONHCH_3$ groups and the fused benzodiazepine rings (A and B) are deeply embedded in the binding pockets. Whereas, the pendant phenyl rings (C) interact with the residues in front of the pockets (Figure IV.1 (b) and (c)). By superposing the binding modes, a rotation of 98.8° was observed between the two poses of the phenyl rings (C) (Appendix G). The binding interactions at $\beta_2^+(\mathrm{C})/\alpha_1^-(\mathrm{D})$ interface exhibit the formation of halogen interaction between the bound fluorine atom at $C_{2'}$ and N_1 (Table IV.6). This intramolecular interaction leads the fluorine to orient toward the principal chain of β_2Met 286 where it stabilizes through the formation of a strong H-bond interaction (Figure IV.2 (c)). Consequently, leads the phenyl ring (C) to deviate by an angle of 98.8° from the phenyl plane observed at the $\beta_2^+(\mathrm{A})/\alpha_1^-(\mathrm{B})$ interface (Figure IV.2 (b)). At this latter, the binding interactions exhibit the formation of additional electrostatic and hydrophobic interactions between the substitutes groups at N_1 and C_7 and the side chains of four residues from the β2 subunit: Val258, Leu285, Met286, and Phe289 (Table IV.5).

Table IV.5. Detailed binding interactions resulting from the molecular docking of co-crystallized DZP, Ro12-6377, and proflazepam (PLZ) at the TMD β_2^+(A)/α_1^-(B) interface.

TMD β_2^+(A)/α_1^-(B) interface								
Ligand Name	**Posit** [a]	**category**	**Type of interactions**	**From**	**From chemistry**	**To**	**To chemistry**	**Dist** [b] **(A°)**
DZP	1	Hydrogen Bond	C-H Bond	A:DZP406:H201	H-Donor	B:ILE228:O	H-Acceptor	2,93
				A:DZP406:H202		B:ILE228:O		2,89
	4	Electrostatic	Pi-Cation	A:DZP406:N15	Positive	A:PHE289	Pi-Orbitals	4,22
	1	Hydrophobic	Alkyl	A:DZP406:C20	Alkyl	A:MET286	Alkyl	4,50
				A:DZP406:C20		B:ILE228		4,46
	7			A:DZP406:CL		A:MET261		4,42
				A:DZP406:CL		A:LEU285		4,79
			Pi-Alkyl	A:PHE289	Pi-Orbitals	A:DZP406:CL		4,86
	Ring C			A:DZP406		B:PRO233		4,28
Ro12-6377	3	Hydrogen Bond	C-H Bond	:*0:H	H-Donor	B:ILE228:O	H-Acceptor	2,78
				:*0:H		B:ILE228:O		2,89
	1			:*0:H		A:LEU285:O		3,04
	7			:*0:H		B:THR237:OG1		2,51
	Ring B	Hydrophobic	Pi-Pi Stacked	A:PHE289	Pi-Orbitals	:*0	Pi-Orbitals	4,11
	1		Alkyl	:*0:C	Alkyl	A:LEU285	Alkyl	4,64
				:*0:C		A:MET286		5,39
	7			:*0:C		A:VAL258		5,37
				:*0:C		B:LEU240		5,04
			Pi-Alkyl	A:PHE289	Pi-Orbitals	:*0:C		5,30
	Ring B			:*0		B:PRO233		4,65
	Ring C			:*0		A:MET286		4,67
				:*0		B:LEU232		5,20
				:*0		B:MET236		4,66
PLZ	1	Hydrogen Bond	Conventional H-Bond	A:ARG269:HH12	H-Donor	:*0:O	H-Acceptor	2,10
				:*0:H		A:ASN265:OD1		1,97
	Internal [c]			:*0:H		:*0:O		2,44
	1		C-H Bond	A:ARG269:HD3		:*0:O		2,72
	3			:*0:H		A:LEU285:O		2,90
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:LEU232	Alkyl	4,20
				:*0:CL		B:MET236		3,73
	Ring B		Pi-Alkyl	:*0	Pi-Orbitals	A:MET286		4,91
				:*0		B:LEU232		5,42
	Ring C			:*0		B:PRO233		4,28

a: position of interaction, **b**: distance of interaction, **c**: internal interaction, ***0:** ligand structure

Table IV.6. Detailed binding interactions resulting from the molecular docking of co-crystallized DZP, Ro12-6377, and proflazepam (PLZ) at the TMD $\beta_2^+(C)/\alpha_1^-(D)$ interface.

TMD $\beta_2^+(C)/\alpha_1^-(D)$ interface								
Ligand Name	**Posit [a]**	**category**	**Type of interactions**	**From**	**From chemistry**	**To**	**To chemistry**	**Dist [b] (A°)**
DZP	1	Hydrogen Bond	C-H Bond	C:DZP406:H201	H-Donor	D:ILE228:O	H-Acceptor	3,08
				C:DZP406:H202		D:ILE228:O		2,56
	4	Hydrogen Bond; Electrostatic	Pi-Cation; Pi-Donor H-Bond	C:DZP406:H152	Positive; H-Donor	C:PHE289	Pi-Orbitals; Pi-Orbitals	2,65
	Ring C	Hydrophobic	Pi-Sigma	D:PRO233:HB2	C-H	C:DZP406	Pi-Orbitals	2,60
	1		Alkyl	C:DZP406:C20	Alkyl	C:MET286	Alkyl	4,68
				C:DZP406:C20		D:ILE228		4,76
				C:DZP406:C20		D:LEU232		4,82
				C:DZP406:C20		D:PRO233		4,46
	7			C:DZP406:CL		C:MET261		4,61
				C:DZP406:CL		C:LEU285		4,41
			Pi-Alkyl	C:PHE289	Pi-Orbitals	C:DZP406:CL		4,99
	Ring C			C:DZP406		D:LEU269		5,50
	Ring B			C:DZP406		C:MET286		5,39
Ro12-6377	2’	Hydrogen Bond; Halogen	C-H Bond; Halogen (Fluorine)	C:MET286:HA	H-Donor; Halogen Acceptor	:*0:F	H-Acceptor; Halogen	2,19
	3	Hydrogen Bond	C-H Bond	:*0:H	H-Donor	D:ILE228:O	H-Acceptor	2,95
				:*0:H		D:ILE228:O		3,03
	7			:*0:H		D:THR237:OG1		2,79
	Internal [c]	Halogen	Halogen (Fluorine)	:*0:N	Halogen Acceptor	:*0:F	Halogen	3,50
	Ring B	Hydrophobic	Pi-Pi Stacked	C:PHE289	Pi-Orbitals	:*0	Pi-Orbitals	4,14
	7		Alkyl	:*0:C	Alkyl	D:LEU240	Alkyl	4,63
	Ring B		Pi-Alkyl	:*0	Pi-Orbitals	D:PRO233		4,40
	Ring C			:*0		C:MET286		4,54
				:*0		D:LEU232		4,94
				:*0		D:MET236		4,87
PLZ	2’	Hydrogen Bond; Halogen	Conventional H-Bond; Halogen (Fluorine)	C:ASN265:HD21	H-Donor; Halogen Acceptor	:*0:F	H-Acceptor; Halogen	2,46
	4	Hydrogen Bond	Conventional H-Bond	C:ASN265:HD22	H-Donor	:*0:N	H-Acceptor	1,90
	1			:*0:H		D:ILE228:O		2,25
			C-H Bond	D:PRO233:HD3		:*0:O		2,58
	Internal [c]			:*0:H		:*0:F		2,53
	Internal [c]			:*0:H		:*0:O		3,07
	2’	Halogen	Halogen (Fluorine)	D:ILE228:O	Halogen Acceptor	:*0:F	Halogen	3,38
	Ring B	Hydrophobic	Pi-Pi Stacked	C:PHE289	Pi-Orbitals	:*0	Pi-Orbitals	4,34
			Pi-Alkyl	:*0		D:PRO233	Alkyl	4,46
	Ring C			:*0		D:PRO233		4,52
				:*0		D:LEU269		5,19

a: position of interaction, **b**: distance of interaction, **c**: internal interaction, ***0:** ligand structure

Unlike Ro12-6377, the binding modes of proflazepam for both TMD $\beta_2^+(A)/\alpha_1^-(B)$ and $\beta_2^+(C)/\alpha_1^-(D)$ interfaces reveal significant differences in the docking orientations and binding interactions (Figure IV.1 (b) and (c)). At $\beta_2^+(A)/\alpha_1^-(B)$ interface (Figure IV.3 (b),

Table IV.5), the chlorobenzene ring (B) was oriented to interact hydrophobically with side chains of residues located in front of the pocket: β_2Met286, α_1Met236, and α_1Leu232. The pendant phenyl ring (C) was embedded deeper into the binding pocket, in a manner similar to that observed in DZP, and shares Pi-Alkyl interaction at a distance of 4.28 A° with side chain of α_1Pro233. The diazepine ring (A) established a moderate H-bond between one of the hydrogens of C_3 and the main carbonyl of β_2Leu 285. The $CH_2CHOHCH_2OH$ group was oriented towards the ECD where its first hydroxyl group was stabilized by strong intramolecular interaction type H-Donor with lone pairs of the oxygen bond at C_2, and with three intermolecular H-bonds formed with side chains of Arg269 and Asn265 of subunit β_2. At $\beta_2^+(C)/\alpha_1^-(D)$ interface (Figure IV.3 (c), Table IV.6), the pendant phenyl (C) is located at a higher level than rings (A) and (B), with its fluorine atom pointing toward the front of the binding-site. This orientation drives the π-electron cloud into Pi-Alkyl interactions with side chains of Pro233 and Leu269 of subunit α_1. As well as, driven the Lone pairs of the fluorine atom to form Halogen-Halogen interaction with the main oxygen of α_1Ile228 and receive a strong H-bond from the side chain of β_2Asn265. This latter also shares a strong H-bond with N_4. Phenyl (B) is located between the side rings of β_2Phe289 and α_1Pro233, which contributes to the formation of two hydrophobic interactions types Pi-Pi Stacked and Pi-Alkyl. Its chlorine atom is oriented towards the bottom of the binding pocket, more precisely, toward Thr265 of α_1 subunit. Here, no interactions were observed. The $CH_2CHOHCH_2OH$ group exerts its influence by occupying the front of the binding-pocket and sharing two H-bonds with α_1Ile228 and α_1Pro233. The length of its backbone skeleton also leads to forming two additional intramolecular bonds with the oxygen atom at C_2 and the fluorine at $C_{2'}$.

By examining the binding mode of Ro12-6377 at the γ_2^+/β_2^- interface (Figure IV.1 (d), Figure IV.2 (d), Table IV.7), both the diazepine and phenyl (C) are positioned in front of the pocket. The methyl group attached at N_1 forms two types of interactions. The first is a moderate H-Donor bond is given to the carbonyl of γ_2Ser280. The seconds are three hydrophobic bonds type Alkyl-Alkyl formed with side chains of γ_2Ala283, γ_2Arg284, and γ_2Val300. Simultaneously, γ_2Asp297 established two hydrogen bonds with the diazepine ring (A); strong H-Donor has given to the oxygen attached at C_2, and moderate H-acceptor bond received from C_3. The Phenyl (C) involved in two hydrophobic interactions; type Pi-Pi T-Shaped with side chain of γ_2Phe304, and type Pi-Alkyl with the side chain of β_2Pro228. Its fluorine atom was oriented toward TM_1:β_2 subunit, which resulted in Halogen interaction with the lone pair of the oxygen situated in the main chain of Leu223, and in moderate H-Acceptor bond forms with the side chain of Pro228. On the other hand, the phenyl (B) was inserted deeper in the binding pocket as its π-electron cloud shared Pi-Alkyl interaction with side chain of β_2Pro228. Its $NHCONHCH_3$ group attached at C_7 was pointed towards the TM_2:β_2 helix. In addition to the two Alkyl-Alkyl interactions shared with side chains of Pro228 and Ile264 of β_2 subunit, the $NHCONHCH_3$ group also received one moderate H-bond from side chain of γ_2Thr277 and simultaneously gave four moderate H-bonds, three of which to β_2Gln224. Thus, leads to the formation of two intermolecular pentameric rings and one butameric ring contribute, as in the classical site, to the stability of Ro12-6377 at the binding site.

Proflazepam adopted a similar binding mode of DZP in the binding locus (Appendix F (d)). The binding mode of DZP at the γ_2^+/β_2^- interface was previously discussed in detail by *J. J. Kim et al* [7]. The uncommon binding interactions observed between the both modulators are related to the presence of two distinct binding groups on the structure of proflazepam: the fluorine atom at $C_{2'}$ and the long backbone skeleton of $CH_2CHOHCH_2OH$ at N_1. The phenyl (C) is placed in front of the pocket (Figure IV.1 (d), Figure IV.3 (d), Table IV.7), its π-electron cloud establishes two interactions of types Pi-Pi T-shaped and Pi-Alkyl with side chains of γ_2Phe304 and β_2Leu231, respectively. Its fluorine atom points towards the TM_1:β_2 subunit, where receives a moderate H-bond from the side chain of Pro228 and establishes Halogen interaction with the main oxygen of Leu223. The phenyl (B) binds at a higher level than rings (A) and (C), with its chlorine atom pointing towards the TM_3:γ_2 subunit. This orientation results in Pi-Alkyl interaction between the π-electron cloud and γ_2Val300. Also, leads the chlorine to accept the moderate H-bond from γ_2Arg284, forms halogen interaction with the side chain of γ_2Asp297, and interacting hydrophobically with β_2Leu223. The binding modes of the oxygen atom attached at C_2 in the three α_1^+/γ_2^-, $\beta_2^+(A)/\alpha_1^-(B)$, and $\beta_2^+(C)/\alpha_1^-(D)$ interfaces reflect the insensitivity of Ro12-6377 and proflazepam to participate in any interaction with the neighboring residues using this position. Otherwise, the γ_2^+/β_2^- interface reflects the contribution of the oxygen atom at this position to enable the receptor potentiation by accepting, respectively, strong and moderate H-bonds from Asp297 and Thr277 of subunit γ_2. The $CH_2CHOHCH_2OH$ group penetrates deep into the binding site. Its pose and orientation towards the β_2:A:TM_2 helix are identical to that observed for the $NHCONHCH_3$ group of Ro12-6377. The length of its backbone skeleton allowed it to simultaneously influence the γ_2:TM_2 helix and the β_2:A:TM_1 by forming six H-bonds with γ_2:TM_2:Ser280, γ_2:TM_2:Thr281, and β_2:A:TM_1:Gln224.

Table IV.7. Detailed binding interactions resulting from the molecular docking of co-crystallized DZP, Ro12-6377, and proflazepam (PLZ) at the TMD γ_2^+/β_2^- interface.

TMD γ_2^+/β_2^- interface								
Ligand Name	**Posit** [a]	**category**	**Type of interactions**	**From**	**From chemistry**	**To**	**To chemistry**	**Dist** [b] **(A°)**
DZP	3	Hydrogen Bond	C-H Bond	E:DZP403:H171	H-Donor	E:SER280:OG	H-Acceptor	2,77
	1			E:DZP403:H202		E:THR281:OG1		2,80
	Ring C	Hydrophobic	Pi-Pi T-shaped	E:PHE304	Pi-Orbitals	E:DZP403	Pi-Orbitals	5,81
	1		Alkyl	E:DZP403:C20	Alkyl	A:PRO228	Alkyl	4,25
	7			E:DZP403:CL		A:LEU223		4,64
				E:DZP403:CL		E:VAL300		4,38
	Ring B		Pi-Alkyl	E:DZP403	Pi-Orbitals	E:VAL300		4,92
Ro12-6377	7	Hydrogen Bond	Conventional H-Bond	:*0:H	H-Donor	A:GLN224:O	H-Acceptor	2,43
	2’		C-H Bond	A:PRO228:HD3		:*0:F		2,61
	7			E:THR277:HB		:*0:O		2,73
	2			E:ASP297:HA		:*0:O		2,28
	3			:*0:H		E:ASP297:OD1		2,89
	1			:*0:H		E:SER280:O		2,71
	7			:*0:H		A:GLN224:O		2,97
				:*0:H		E:THR281:OG1		2,78
				:*0:H		A:GLN224:O		2,94
	2’	Halogen	Halogen (Fluorine)	A:LEU223:O	Halogen Acceptor	:*0:F	Halogen	2,56
	Ring C	Hydrophobic	Pi-Pi T-shaped	E:PHE304	Pi-Orbitals	:*0	Pi-Orbitals	5,40
	1		Alkyl	E:ALA283	Alkyl	:*0:C	Alkyl	4,28
				:*0:C		E:ARG284		3,96
				:*0:C		E:VAL300		3,90
	7			:*0:C		A:PRO228		4,75
				:*0:C		A:ILE264		3,81
	Ring B		Pi-Alkyl	:*0	Pi-Orbitals	A:PRO228		5,11
	Ring C			:*0		A:PRO228		5,29
PLZ	1	Hydrogen Bond	Conventional H-Bond	A:GLN224:HE21	H-Donor	:*0:O	H-Acceptor	2,32
				A:GLN224:HE22		:*0:O		2,59
				:*0:H		E:THR281:OG1		1,92
	2’		C-H Bond	A:PRO228:HD3		:*0:F		2,80
	2			E:THR277:HA		:*0:O		2,79
	7			E:ARG284:HD3		:*0:CL		2,83
	1			:*0:H		E:SER280:OG		2,74
				:*0:H		A:GLN224:O		2,81
				:*0:H		E:THR281:OG1		2,37
	7	Halogen	Halogen (Cl, Br, I)	E:ASP297:OD1	Halogen Acceptor	:*0:CL	Halogen	3,18
	2’		Halogen (Fluorine)	A:LEU223:O		:*0:F		2,90
	Ring C	Hydrophobic	Pi-Pi T-shaped	E:PHE304	Pi-Orbitals	:*0	Pi-Orbitals	5,82
	7		Alkyl	:*0:CL	Alkyl	A:LEU223	Alkyl	4,29
	Ring B		Pi-Alkyl	:*0	Pi-Orbitals	E:VAL300		4,92
	Ring C			:*0		A:LEU231		5,35
	1	Unfavorable	Unfavorable Acceptor-Acceptor	A:GLN224:O	H-Acceptor	:*0:O	H-Acceptor	2,83

a: position of interaction, **b**: distance of interaction, **c**: internal interaction, ***0:** ligand structure

The detailed mechanisms of interactions by which proflazepam modulates the two TMD $\beta_2^+(A)/\alpha_1^-(B)$ and $\beta_2^+(C)/\alpha_1^-(D)$ sites are appearing to be complementary to each other. As can be seen, the $\beta_2^+(A)/\alpha_1^-(B)$ interface has been predicted to specifically modulate by the first hydroxyl group of $CH_2CHOHCH_2OH$, C_3, the chlorine atom at C_7, and the π-electron clouds of both phenyls (B) and (C). Whereas, the $\beta_2^+(C)/\alpha_1^-(D)$ interface has been predicted to specifically modulate by the second hydroxyl group of $CH_2CHOHCH_2OH$, N_4, the fluorine atom at $C_{2'}$, and the π-electron cloud of both phenyls (B) and (C). This selectivity in binding interactions between the two orthosteric binding sites allows us to hypothesize that they may be functionally in a complementary manner, as it previously observed for the two agonists (GABA) binding sites which do not yet know an explanation for why they are structurally identical and functionally not equivalent [43]. Furthermore, the bulky structure of proflazepam allowed it to directly induce influence on the pore-lining helices β_2:A:TM_2, β_2:C:TM_2, and α_1:D:TM_2 by creating H-bonds with β_2:A:Arg269, β_2:A:Asn265, β_2:C:Asn265, and interacting hydrophobically with α_1:D:Leu269. Likewise, three hydrophobic interactions with β_2:A:TM_2:Met261, β_2:C:TM_2:Met261, and α_1:D:TM_2:Leu269 are observed for the DZP. Whereas, just one hydrophobic interaction with β_2:A:TM_2:Val258 is established for Ro12-6377. Evidently, rings (C) of both proflazepam and DZP adopt a similar hydrophobic interaction with α_1:D:TM_2:Leu269.

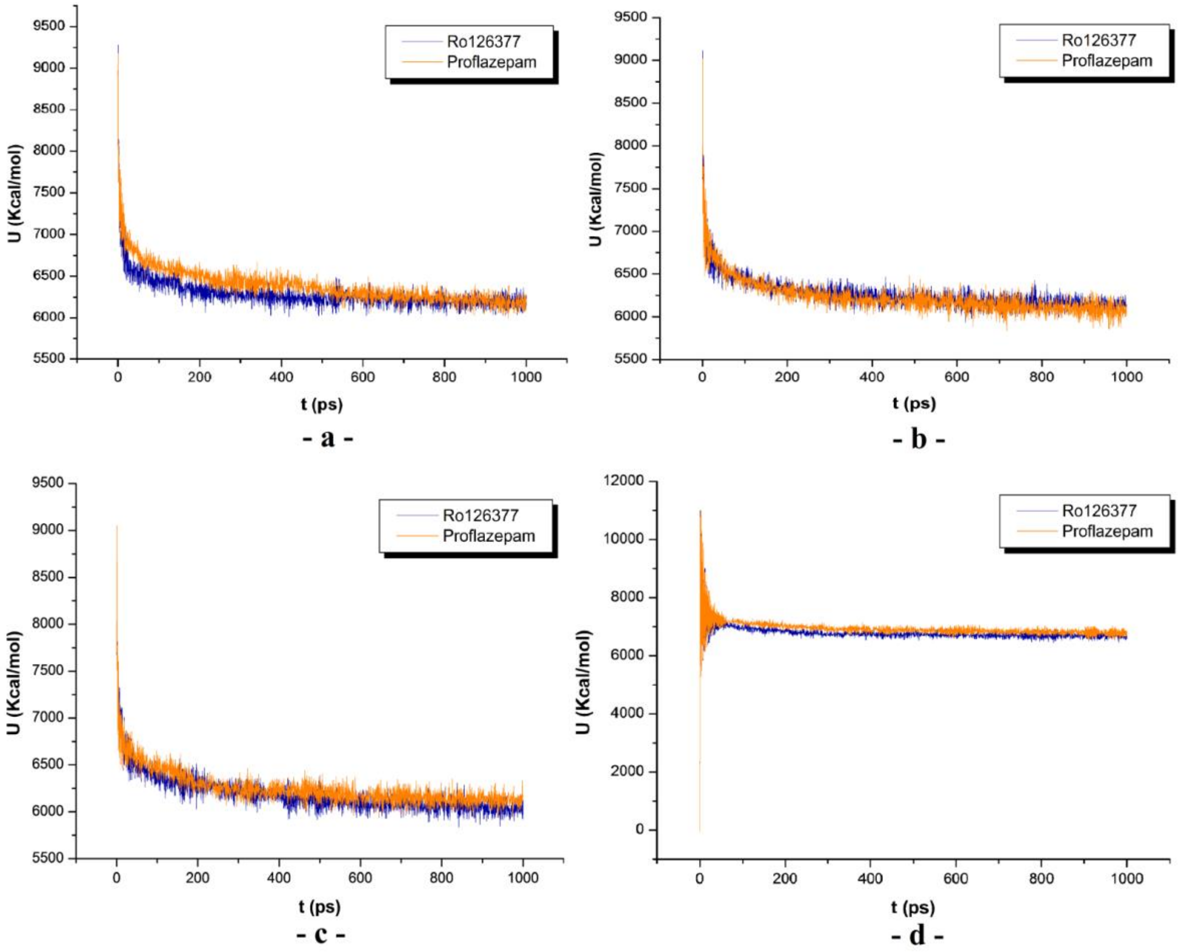
Obviously, at γ_2^+/β_2^- interface, the binding mode of Ro12-6377 is the most influential on pore-lining by sharing five interactions with the residues of γ_2:TM_2 helix and one interaction with β_2:A:TM_2 helix. Likewise, the binding mode of proflazepam is connected to the pore-lining by participating in five interactions with γ_2:TM_2 helix. Accordingly, both modulators exhibit common interactions with γ_2:TM_2 residues: Thr281, Thr277, Arg284, and Ser280. Otherwise, by examining the binding interactions of DZP, its structure predicted to enrich the skeleton of γ_2:TM_2 through two moderate H-bonds originating from the methyl group attached at N_1 to Ser280 and from C_3 to Thr281. These two interactions are identical between the three modulators.

The feature of interacting with the residues of the TM_2 helices is of great importance as it leads both Ro12-6377 and proflazepam to directly induce motions in the chloride-channel lining. Thereby, possibly contributing to the expansion of its diameter by opening the 9' gate by orienting the β_2:C:Leu259 side chain towards one of the two adjacent α subunits. As mentioned earlier, this rotation is the main factor in the activation of the pLGICs family [7,48].

3.2 Molecular dynamics simulation

MD simulation is used as a complementary tool to validate the docking results before they are approved in the drug-design process. MD simulation offers the peculiarity of treating biological systems as flexible entities. This flexibility allows for free integrations between the macromolecule binding site and the binding-ligand, resulting in binding modes (poses or interactions) that confirm or refute the results of molecular docking [49]. Or in some cases, may lead to the release of the ligand from the binding-site, and, this latter is an undesirable defect especially if the ligand shows the highest stability in the docking simulation [50].

For these reasons, Ro12-6377 and proflazepam in complex with the four binding interfaces were subjected to MD simulations using the settled parameters cited in the material and methods section. Their dynamic behaviors were investigated by evaluating the response of the potential energy U (Kcal/mol) over a time period of 1000 picoseconds (ps) (Figure IV.4).



- a -

- b -

- c -

- d -

Figure IV.4. Evaluation the response of potential energy U (Kcal/mol) as function of time t(ps) for Ro12-6377 and proflazepam in complex with: **a**, ECD α_1^+/γ_2^-. **b**, TMD β_2^+(A)/α_1^-(B). **c**, TMD β_2^+(C)/α_1^-(D), and **d**, TMD γ_2^+/β_2^- interfaces.

During the first 500 ps of the simulation, the ECD α_1^+/γ_2^- interface exhibits higher stability in complex with Ro12-6377 than in complex with proflazepam. Afterwise, from 500 ps until the end of the simulation, both complexes tend to have equivalent stability (6270.7 Kcal/mol for Ro12-6377 and 6226.86 Kcal/mol for proflazepam) (Figure IV.4 (a)). Moreover, within the three TMD binding interfaces, the two modulators exhibited stability equivalence throughout all the simulation periods (Figure IV.4 (b), (c), and (d)).

Later, deep analyzes of binding modes, binding orientations, and binding interactions of Ro12-6377 and proflazepam within the four binding interfaces were performed and discussed between the two simulations.

3.2.1 MD simulation analysis of the classical binding site

At ECD α_1^+/γ_2^- interface, the binding poses of Ro12-6377 are equivalents for the two simulations (Figure IV.5 (a), Figure IV.6 (a), and Table IV.8). However, significant differences between the two binding orientations were detected, notably in substitutes at C_2, C_5, and C_7. The new binding orientation predicted for phenyl (C) is driving the $C_{2'}$ bound fluorine atom to move away from the diazepine ring (A), which leads to the disappearance of the intramolecular interaction formed with C_2. All the H-bonds formed with α_1Ser206, γ_2Phe77, and α_1His102 were vanished and replaced by H-bonds given from side chain of α_1Lys156 to the oxygen atom at C_2, and from the $NHCONHCH_3$ group at C_7 to both α_1Tyr160 and γ_2Asp56. Similarly, the hydrophobic interactions suggested with Phe100 and Val203 of α_1 subunit were also replaced by hydrophobic interactions created between the π-orbitals of Tyr160 and Tyr210 of the same subunit and the π-electron cloud of phenyl (C). In contrast, the hydrophobic interactions established with the subunit γ_2 residues (Tyr58, Phe77, and Ala79) were preserved as the same.

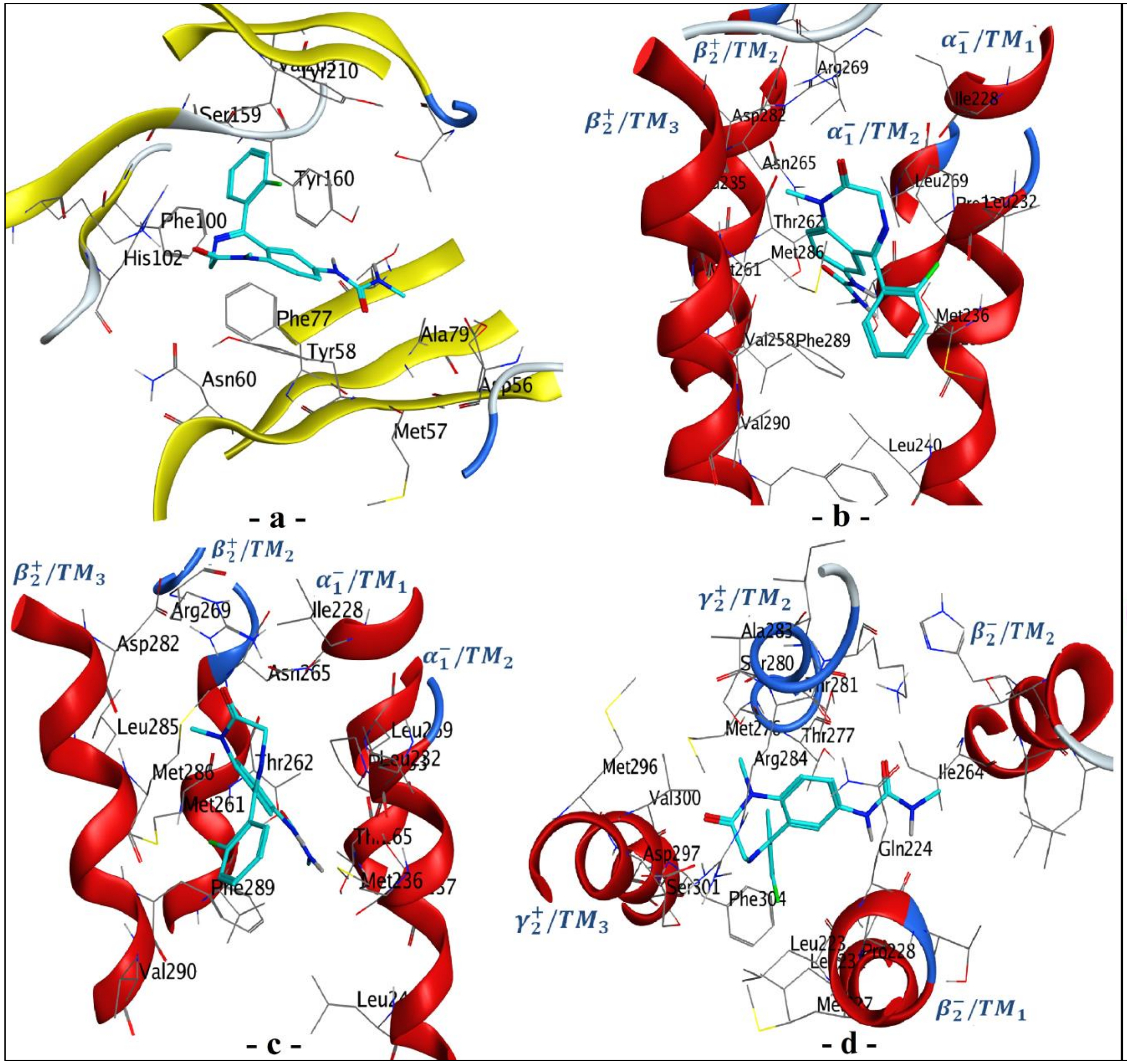


Figure IV.6. Binding modes resulting from the molecular dynamics simulation of Ro12-6377 at the interfaces of: **a**, ECD α_1^+/γ_2^-. **b**, TMD β_2^+(A)/α_1^-(B). **c**, TMD β_2^+(C)/α_1^-(D), and **d**, TMD γ_2^+/β_2^- interfaces.

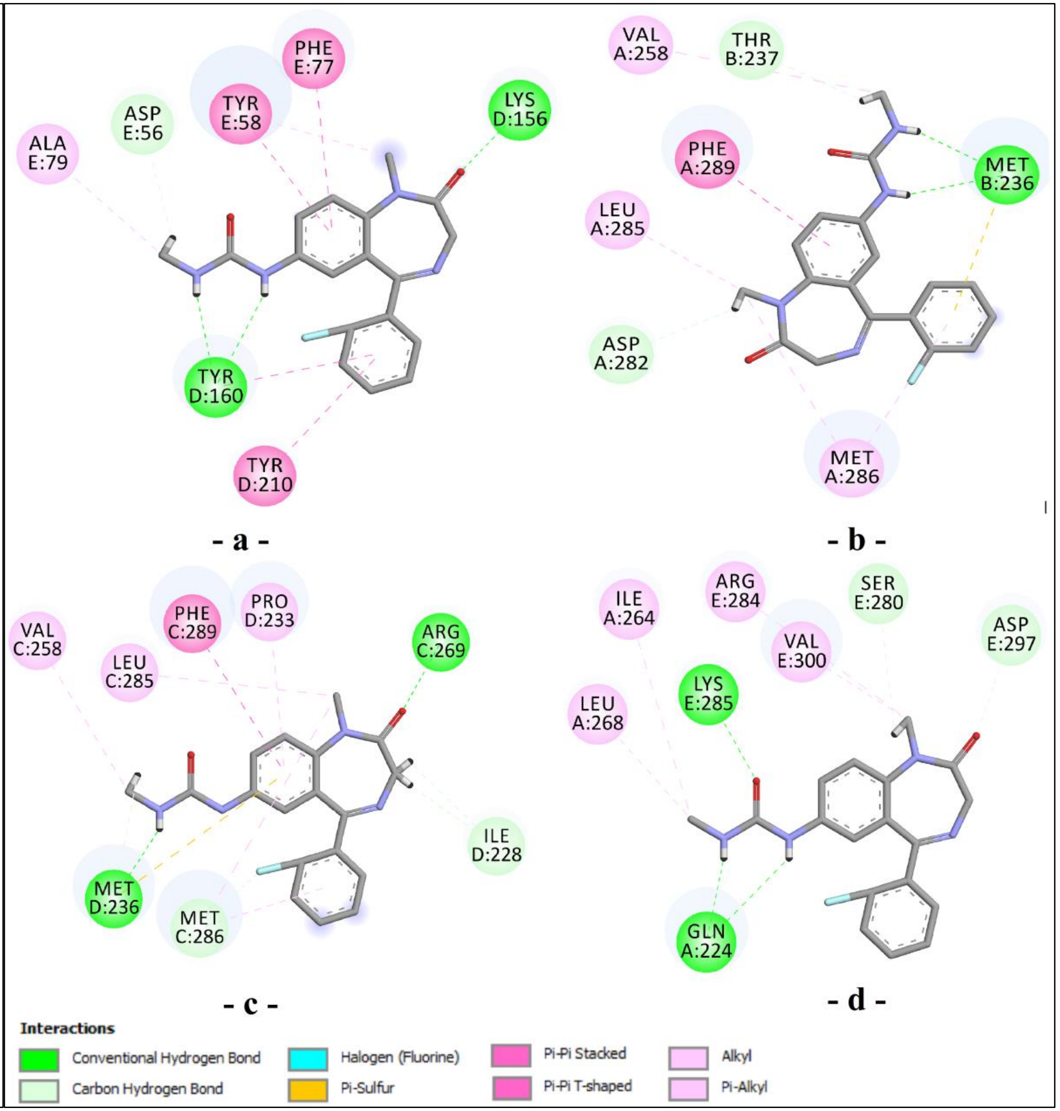


Figure IV.5. Binding interactions resulting from the molecular dynamics simulation of Ro12-6377 at the interfaces of: **a**, ECD α_1^+/γ_2^-. **b**, TMD β_2^+(A)/α_1^-(B). **c**, TMD β_2^+(C)/α_1^-(D), and **d**, TMD γ_2^+/β_2^- interfaces.

Table IV.8. Detailed binding interactions resulting from the molecular dynamics simulation Ro12-6377 and proflazepam (PLZ) at the classical site.

ECD α_1^+/γ_2^- interface								
Ligand Name	**Posit** [a]	**category**	**Type of interactions**	**From**	**From chemistry**	**To**	**To chemistry**	**Dist** [b] **(A°)**
Ro12-6377	2	Hydrogen Bond	Conventional Hydrogen Bond	D:LYS156:HZ3	H-Donor	:*0:O	H-Acceptor	1,67
	7			:*0:H		D:TYR160:OH		2,23
				:*0:H		D:TYR160:OH		2,27
			C-H Bond	:*0:H		E:ASP56:OD1		2,63
	Ring C	Hydrophobic	Pi-Pi Stacked	D:TYR160	Pi-Orbitals	:*0	Pi-Orbitals	4,46
	Ring B			E:TYR58		:*0		4,29
	Ring C		Pi-Pi T-shaped	D:TYR210		:*0		5,39
	Ring B			E:PHE77		:*0		4,68
	7		Alkyl	E:ALA79	Alkyl	:*0:C	Alkyl	3,59
	1		Pi-Alkyl	E:TYR58	Pi-Orbitals	:*0:C	Alkyl	4,47
PLZ	4	Hydrogen Bond	Conventional Hydrogen Bond	D:LYS156:HZ2	H-Donor	:*0:N	H-Acceptor	1,57
	1			:*0:H		D:SER159:O		1,79
	2’		C-H Bond	D:LYS156:HE2		:*0:F		2,51
	Internal [c]			:*0:H		:*0:O		2,51
	Ring C	Electrostatic	Pi-Cation	D:LYS156:NZ	Positive	:*0	Pi-Orbitals	4,25
	1	Hydrogen Bond	Pi-Donor Hydrogen Bond	:*0:H	H-Donor	D:TYR210		2,69
	Ring B	Hydrophobic	Pi-Alkyl	:*0	Pi-Orbitals	D:VAL203	Alkyl	4,72

a: position of interaction, **b**: distance of interaction, **c**: internal interaction, ***0:** ligand structure

The binding modes of proflazepam are inconsistent in the binding poses for the two simulations. The adequate binding pose generated by MD simulation was placed less deeply within the binding-locus than that generated by molecular docking simulation (Figure IV.7 (a)). Unlike molecular docking simulation, this binding-position emerged its structure to react as donor and acceptor with neighboring residues. The most notable differences in the binding orientations appear in the substitutes groups at N_1, C_5, and C_7. By examining the binding-interactions (Figure IV.8 (a) and Table IV.8), the moderate H-bond established between the first hydroxyl group of $CH_2CHOHCH_2OH$ and the side chain of α_1Tyr210 was preserved between the two simulations. Whereas, the remaining interactions were completely vanished and replaced by five interactions with the residues of the same subunit: side chain of Lys156 gives two H-Donors to N_4 and the fluorine atom bond at $C_{2'}$. Also, involved in Pi-Cation interaction with the π-electron cloud of phenyl (C). The main chain of Ser159 receives one H-acceptor bond from the second hydroxyl group of $CH_2CHOHCH_2OH$. A moderate intramolecular H-bond was formed between the oxygen atom at C_2 and the $CH_2CHOHCH_2OH$ group. Finally, hydrophobic interaction type Pi-Alkyl was observed between the π-electron cloud of phenyl (B) and the side chain of Val203.

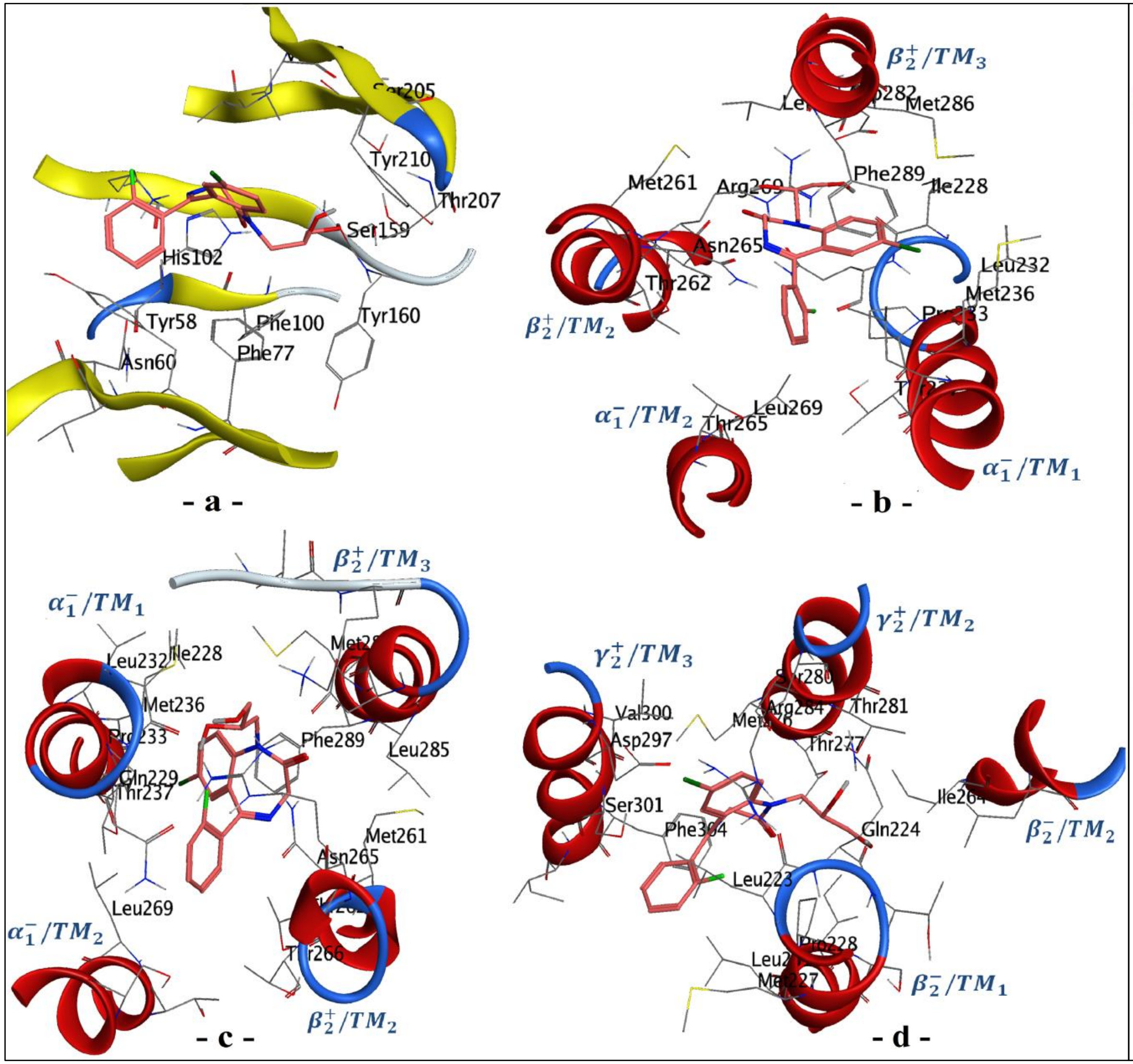


Figure IV.8. Binding modes resulting from the molecular dynamics simulation of proflazepam at the interfaces of: **a**, ECD α_1^+/γ_2^-. **b**, TMD $\beta_2^+(A)/\alpha_1^-(B)$. **c**, TMD $\beta_2^+(C)/\alpha_1^-(D)$, and **d**, TMD γ_2^+/β_2^- interfaces.

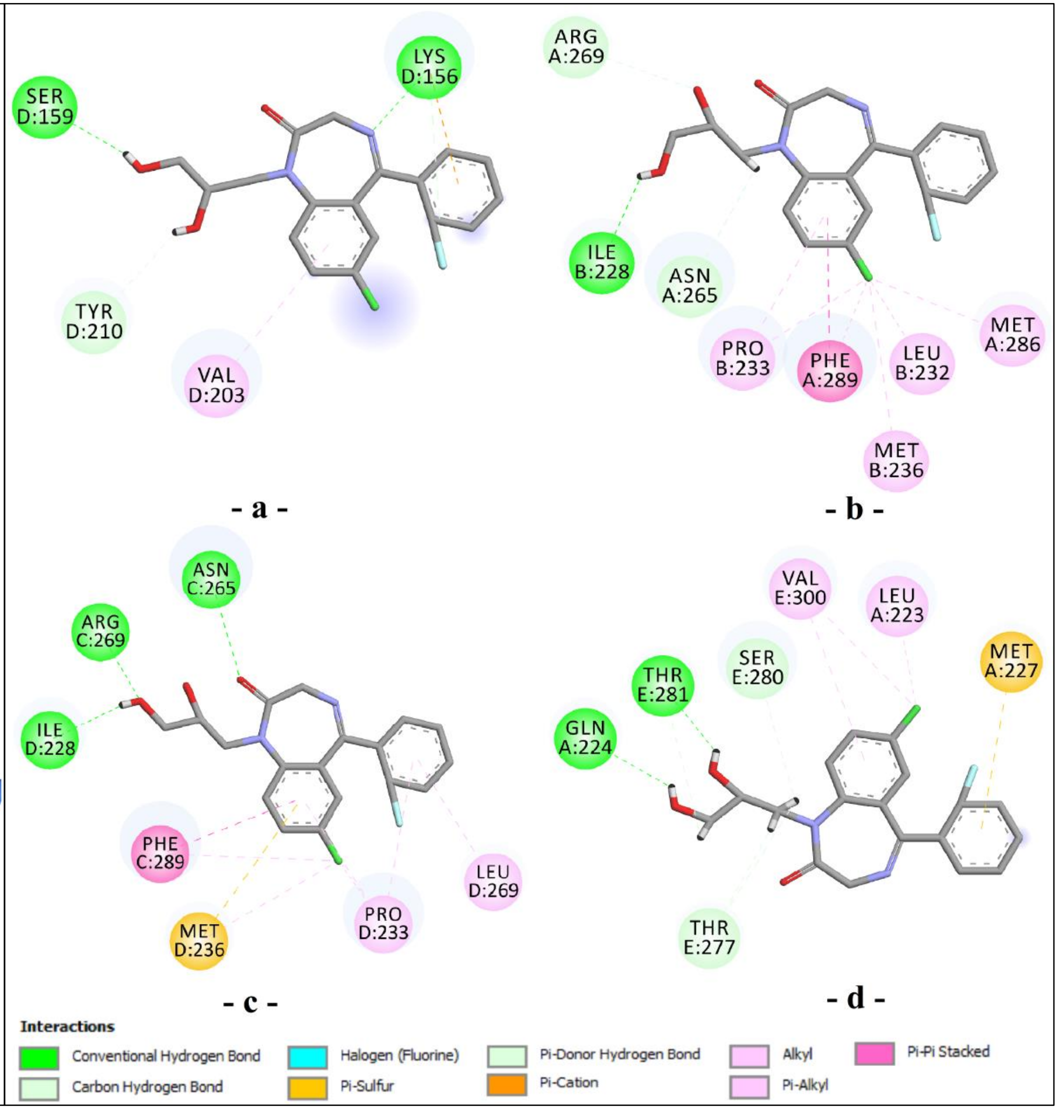


Figure IV.7. Binding interactions resulting from the molecular dynamics simulation of proflazepam at the interfaces of: **a**, ECD α_1^+/γ_2^-. **b**, TMD $\beta_2^+(A)/\alpha_1^-(B)$. **c**, TMD $\beta_2^+(C)/\alpha_1^-(D)$, and **d**, TMD γ_2^+/β_2^- interface.

3.2.2 MD simulation analysis of the three TMD binding sites

The binding modes of Ro12-6377 for both TMD $\beta_2^+(A)/\alpha_1^-(B)$ and $\beta_2^+(C)/\alpha_1^-(D)$ interfaces are consistent in docking-poses for the two simulations (Figure IV.5 (b) and (c)). However, notable differences were observed between the binding orientations of C_5 and C_7 substitutes. The binding interactions at both interfaces exhibit the tendency of Ro12-6377 to enhance its interactions with the residues of β_2 subunit relative to those established with subunit α_1 (Figure IV.6 (b) and (c), Table IV.9, Table IV.10).

Table IV.9. Detailed binding interactions resulting from the molecular dynamics simulation of Ro12-6377 and proflazepam (PLZ) at the TMD $\beta_2^+(A)/\alpha_1^-(B)$ interface.

TMD $\beta_2^+(A)/\alpha_1^-(B)$ interface								
Ligand Name	**Posit** [a]	**category**	**Type of interactions**	**From**	**From chemistry**	**To**	**To chemistry**	**Dist** [b] **(A°)**
Ro12-6377	7	Hydrogen Bond	Conventional	:*0:H	H-Donor	B:MET236:O	H-Acceptor	2,51
			Hydrogen Bond	:*0:H		B:MET236:O		1,85
	1		C-H Bond	:*0:H		A:ASP282:O		2,79
	7			:*0:H		B:THR237:OG1		3,06
	Ring C	Other	Pi-Sulfur	B:MET236:SD	Sulfur	:*0	Pi-Orbitals	3,90
	Ring B	Hydrophobic	Pi-Pi Stacked	A:PHE289	Pi-Orbitals	:*0		4,62
	1		Alkyl	:*0:C	Alkyl	A:LEU285	Alkyl	5,05
				:*0:C		A:MET286		4,81
	7			:*0:C		A:VAL258		5,23
	Ring C		Pi-Alkyl	:*0	Pi-Orbitals	A:MET286		4,84
PLZ	Internal [c]	Hydrogen Bond	Conventional	:*0:H	H-Donor	:*0:O	H-Acceptor	1,63
	1		Hydrogen Bond	:*0:H		B:ILE228:O		1,85
			C-H Bond	A:ARG269:HD2		:*0:O		2,82
				:*0:H		A:ASN265:OD1		2,71
	Ring B	Hydrophobic	Pi-Pi Stacked	A:PHE289	Pi-Orbitals	:*0	Pi-Orbitals	4,73
	7		Alkyl	:*0:CL	Alkyl	A:MET286	Alkyl	5,41
				:*0:CL		B:LEU232		4,88
				:*0:CL		B:PRO233		4,64
				:*0:CL		B:MET236		4,14
	7		Pi-Alkyl	A:PHE289	Pi-Orbitals	:*0:CL		4,39
	Ring B			:*0		B:PRO233		4,39

a: position of interaction, **b**: distance of interaction, **c**: internal interaction, ***0:** ligand structure

Briefly, at the $\beta_2^+(A)/\alpha_1^-(B)$ interface, all interactions created with residues of subunit β_2:A were preserved and stabilized by removing the moderate H-bond and the Pi-Alkyl interaction linking, respectively, the substitutes at N_1 and C_7 with the side chains of Leu285 and Phe289, and, creating instead of them, new moderate H-bond oriented from the methyl group at N_1 towards the main carbonyl group of Asp282. In contrast, only interactions with two residues from α_1:B subunit (Met236 and Thr237) were preserved. The Pi-Alkyl interaction bond the side chain of Met236 to phenyl (C) was vanished and replaced with two

H-bonds oriented from C_7 substitute towards the main carbonyl group, and Pi-Sulfur interaction given from the sulfur atom in the side chain to phenyl (C). At the $\beta_2^+(C)/\alpha_1^-(D)$ interface, the binding interactions with Thr237, Leu240, and Leu232 of α_1:D subunit were vanished and replaced by new interactions with residues from the β_2:C subunit: Arg269, Leu285, and Val258. Side chains of both Leu285 and Val258 were, respectively, involved in Alkyl-Alkyl interactions with the substitutes at N_1 and C_7. While the side chain of Arg269 established a strong H-bond with the oxygen atom at C_2. Likewise, the intramolecular interaction observed between N_1 and the fluorine atom at $C_{2'}$ also vanished and created, instead of it, additional interaction between this latter and the main carbonyl of Met286. At the α_1:D subunit, Likewise to B:Met236, D:Met236 was replacing the Pi-Alkyl interaction connected its side chain to phenyl (C) by two moderate H-bonds and one Pi-Sulfur interaction. The two moderate H-bonds are oriented from the $NHCONHCH_3$ group at C_7 toward the main carbonyl group, whereas the Pi-Sulfur interaction appears between the sulfur atom in the side chain and the phenyl (B).

The binding modes of proflazepam for the TMD $\beta_2^+(A)/\alpha_1^-(B)$ interface are inconsistent in docking-poses for the two simulations (Figure IV.7 (b)). Accordingly, the differences between the two binding orientations were detected for the entire structure of proflazepam. The adequate binding pose resulting from the MD simulation was inserted more deeply into the binding-locus in such a way that the fused benzodiazepine rings and the substitute at N_1 have facial alignment with the β_2:TM_3 helix. The binding interactions after the MD simulation (Figure IV.8 (b), and Table IV.9) suggested the lack of the two H-bonds established with β_2Leu285 and β_2Arg269, As well as, the three Pi-Alkyl interactions created with β_2Met286, α_1Leu232, and α_1Pro233. Alternatively, α_1Ile228, β_2Phe289, α_1Pro233, and β_2Met286 were estimated to participate in six interactions with phenyl (B) and the substitutes at N_1 and C_7. The second hydroxyl group of $CH_2CHOHCH_2OH$ gave a strong H-bond to the main carboxyl function of α_1Ile228. The chlorine at C_7 was involved in three hydrophobic interactions, two are of the Alkyl-Alkyl type created with side chains of β_2Met286 and α_1Pro233 and one is of the Pi-Alkyl type created with the Pi-orbitals of β_2Phe289. Finally, the Pi-orbitals of phenyl (B) have participated through Pi-Pi stacked and Pi-Alkyl interactions with the side chains of β_2Phe289 and α_1Pro233, respectively.

Table IV.10. Detailed binding interactions resulting from the molecular dynamics simulation of Ro12-6377 and proflazepam (PLZ) at the TMD $\beta_2^+(\mathrm{C})/\alpha_1^-(\mathrm{D})$ interface.

TMD $\beta_2^+(\mathrm{C})/\alpha_1^-(\mathrm{D})$ interface								
Ligand Name	**Posit [a]**	**category**	**Type of interactions**	**From**	**From chemistry**	**To**	**To chemistry**	**Dist [b] (A°)**
Ro12-6377	2	Hydrogen Bond	Conventional Hydrogen Bond	C:ARG269:HH12	H-Donor	:*0:O	H-Acceptor	1,62
	7			:*0:H	H-Donor	D:MET236:O		2,55
	2’	Hydrogen Bond; Halogen	C-H Bond; Halogen (Fluorine)	C:MET286:HA	H-Donor; Halogen Acceptor	:*0:F	H-Acceptor; Halogen	2,53
	3	Hydrogen Bond	C-H Bond	:*0:H	H-Donor	D:ILE228:O	H-Acceptor	3,03
				:*0:H		D:ILE228:O		2,75
	7			:*0:H		D:MET236:O		2,81
	2’	Halogen	Halogen (Fluorine)	C:MET286:O	Halogen Acceptor	:*0:F	Halogen	3,32
	Ring B	Other	Pi-Sulfur	D:MET236:SD	Sulfur	:*0	Pi-Orbitals	5,26
		Hydrophobic	Pi-Pi Stacked	C:PHE289	Pi-Orbitals	:*0		4,95
	1		Alkyl	:*0:C	Alkyl	C:LEU285	Alkyl	4,68
				:*0:C		C:MET286		4,54
	7			:*0:C		C:VAL258		4,89
	Ring B		Pi-Alkyl	:*0	Pi-Orbitals	D:PRO233		4,88
	Ring C			:*0		C:MET286		5,12
PLZ	2	Hydrogen Bond	Conventional Hydrogen Bond	C:ASN265:HD22	H-Donor	:*0:O	H-Acceptor	2,20
	1			C:ARG269:HH12		:*0:O		1,42
	Internal [c]	Hydrogen Bond; Halogen	Conventional Hydrogen Bond;Halogen (Fluorine)	:*0:H	H-Donor; Halogen Acceptor	:*0:F	H-Acceptor; Halogen	1,81
	1	Hydrogen Bond	Conventional Hydrogen Bond	:*0:H	H-Donor	D:ILE228:O	H-Acceptor	1,65
	Internal [c]		C-H Bond	:*0:H	H-Donor	:*0:F		2,50
	Internal [c]	Halogen	Halogen (Fluorine)	:*0:N	Halogen Acceptor	:*0:F	Halogen	3,44
	Ring B	Other	Pi-Sulfur	D:MET236:SD	Sulfur	:*0	Pi-Orbitals	5,49
		Hydrophobic	Pi-Pi Stacked	C:PHE289	Pi-Orbitals	:*0		4,64
	7		Alkyl	:*0:CL	Alkyl	D:PRO233	Alkyl	4,81
				:*0:CL		D:MET236		5,26
			Pi-Alkyl	C:PHE289	Pi-Orbitals	:*0:CL		4,94
	Ring B			:*0		D:PRO233		4,44
	Ring C			:*0		D:PRO233		5,31
				:*0		D:LEU269		5,36

a: position of interaction, **b**: distance of interaction, **c**: internal interaction, ***0:** ligand structure

The binding modes of proflazepam for the TMD $\beta_2^+(\mathrm{C})/\alpha_1^-(\mathrm{D})$ interface are equivalents in the docking-poses for the two simulations (Figure IV.7 (c)). The most pronounced differences in the binding orientations have appeared for the substitutes groups at N_1, C_5, and C_7. The molecular docking and MD simulations overlap in that the residues β_2Asn265, β_2Phe289, α_1Ile228, α_1Pro233, and α_1Leu269 are essential parts of the $\beta_2^+(\mathrm{C})/$

α_1^-(D) modulation by proflazepam (Figure IV.8 (c), and Table IV.10). However, this does not necessarily mean that they have retained the same nature and orientation of their binding interactions shared with the proflazepam structure for both simulations, which is an observation common to all the binding sites studied in this paper. The binding pose after the MD simulation was stabilized by three intramolecular interactions orienting from the $CH_2CHOHCH_2OH$ group and N_1 towards the fluorine atom bond at $C_{2'}$. The four hydrophobic interactions created for both phenyls (B) and (C) and the H-bond linked α_1Ile 228 to the $CH_2CHOHCH_2OH$ group were preserved as the same. β_2Asn265 was reduced the two H-bonds established with N_4 and the fluorine atom at $C_{2'}$ to a single strong H-bond directed from its amide group towards the bound oxygen atom at C_2. α_1Ile228 loses the Halogen interaction established with the fluorine atom at $C_{2'}$ and conserves the H-bond accepted from the $CH_2CHOHCH_2OH$ group. α_1Pro233 tends to involve in hydrophobic interactions using its side chain. Thus, in addition to the two Pi-Alkyl interactions created with phenyl (B) and phenyl (C), it prefers to create an Alkyl-Alkyl interaction with the chlorine atom at C_7 rather than the strong H-bond created with the substitute at N_1. β_2Phe289 shows an additional hydrophobic interaction type Pi-Alkyl with the chlorine atom at C_7. Finally, three interactions with two new residues were observed for phenyl (B) and the substitutes at N_1 and C_7. The second hydroxyl group of $CH_2CHOHCH_2OH$ accepts a strong H-bond from the side chain of β_2Arg269, and, both the π-electron cloud of phenyl (B) and the chlorine at C_7 were formed, respectively, a Pi-Sulfur and Alkyl-Alkyl interactions with the side chain of α_1Met236.

At TMD γ_2^+/β_2^- interface, the binding modes of Ro12-6377 show significant differences for the two simulations. The adequate binding mode generated by MD simulation was inserted more deeply into the binding-locus, so that phenyl (B) is positioned between β_2:TM_1 and γ_2:TM_2 helices and phenyl (C) has facial alignment with the β_2:TM_2 and γ_2:TM_3 helices (Figure IV.5 (d)). MD simulation reduced the number of interactions suggested by molecular docking to less than half (Figure IV.6 (d) and Table IV.11). Else, suggested three interactions with γ_2Lys285 and β_2Leu268 instead of those established with β_2Pro228, β_2Leu223, γ_2Thr277, γ_2Thr281, γ_2Phe304, and γ_2Ala283. Both γ_2Lys285 and β_2Leu268 interact with the substitution group at C_7 through two H-bonds and one Alkyl-Alkyl interaction orienting, respectively, from their side chains towards the lone pairs of the oxygen atom and the methyl group. On the other hand, the number of interactions created with γ_2Asp297 and β_2Gln224 is reduced by the factor of one H-bond for each. In contrast, interactions with γ_2Ser280, γ_2Arg284, γ_2Val300, and β_2Ile264 residues are kept identical between the two simulations.

Table IV.11. Detailed binding interactions resulting from the molecular dynamics simulation of Ro12-6377 and proflazepam (PLZ) at the TMD γ_2^+/β_2^- interface.

TMD γ_2^+/β_2^- interface								
Ligand Name	**Posit** [a]	**category**	**Type of interactions**	**From**	**From chemistry**	**To**	**To chemistry**	**Dist** [b] **(A°)**
Ro12-6377	7	Hydrogen Bond	Conventional Hydrogen Bond	E:LYS285:HZ2	H-Donor	:*0:O	H-Acceptor	2,09
				:*0:H		A:GLN224:O		2,94
				:*0:H		A:GLN224:O		1,74
			C-H Bond	E:LYS285:HE2		:*0:O		2,73
	2			E:ASP297:HA		:*0:O		2,40
	1			:*0:H		E:SER280:O		2,96
		Hydrophobic	Alkyl	:*0:C	Alkyl	E:ARG284	Alkyl	4,71
				:*0:C		E:VAL300		4,13
	7			:*0:C		A:ILE264		4,75
				:*0:C		A:LEU268		4,66
PLZ	1	Hydrogen Bond	Conventional Hydrogen Bond	:*0:H	H-Donor	E:THR281:OG1	H-Acceptor	1,81
				:*0:H		A:GLN224:O		2,06
			C-H Bond	:*0:H		E:THR277:O		2,86
				:*0:H		E:THR277:O		2,94
				:*0:H		E:SER280:OG		2,51
				:*0:H		E:THR281:OG1		2,60
	Ring C	Other	Pi-Sulfur	A:MET227:SD	Sulfur	:*0	Pi-Orbitals	5,18
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	A:LEU223	Alkyl	4,37
				:*0:CL		E:VAL300		4,07
	Ring B		Pi-Alkyl	:*0	Pi-Orbitals	E:VAL300		5,15

a: position of interaction, **b**: distance of interaction, ***0:** ligand structure

Unlike Ro12-6377, the binding modes of proflazepam for the TMD γ_2^+/β_2^- interface are equivalents for the two simulations (Figure IV.7 (d)). However, notable differences in the binding orientations of the substitution groups at C_5 and C_7 have appeared. As previously noted for Ro12-6377, after MD simulation, the number of interactions was almost reduced to half ((Figure IV.8 (d) and Table IV.11). All the interactions established with β_2Pro228, γ_2Arg284, γ2Asp297, β_2Leu231, and γ_2Phe304 have vanished, and instead of them, Pi-Sulfur interaction was observed between the sulfur atom at β_2Met227 side chain and the Pi-Orbitals of phenyl (C). The three H-bonds linked β_2Gln224 to the $CH_2CHOHCH_2OH$ group were reduced into one moderate H-acceptor bond oriented from the second hydroxyl group toward the main carbonyl group of β_2Gln224. The Halogen interaction established between β_2Leu223 and the fluorine at $C_{2'}$ vanished, and the H-bond connected C_2 to γ_2Thr277 was replaced by two moderate H-bonds given from the $CH_2CHOHCH_2OH$ group to the main carbonyl group of the same residue. γ_2Val300 involved in additional hydrophobic interaction type Alkyl-Alkyl with the chlorine atom at C_7. Finally, interactions with γ_2Thr281 and γ_2Ser280 residues are kept identical between the two simulations.

Overall, the MD results do not agree with those observed in the molecular docking analysis that showed interactions with α_1His102. As mentioned earlier, this residue ensures the recognition of classical BDZ by the ECD interface. The MD results coincided well with those noted in the molecular docking analysis that indicated the features of Ro12-6377 and proflazepam to directly connect the pore-lining residues. The bulky structure of Ro12-6377 and proflazepam is the key factor in the deep penetration towards the TM_2 helices. In particular, the long backbone skeleton of the $NHCONHCH_3$ and $CH_2CHOHCH_2OH$ groups; where most interactions with the TM_2 helices have been observed. At TMD $\beta_2^+(C)/\alpha_1^-(D)$, the MD simulation generated two interactions for Ro12-6377 with β_2:C:TM_2:Arg269 and β_2:C:TM_2:Val 258 residues. In addition, at TMD γ_2^+/β_2^- interface, generated three interactions with β_2:A:TM_2:Leu268 and γ_2:TM_2:Lys285 instead of those previously observed with TM_2:Thr277, TM_2:Thr281, and TM_2:Ala283 of subunit γ_2. Likewise, at TMD $\beta_2^+(C)/\alpha_1^-(D)$ interface, it suggested an additional interaction for the proflazepam with β_2:C:TM_2:Arg269, and at TMD γ_2^+/β_2^- interface, suggested a lack of interactions with γ_2:TM_2:Arg 284.

3.3 Pharmacokinetic and drug-likeness prediction

As shown in Appendix H, all BDZ compounds are estimated to have high gastrointestinal absorption and nearly all are able to cross the blood–brain barrier. Almost all BDZ molecules are not affected by p-gp efflux pump. Besides, all tested compounds respect Lipinski, Veber, Egan, Ghose, and Muegge rules.

Moreover, Ro12-6377 and proflazepam can successfully penetrate the blood–brain barrier. They are also estimated to be actively effluxes by P-glycoprotein transporter and to act as non-inhibitors towards CYP isoenzymes, except for CYP2D6 in proflazepam. Interestingly, these two compounds share one favourable characteristic in which they do not inhibit CYP2C19 and CYP3A4 enzymes that might be responsible for the hepatic clearance or the formation of active metabolites of BDZ. Indeed, it is well established that the BDZs are primarily metabolized via CYP2C19 and CYP3A4 and the inhibition of them can cause drug-drug interactions (DDIs) [51,52].

3.4 QSAR analysis

The QSAR study was carried out on the BDZ data set previously investigated through molecular docking simulation (Table IV.1). The data set includes a total of 65 compounds that were used to generate the PLS regression model and evaluate its performance. Using 5-fold cross-validation, we randomly split the 65 observations into two sets, a training set containing 52 of the data points, and a test set containing the remaining 13 observations. The molecular descriptors are coded into the term of variables: $x_1=q_{N1}$, $x_2=q_{C3}$, $x_3=q_{N4}$, $x_4=q_{C6}$, $x_5=q_{C7}$, $x_6=q_{C8}$, $x_7=q_{C9}$, $x_8=q_{C2'}$, $x_9=q_{C6'}$, $x_{10}=\pi_{C7}$, $x_{11}=\pi_{C2'}$, x_{12}=HA, x_{13}=HD, x_{14}=DM, x_{15}=Log (P), $x_{16}=M_R$, $x_{17}=F_T$. The numerical values are summarized in Appendix I.

Table IV.12 reported the observed and predicted biological activities with the corresponding studentized deleted residual (r_i^*) and the leverage (h_{ii}) values of the training and test set compounds of a sample among the 10 000 simulations generated in this study. In this case, six outliers (marked in bold) were detected for the activity response. Four observations have r_i^* greater than the threshold |2| and three observations with large h_{ii} values. Consequently, the training and test sets were reduced to 47 and 12 observations, respectively.

Table IV.12. Studentized deleted residual values ($\mathbf{r_i^*}$) and the leverage values (h_{ii}) of the training and test set compounds.

Comp N°	**Y_i norm**	**Y_{hat}**	**r_i^***	**h_{ii}**	**Comp N°**	**Y_i norm**	**Y_{hat}**	**r_i^***	**h_{ii}**
				Training set					
1	8.2954	8.1891	0.1713	0.2220	**35**	8.2954	8.5488	-0.4167	0.2493
3	8.4000	9.1331	-1.1372	0.1352	**36**	9.2113	9.1223	0.1354	0.1266
4	8.4393	9.7341	-2.0056	0.0804	**37**	9.1982	9.5757	-0.5616	0.0800
6	9.3159	9.8727	-0.9261	0.2553	**39**	8.3346	8.8803	-0.8621	0.1764
7	9.6430	9.2510	0.5873	0.0922	**40**	8.1253	8.9133	-1.2514	0.1698
8	9.6823	10.0519	-0.5490	0.0775	**43**	8.9627	9.6633	-1.4642	**0.5146**
9	9.7346	9.7370	-0.0036	0.1082	**44**	8.5309	9.4693	-1.5789	0.2455
10	9.8393	9.5026	0.4970	0.0669	**45**	11.0692	10.9866	0.1345	0.2361
11	9.9701	9.8377	0.1953	0.0698	**47**	10.6636	9.8739	1.2419	0.1540
13	10.1271	9.7364	0.5867	0.0962	**48**	9.5645	10.1277	-0.8436	0.0843
14	10.3234	10.4628	-0.2098	0.1072	**49**	10.1926	9.7383	0.6890	0.1121
15	10.5851	9.9585	0.9361	0.0763	**50**	10.8206	8.6790	**4.5068**	**0.3328**
16	10.8075	11.0797	-0.4154	0.1287	**51**	8.8580	9.5543	-1.0520	0.0922
17	10.8206	10.0065	1.2313	0.0859	**52**	10.1010	10.2745	-0.2602	0.1000
19	11.0561	10.9483	0.1691	0.1788	**53**	10.2711	9.9006	0.5519	0.0827
20	11.3832	10.6639	1.0698	0.0625	**54**	11.0169	10.7783	0.3697	0.1553
21	11.4356	11.6426	-0.3182	0.1425	**55**	8.4916	9.3474	-1.2840	0.0683
22	11.4356	11.1475	0.4431	0.1420	**56**	8.9234	9.4633	-0.7967	0.0581
23	11.5402	10.0976	**2.2560**	0.0778	**58**	10.1010	9.8097	0.4441	0.1263
25	10.1664	10.7639	-1.0675	**0.3503**	**59**	9.3421	9.2794	0.0929	0.0785
27	10.5066	10.2804	0.3321	0.0600	**60**	8.5309	9.3314	-1.2208	0.1014
29	10.9776	10.8453	0.2004	0.1183	**61**	9.6823	9.8128	-0.1930	0.0745
30	11.1477	11.5546	-0.6271	0.1413	**62**	11.0430	10.4101	0.9384	0.0621
31	10.6636	11.4656	-1.2521	0.1410	**63**	10.2711	10.0263	0.3632	0.0791
32	10.4542	9.7895	1.0285	0.1353	**64**	9.7215	9.7755	-0.0789	0.0547
33	11.3309	9.9877	**2.2297**	0.1836	**65**	9.5253	9.8470	-0.4706	0.0503

Table IV.12. Continued

Test set									
2	8.5167	8.6018	-1.2800	0.3654	**34**	9.5531	9.1409	0.3076	0.3440
5	9.3272	8.8671	0.5421	0.3889	**38**	9.2475	9.0331	-0.5152	0.9222
12	10.2041	10.2062	-0.3840	0.5716	**41**	10.1111	10.3151	**-2.3002**	0.4636
18	11.1873	9.7623	1.8961	0.5383	**42**	9.7922	9.4752	0.8742	0.8557
24	11.0146	10.1997	0.7054	0.3238	**46**	11.8516	12.0887	-1.4227	0.9560
26	10.8020	10.4640	-0.1934	0.2681	**57**	10.0314	10.2422	-0.7454	0.4415
28	11.6789	10.8360	1.6756	0.5610					

Figures IV.9 and IV.10 show the results of the 10 000 simulations obtained for the selected biological activity. According to Figure IV.9, the best subsets of variables are those with the highest probability of occurrence according to the BIC criterion. As a result, the best subset of variables to select is that containing six variables. The intercept is systematically included whatever the model. The next step is determining the variables retained in the subset. According to Figure IV.10, it appears that the six best variables are x_3, x_{11}, x_{14}, x_{10}, x_{15}, and x_{16}. This ranking order is a function of their probability of occurrence. By combining the results of Figures IV.9 and IV.10, the best variable subset contains: q_{N4}, $\pi_{C2'}$, DM, π_{C7}, log (P), and M_R (Table IV.13).

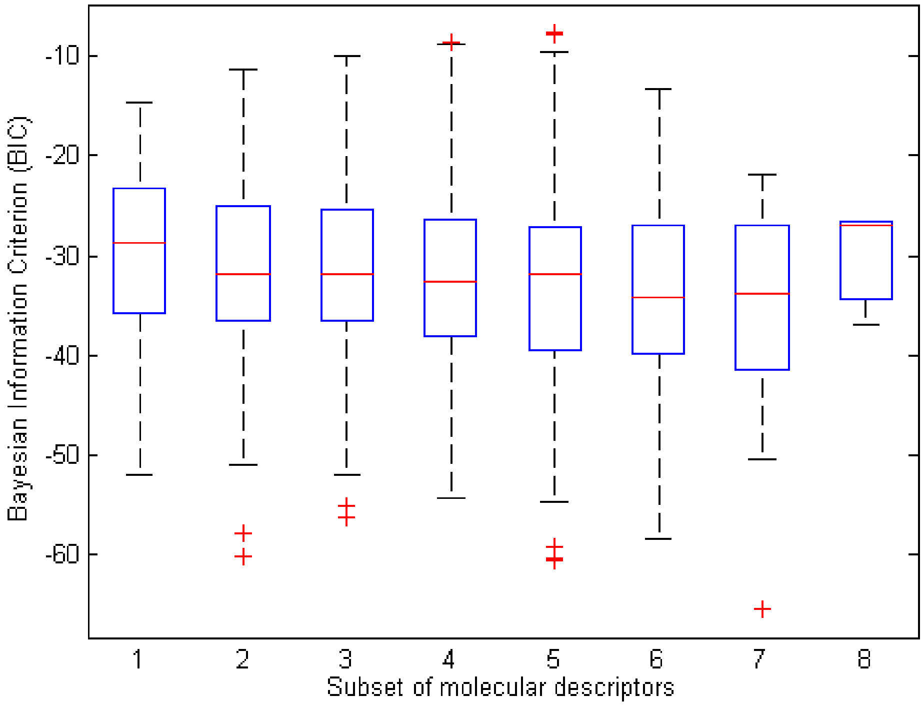


Figure IV.9. Box plots of the distribution of the Bayesian Information Criterion (BIC) by number of molecular descriptors.

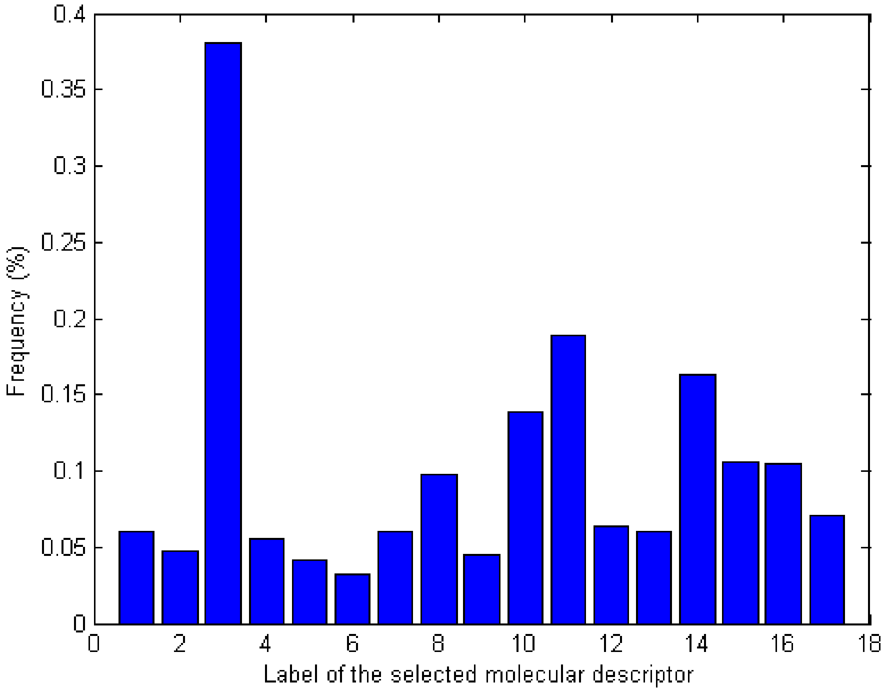


Figure IV.10. Probability of occurrence of selected molecular descriptors.

Figure IV.11 depict the plots of the internal indicators R^2, R^2_{adj}, and F-statistics versus their Probabilities of occurrence derived for the 10 000 training sets. The plots show a bimodal distribution that is symmetrically distributed around two mean values. The mean value with the highest probability is picked in those plots based on where the majority of the variables tend to be distributed. Figure IV.12 depict the plots of the external indicators Q^2_{loo} and Q^2_{F3} versus their Probabilities of occurrence derived for the 10 000 test sets. The plots of $10^4 Q^2_{loo}$ and $10^4 Q^2_{F3}$ demonstrate a monomodal distribution that is symmetrically distributed around its mean. To estimate the Q^2_{loo} and Q^2_{F3} coefficients, the mean values with the greatest probability are used.

Table IV.13. The optimal variables to generate the PLS model.

Comp N°	q_{N4}	π_7	$\pi_{2'}$	MD (debye)	Log P	M_R
1	-0,608	-1,230	0,000	6,301	1,643	79,708
2	-0,618	-1,230	0,000	6,583	1,779	76,592
3	-0,590	-0,570	0,000	2,121	2,328	80,729
4	-0,620	0,000	0,000	5,061	2,608	71,891
5	-0,655	-1,340	0,140	6,721	2,123	82,211
6	-0,698	-1,230	0,710	6,791	2,383	81,396
7	-0,620	-0,650	0,000	3,137	2,320	78,475
8	-0,613	0,140	0,000	3,369	2,751	72,108
9	-0,619	1,020	0,000	5,662	3,566	81,533
10	-0,634	-0,570	0,140	3,218	2,471	80,945
11	-0,631	0,820	0,000	4,936	3,345	81,578
12	-0,657	0,000	0,140	5,752	2,751	72,108
13	-0,670	-0,550	0,140	3,274	2,308	82,510
14	-0,614	0,880	0,000	1,909	3,486	77,865
15	-0,601	0,710	0,000	3,228	3,076	79,812
16	-0,679	0,710	0,710	3,462	4,284	89,421
17	-0,659	0,460	0,140	4,057	4,085	82,395
18	-0,622	-0,280	0,140	3,209	2,555	81,544
19	-0,673	-0,280	0,880	3,692	3,426	84,185
20	-0,665	0,710	0,140	4,140	3,355	76,912
21	-0,696	0,710	0,710	3,938	3,816	81,501
22	-0,696	-0,280	0,710	2,488	3,152	83,016
23	-0,649	-0,280	0,140	2,378	2,691	78,428
24	-0,637	0,140	0,140	4,373	2,757	75,440
25	-0,597	0,140	0,000	3,338	2,615	57,224
26	-0,654	0,140	0,140	4,242	2,893	72,324
27	-0,624	0,710	0,000	3,257	3,212	76,696
28	-0,689	0,710	0,140	3,645	3,497	77,129
29	-0,700	0,710	0,140	3,646	3,361	80,245
30	-0,681	0,710	0,710	3,557	3,959	81,717
31	-0,710	0,710	0,710	3,486	4,420	86,306
32	-0,615	-0,280	0,000	1,381	2,548	78,212
33	-0,663	-0,280	0,710	3,365	3,016	97,151
34	-0,665	-1,230	0,140	6,893	1,786	79,924
35	-0,655	-1,030	0,140	4,302	1,743	93,660
36	-0,573	0,710	0,000	4,083	4,031	85,262
37	-0,607	0,710	0,000	3,104	3,304	87,392
38	-0,584	0,710	0,000	3,049	3,857	91,754
39	-0,569	-0,280	0,000	1,785	2,475	87,181
40	-0,587	0,710	0,000	3,222	4,130	93,618
41	-0,642	0,710	0,140	3,078	2,529	86,321
42	-0,640	0,710	0,140	5,437	1,843	98,979
43	-0,634	0,710	0,140	5,140	1,898	92,283
44	-0,657	-0,280	0,710	4,125	4,070	99,938
45	-0,736	0,710	0,140	3,539	3,923	81,406
46	-0,766	-0,280	0,710	2,404	3,721	87,510
47	-0,698	-0,280	0,140	3,063	3,123	86,038
48	-0,660	0,710	0,000	2,958	3,645	84,306
49	-0,609	0,710	0,000	4,751	2,787	81,010
50	-0,541	0,710	0,140	6,653	4,124	84,764
51	-0,647	0,560	0,000	5,677	3,635	81,974
52	-0,674	0,000	0,140	5,429	2,893	72,324
53	-0,648	0,000	0,140	5,768	2,615	75,224
54	-0,666	0,000	0,710	5,557	3,076	79,812
55	-0,605	0,000	0,000	4,500	3,212	76,696
56	-0,635	0,000	0,140	5,583	3,219	80,028
57	-0,684	0,000	0,140	4,474	3,497	77,129
58	-0,676	0,000	0,140	6,296	3,264	77,149
59	-0,620	0,000	0,140	5,954	3,219	80,028
60	-0,632	0,000	0,140	4,116	3,823	84,833
61	-0,611	0,710	0,000	2,420	3,680	84,617
62	-0,662	0,710	0,140	3,363	3,959	81,717
63	-0,668	0,560	0,140	5,148	3,868	81,954
64	-0,614	0,710	0,000	3,589	3,816	81,501
65	-0,617	0,710	0,000	3,336	3,725	81,737

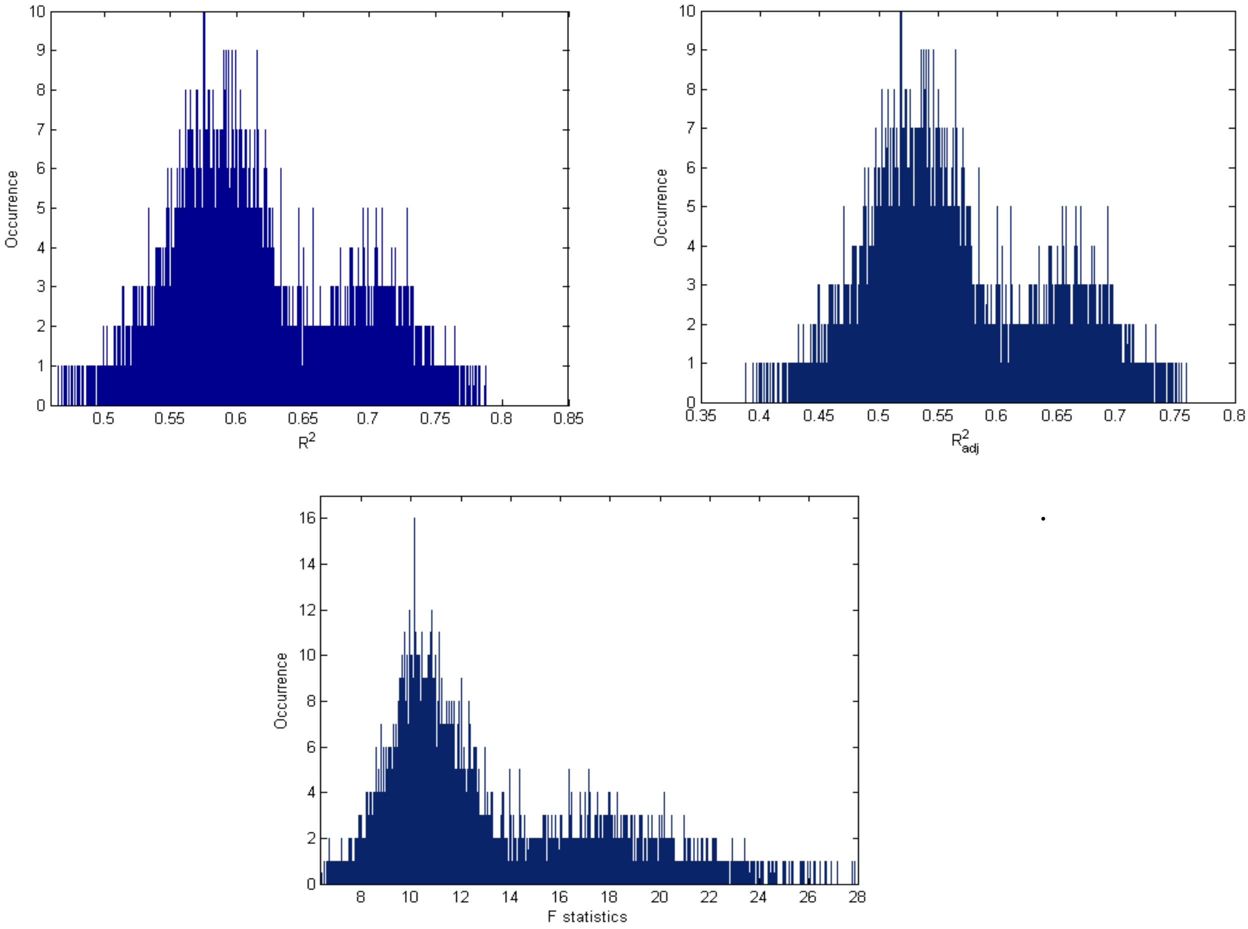


Figure IV.11. Distribution of 10^4R^2, $10^4R^2_{adj}$, and $10^4F_{statistics}$ as function of their probability of occurrences.

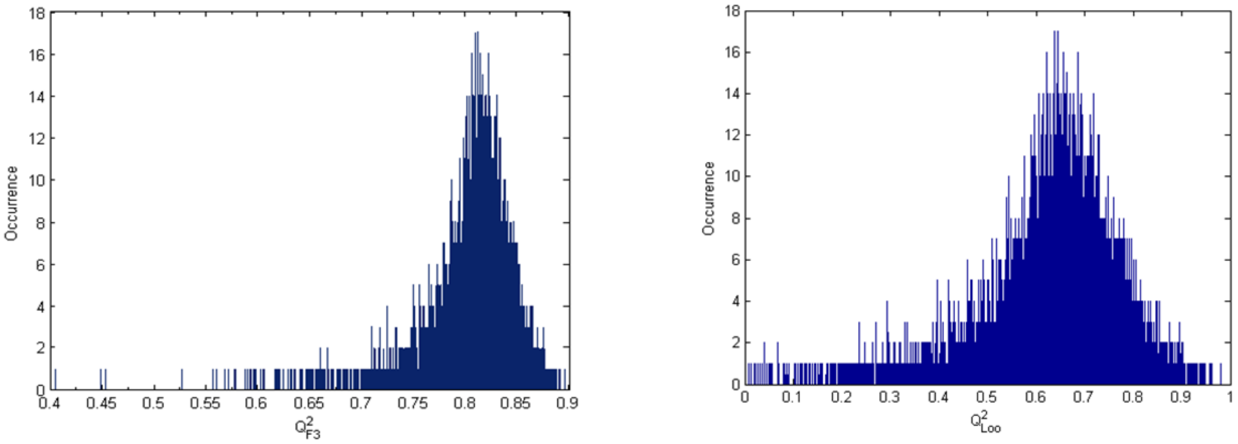


Figure IV.12. Distribution of $\mathbf{10^4Q^2_{F3}}$ and $\mathbf{10^4Q^2_{Loo}}$ as function of their probability of occurrences.

In order to highlight the weight of each molecular descriptor, the PLS regression models are written with scaled variables. The standardized regression coefficient value of each descriptor highlights the relative importance of the descriptors in determination of biological activity of the compounds.

The final QSAR model with 95% confidence interval of the regression coefficient is:

$$y' = \underset{(\pm 0.3427)}{10.0747} - \underset{(\pm 0.0908)}{0.3546\,x'_3} + \underset{(\pm 0.0679)}{0.4694\,x'_{10}} + \underset{(\pm 0.0665)}{0.4460\,x'_{11}} - \underset{(\pm 0.0676)}{0.2381\,x'_{14}} - \underset{(\pm 0.0882)}{0.1675\,x'_{15}} - \underset{(\pm 0.0496)}{0.3665\,x'_{16}}$$

Where $y' = \log(1/C)_{obs}/s_y$ and $x'_j = x_j/s_{x_j}$. s_y: The standard deviation corresponding to the biological response. s_{xj}: The standard deviation corresponding to the *jth* descriptor.

Table IV.14 summarizes the statistical indicators used for internal and external validation. According to the goodness of fit statistics, 63.2% of the variability in BDZ activity around its mean is explained by the PLS regression equation. The quality of models can be judged and compared based on the R^2_{adj} values. The F-statistics reveal the significance of the PLS regression equations. The obtained p-value shows that the model is statistically highly significant at 95%. Moreover, it is well known that cross validation is useful for overcoming the problem of overfitting [22]. This problem refers to a situation when the model requires more information than the data can provide. Indeed, in our case the difference between R^2 and Q^2_{loo} is much less than the threshold of 0.30, confirming that the PLS regression models are not overfitted. Additionally, our model exhibits a high value of Q^2_{F3} (0.813). This result confirms that the resulting QSAR model has good external predictability and robustness.

Table IV.14. Quality and validation metrics.

Model	**Goodness of fit**				**Goodness of prediction**	
	$\mathbf{R^2}$	$\mathbf{R^2_{adj}}$	**F**	$\mathbf{p_{value}}$	$\mathbf{Q^2_{loo}}$	$\mathbf{Q^2_{F3}}$
Y ($\mathbf{n_{tr}}$=52 ; $\mathbf{n_{ts}}$=13)	0.632	0.584	12.806	6.2050e-07	0.639	0.813

n_{tr} is the training set used for building the PLS regression equation and n_{ts} is the test set used to verify a model's predictive ability for new untested molecules.

Molecular descriptors with positive regression coefficients indicate positive correlations with observed activity. Therefore, their increase will improve the $GABA_A$/BDZ response. Our model demonstrates a positive contribution from the hydrophobic constants of the substitutes at C_7 and $C_{2'}$. As well as, demonstrates a negative contribution from molecular lipophilicity (log (P)), molecular polarity (DM), molecular size (M_R), and net charge of N_4 (q_{N4}).

According to the standardized regression coefficient values, the hydrophobicity in positions 7 and 2' are the main properties for determining the biological activities of the studied compounds. This result is in good agreement with previous reports that the 7 and 2' positions are, respectively, the first and second key positions in the BDZ structures to influence binding with $GABA_A$Rs [5]. Otherwise, the lipophilic behavior shows the least effect on activity. Molecular lipophilicity is directly influenced by both molecular hydrophobicity and polarity [53]. In our case, the positive effects of hydrophobicity at positions 7 and 2’ on the observed activity seem to be two times more significant than the negative effect of molecular polarity. The lipophilic feature plays a pivotal role in understanding the pharmacokinetics parameters, pharmacodynamics, and toxicological profile of drugs. Besides, exhibits an important influence on host-guest interaction and drug binding affinity [34]. BDZ derivatives are relatively weak bases with highly lipophilic characters. At physiological pH, this lipophilicity explains their strong binding to plasma proteins (70-99%) as well as their high penetration through the blood-brain barrier [54].

N_4 net charge and molar refractivity exhibit about twice the negative influence of molecular lipophilicity on activity. Reducing the negative charge of N_4 is important to avoid the formation of water-unstable BDZ salts. The imine group (N_4) is the most basic nitrogen in the classical BDZ structure. Since the amide group at positions 1 and 2 has a non-basic character, the lone pair of N_4 can easily protonate when placed in a strongly acidic environment. Thus, it leads to the formation of BDZ salts (iminium ion). Unfortunately, the salts of strong acids are unstable in an aqueous medium and are therefore undergoing sequential hydrolysis of the imine and amide groups, respectively. The imine hydrolysis reaction is reversible. In contrast, hydrolysis of amide leads to the formation of inactive products. Consequently, eliminate activity towards $GABA_A$ receptors [55]. Most of the classical BDZ agents we study here provide unstable salts and are relatively water-insoluble than drugs formed from heterocyclic BDZs. In the latter, the amide group is protected in the form of heterocyclic groups such as ImidazoBDZs, TriazoloBDZs, etc. Hence, hydrolysis of the amide does not occur and the reaction does not lead to inactive products. To improve the solubility of classical BDZ salts in water-soluble injections, in addition to water, it is necessary to use co-solvents such as PEG 400, propylene glycol, 10% ethyl alcohol, 2% benzyl alcohol, etc [56].

Our results support those of *D.J. Maddalena* and *S. S. So* [5,6] by asserting that increased the hydrophobicity at position 7 is necessary to obtain highly potent BDZ analogues. Besides, it provides complementary insights on how hydrophobicity at $C_{2'}$ position, the net charge of N_4, and molar refractivity, polarity, and lipophilicity of the entire molecules would be ameliorated to achieve the optimal activity. In particular, the challenge is to provide structures meeting simultaneously the requirements of low lipophilicity/low polarity, since they are naturally anti-correlated. Here, our model suggests that the influence of molecular polarity on activity outweighs that of molecular lipophilicity by 7%. Accordingly, efforts should be devoted primarily to reducing molecular polarity rather than molecular lipophilicity.

After excluding outliers, a comprehensive analysis of the detailed binding interactions with the four binding interfaces was performed for the training and test data sets. Subsequently, the hydrophobic interactions with C_7 and $C_{2'}$ substitutes and the electrostatic interactions with N_4 were selected and collected in Appendices J, K, L, and M. As can be seen, the richness of the C_7 position by hydrophobic interactions is attributed in most cases to the presence of chlorine atoms that exhibited Alkyl-Alkyl or Pi-Alkyl interactions with neighboring residues. Significantly, the response of chlorine to establish hydrophobic interactions with the four binding interfaces is different between ECD and TMD. At ECD α_1^+/γ_2^- interface, it tends to act as an acceptor of interactions, while at the three TMD interfaces it acts as a generator of interactions. $C_{2'}$ position shows a total lack of hydrophobic interactions with the four binding interfaces, except for the case of the CF_3 group in Ro05-3590 and the chlorine atom in Ro05-4608. This deficiency maybe because most of the compounds of our data set contains small substituents (H and F) which mainly tend to interact electrostatically rather than hydrophobically. The N_4 atom tends to react as a hydrogen-acceptor from neighboring residues. Importantly, in the ECD α_1^+/γ_2^- interface, all interactions observed for N_4 are received from α_1His 102 side chain. As reported previously in the molecular docking section, this residue is known to be important in the recognition of classical BDZ. This result does not correlate with previous analyses that support cationic interactions at this position [57].

4 Conclusion

In this investigation, a combination of in silico approaches including molecular docking/dynamic simulations, and QSAR analysis have been performed to achieve two purposes: elucidate the binding mechanism by which a dataset of [^{3}H]diazepam derivatives allosterically modulates GABA$_A$ receptor $\alpha_1\beta_2\gamma_2$ subtypes and identify the structural details that contribute to ameliorate the $\alpha_1\beta_2\gamma_2$/BDZ response. Examination of binding affinities revealed that the known ECD is the target for the majority of classical benzodiazepines. However, the tendency of the remainder to bind mainly on the binding interfaces included in TMD, or in some cases, to act on both ECD and TMD binding sites simultaneously, cannot be overlooked. This result opens the way for further studies that may combine binding in these binding sites with diversity in the activities of benzodiazepines. Binding affinities-based screening identified Ro12-6377 and proflazepam as the best modulators for the four binding interfaces. By monitoring the dynamic behaviors over a time period of 1000 ps, the two modulators were observed to have equivalent stability within the four binding sites under study. Binding modes after MD simulation were altered from the structures generated by molecular docking. Thus, several differences in the binding interactions with key residues were detected between the two simulations. Importantly, interactions with pore-lining residues have been suggested for both modulators. The combination of ADME prediction/Drug-likeness prediction shows their good pharmacokinetic properties as well as their compliance with all drug-likeness rules. Furthermore, the developed QSAR model yielded satisfactory statistical results that explain 63.2% of the variability in benzodiazepine activity. Its stability and predictive power were ensured based on internal and external validation indicators: R^2_{adj}=0.584, F=12.806; P_{value}=6.2050e-07, Q^2_{loo}=0.639, and Q^2_{F3}=0.813. The model equation demonstrates a positive contribution from the hydrophobicity of the substitutes at C_7 and $C_{2'}$. As well as, demonstrates a negative contribution from molecular lipophilicity, molecular polarity, molecular size, and net charge of N_4. The model results agree well with previous findings indicating that the increase in hydrophobicity at 7-position mainly contributes to the enhancement of BDZ activity. Finally, the combination of the results of both methods: QSAR and molecular docking show that the hydrophobic interactions at the 7-position are mostly attributed to the substitutions of chlorine atoms. The latter tends to act as an acceptor of interactions in the ECD binding interfaces and as a generator of interactions in the three TMD binding interfaces.

Bibliography

[1] Y. Xu, Z. He, H. Liu, Y. Chen, Y. Gao, S. Zhang, M. Wang, X. Lu, C. Wang, Z. Zhao, Y. Liu, J. Zhao, Y. Yu, M. Yang, 3D-QSAR, molecular docking, and molecular dynamics simulation study of thieno[3,2-: B] pyrrole-5-carboxamide derivatives as LSD1 inhibitors, RSC Adv. 10 (2020) 6927–6943. https://doi.org/10.1039/c9ra10085g.

[2] H. Alonso, A.A. Bliznyuk, J.E. Gready, Combining docking and molecular dynamic simulations in drug design, Med. Res. Rev. 26 (2006) 531–568. https://doi.org/10.1002/med.20067.

[3] J. Verma, V. Khedkar, E. Coutinho, 3D-QSAR in Drug Design - A Review, Curr. Top. Med. Chem. 10 (2010) 95–115. https://doi.org/10.2174/156802610790232260.

[4] K. Roy, On some aspects of validation of predictive quantitative structure-activity relationship models, Expert Opin. Drug Discov. 2 (2007) 1567–1577. https://doi.org/10.1517/17460441.2.12.1567.

[5] D.J. Maddalena, G.A.R. Johnston, Prediction of Receptor Properties and Binding Affinity of Ligands to Benzodiazepine/GABAA Receptors Using Artificial Neural Networks, J. Med. Chem. 38 (1995) 715–724. https://doi.org/10.1021/jm00004a017.

[6] S.S. So, M. Karplus, Genetic neural networks for quantitative structure-activity relationships: Improvements and application of benzodiazepine affinity for benzodiazepine/GABAA receptors, J. Med. Chem. 39 (1996) 5246–5256. https://doi.org/10.1021/jm960536o.

[7] J.J. Kim, A. Gharpure, J. Teng, Y. Zhuang, R.J. Howard, S. Zhu, C.M. Noviello, R.M. Walsh, E. Lindahl, R.E. Hibbs, Shared structural mechanisms of general anaesthetics and benzodiazepines, Nature. 585 (2020) 303–308. https://doi.org/10.1038/s41586-020-2654-5.

[8] D. Hadjipavlou-Litina, C. Hansch, Quantitative Structure–Activity Relationships of the Benzodiazepines. A Review and Reevaluation, Chem. Rev. 94 (1994) 1483–1505. https://doi.org/10.1021/cr00030a002.

[9] A. Micheli, A. Sperduti, A. Starita, A.M. Bianucci, Analysis of the Internal Representations Developed by Neural Networks for Structures Applied to Quantitative Structure-Activity Relationship Studies of Benzodiazepines, J. Chem. Inf. Comput. Sci. 41 (2001) 202–218. https://doi.org/10.1021/ci9903399.

[10] HyperChem, (2008).

[11] M. J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, G.A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B.G. Janesko, R. Gomperts, B. Mennucci, H.P. Hratchian, J. V.

Ortiz, A.F. Izmaylov, J.L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V.G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J.A. Montgomery, Jr., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, J.M. Millam, M. Klene, C. Adamo, R. Ca, D.J. Fox, Gaussian 09W, (2010).

[12] L.E. Chirlian, M.M. Francl, Atomic charges derived from electrostatic potentials: A detailed study, J. Comput. Chem. 8 (1987) 894–905. https://doi.org/10.1002/jcc.540080616.

[13] S.P. Gupta, R.N. Saha, V. Mulchandani, Quantitative structure–activity relationship studies on benzodiazepine receptor binding: Recognition of active sites in receptor and modelling of interaction, J. Mol. Recognit. 5 (1993) 75–80. https://doi.org/10.1002/JMR.300050205.

[14] J.O. Rawlings, S.G. Pantula, D.A. Dickey, Applied Regression Analysis: A Research Tool, Second Edition, 1999.

[15] J.O. Rawlings, S.G. Pantula, D.A. Dickey, Applied Regression Analysis : A Research Tool, Second Edi, NY: Springer New York., New York, 1998.

[16] S. De Jong, SIMPLS: an alternative approach to partial least squares regression, Chemom. Intell. Lab. Syst. 18 (1993) 251–263. https://doi.org/10.1016/0169-7439(93)85002-X.

[17] T.S. Wiens, B.C. Dale, M.S. Boyce, G.P. Kershaw, Three way k-fold cross-validation of resource selection functions, Ecol. Modell. 212 (2008) 244–255. https://doi.org/10.1016/j.ecolmodel.2007.10.005.

[18] G. James, D. Witten, T. Hastie, Robert Tibshirani, An Introduction to Statistical Learning - with Applications in R, New York: springer, 2013.

[19] G. Schwarz, Estimating the Dimension of a Model, Ann. Stat. 6 (1978) 461–464. http://www.jstor.org/stable/2958889.

[20] O. Renaud, M.P. Victoria-Feser, A robust coefficient of determination for regression, J. Stat. Plan. Inference. 140 (2010) 1852–1862. https://doi.org/10.1016/j.jspi.2010.01.008.

[21] O. Harel, The estimation of R2 and adjusted R2 in incomplete data sets using multiple imputation, J. Appl. Stat. 36 (2009) 1109–1118. https://doi.org/10.1080/02664760802553000.

[22] R. Veerasamy, H. Rajak, A. Jain, S. Sivadasan, C.P. Varghese, R.K. Agrawal, Validation of QSAR Models - Strategies and Importance, Int. J. Drug Des.

Disocovery. 2 (2011) 511–519.

[23] V. Consonni, D. Ballabio, R. Todeschini, Comments on the definition of the Q2 parameter for QSAR validation, J. Chem. Inf. Model. 49 (2009) 1669–1678. https://doi.org/10.1021/ci900115y.

[24] V. Consonni, D. Ballabio, R. Todeschini, Evaluation of model predictive ability by external validation techniques, J. Chemom. 24 (2010) 194–201. https://doi.org/10.1002/cem.1290.

[25] V. Consonni, R. Todeschini, D. Ballabio, F. Grisoni, On the Misleading Use of QF32 for QSAR Model Comparison, Mol. Inform. 38 (2019) 2–6. https://doi.org/10.1002/minf.201800029.

[26] Molecular Operating Environment (MOE), (2014).

[27] BIOVIA, Dassault Systèmes, Discovery Studio Visualizer, (2020).

[28] S.D. Bond, B.J. Leimkuhler, B.B. Laird, The Nosé-Poincaré Method for Constant Temperature Molecular Dynamics, J. Comput. Phys. 151 (1999) 114–134. https://doi.org/10.1006/jcph.1998.6171.

[29] T.A. Halgren, Performance of MMFF94*, J. Comput. Chem. 17 (1996) 490–519. http://journals.wiley.com/jcc.

[30] A. Belkadi, S. Kenouche, N. Melkemi, I. Daoud, R. Djebaili, K-means clustering analysis, ADME/pharmacokinetic prediction, MEP, and molecular docking studies of potential cytotoxic agents, Struct. Chem. 32 (2021) 2235–2249. https://doi.org/10.1007/s11224-021-01796-w.

[31] M.A.F. Nasution, E.P. Toepak, A.H. Alkaff, U.S.F. Tambunan, Flexible docking-based molecular dynamics simulation of natural product compounds and Ebola virus Nucleocapsid (EBOV NP): a computational approach to discover new drug for combating Ebola, BMC Bioinformatics. 19 (2018) 137–176. https://doi.org/10.1186/s12859-018-2387-8.

[32] K.E. Hevener, W. Zhao, D.M. Ball, K. Babaoglu, J. Qi, S.W. White, R.E. Lee, Dihydropteroate synthase, J. Chem. Inf. Model. (2009) 444–460. https://doi.org/10.1021/ci800293n.

[33] S. Durdagi, B. Aksoydan, B. Dogan, K. Sahin, A. Shahraki, N. Birgul-Iyison, Screening of Clinically Approved and Investigation Drugs as Potential Inhibitors of SARS-CoV-2 Main Protease and Spike Receptor-Binding Domain Bound with ACE2 COVID19 Target Proteins: A Virtual Drug Repurposing Study, ChemRxiv. (2020). https://doi.org/10.26434/chemrxiv.12032712.v2.

[34] T. Ginex, J. Vazquez, E. Gilbert, E. Herrero, F.J. Luque, Lipophilicity in drug design: An overview of lipophilicity descriptors in 3D-QSAR studies, Future Med. Chem. 11 (2019) 1177–1193. https://doi.org/10.4155/fmc-2018-0435.

[35] G.A. Jeffrey, An introduction to hydrogen bonding, Oxford University Press, 1997.

[36] J. Singh, R.C. Petter, T.A. Baillie, A. Whitty, The resurgence of covalent drugs, Nat. Rev. Drug Discov. 10 (2011) 307–317. https://doi.org/10.1038/nrd3410.

[37] S. Raschka, A.J. Wolf, J. Bemister-Buffington, L.A. Kuhn, Protein-ligand interfaces are polarized: Discovery of a strong trend for intermolecular hydrogen bonds to favor donors on the protein side with implications for predicting and designing ligand complexes, BioRxiv. (2018). https://doi.org/10.1101/260612.

[38] C. Janiak, A critical account on π-π stacking in metal complexes with aromatic nitrogen-containing ligands, J. Chem. Soc. Dalt. Trans. (2000) 3885–3896. https://doi.org/10.1039/b003010o.

[39] S.K. Burley, G.A. Petsko, Aromatic-Aromatic Interaction: A Mechanism of Protein Structure Stabilization, Science (80-.). 229 (1985) 23–28. https://doi.org/10.1126/science.3892686.

[40] D. Piovesan, G. Minervini, S.C.E. Tosatto, The RING 2.0 web server for high quality residue interaction networks, Nucleic Acids Res. 44 (2016) W367–W374. https://doi.org/10.1093/nar/gkw315.

[41] M. Peikert, X. Chen, L. Chi, G. Brezesinski, S. Janich, E.U. Würthwein, H.J. Schäfer, Phase behavior and molecular packing of octadecyl phenols and their methyl ethers at the air/water interface, Langmuir. 30 (2014) 5780–5789. https://doi.org/10.1021/la404340h.

[42] S. Zhu, C.M. Noviello, J. Teng, R.M. Walsh, J.J. Kim, R.E. Hibbs, Structure of a human synaptic GABAA receptor, Nature. 559 (2018) 67–88. https://doi.org/10.1038/s41586-018-0255-3.

[43] S. Masiulis, R. Desai, T. Uchański, I. Serna Martin, D. Laverty, D. Karia, T. Malinauskas, J. Zivanov, E. Pardon, A. Kotecha, J. Steyaert, K.W. Miller, A.R. Aricescu, GABAA receptor signalling mechanisms revealed by structural pharmacology, Nature. 565 (2019) 454–459. https://doi.org/10.1038/s41586-018-0832-5.

[44] L. Richter, C. De Graaf, W. Sieghart, Z. Varagic, M. Mörzinger, I.J.P. De Esch, G.F. Ecker, M. Ernst, Diazepam-bound GABAA receptor models identify new benzodiazepine binding-site ligands, Nat. Chem. Biol. 8 (2012) 455–464. https://doi.org/10.1038/nchembio.917.

[45] K.R. Tan, A. Gonthier, R. Baur, M. Ernst, M. Goeldner, E. Sigel, Proximity-accelerated chemical coupling reaction in the benzodiazepine-binding site of γ-aminobutyric acid type A receptors: Superposition of different allosteric modulators, J. Biol. Chem. 282 (2007) 26316–26325. https://doi.org/10.1074/jbc.M702153200.

[46] J.M.C. Derry, S.M.J. Dunn, M. Davies, Identification of a residue in the γ-

aminobutyric acid type A receptor α subunit that differentially affects diazepam-sensitive and -insensitive benzodiazepine site binding, J. Neurochem. 88 (2004) 1431–1438. https://doi.org/10.1046/j.1471-4159.2003.02264.x.

[47] U. Rudolph, H. Möhler, Analysis of GABAA Receptor Function and Dissection of the Pharmacology of Benzodiazepines and General Anesthetics Through Mouse Genetics, Annu. Rev. Pharmacol. Toxicol. 44 (2004) 475–498. https://doi.org/10.1146/annurev.pharmtox.44.101802.121429.

[48] M.A. Dämgen, P.C. Biggin, A Refined Open State of the Glycine Receptor Obtained via Molecular Dynamics Simulations, Structure. 28 (2020) 130-139.e2. https://doi.org/10.1016/j.str.2019.10.019.

[49] A.A. Bliznyuk, J.E. Gready, Combining Docking and Molecular Dynamic Simulations in Drug Design, 26 (2006). https://doi.org/10.1002/med.20067.

[50] I. Daoud, N. Melkemi, T. Salah, S. Ghalem, Combined QSAR, molecular docking and molecular dynamics study on new Acetylcholinesterase and Butyrylcholinesterase inhibitors, Elsevier Ltd, 2018. https://doi.org/10.1016/j.compbiolchem.2018.03.021.

[51] L. Dean, Diazepam Therapy and CYP2C19 Genotype, Med. Genet. Summ. (2020) 1–7. http://www.ncbi.nlm.nih.gov/pubmed/28520370.

[52] T. Fukasawa, A. Suzuki, K. Otani, Effects of genetic polymorphism of cytochrome P450 enzymes on the pharmacokinetics of benzodiazepines, J. Clin. Pharm. Ther. 32 (2007) 333–341. https://doi.org/10.1111/j.1365-2710.2007.00829.x.

[53] F. Tsopelas, C. Giaginis, A. Tsantili-Kakoulidou, Lipophilicity and biomimetic properties to support drug discovery, Expert Opin. Drug Discov. 12 (2017) 885–896. https://doi.org/10.1080/17460441.2017.1344210.

[54] T.I. Saari, M. Uusi-Oukari, J. Ahonen, K.T. Olkkola, Enhancement of GABAergic activity: Neuropharmacological effects of benzodiazepines and therapeutic use in anesthesiology, Pharmacol. Rev. 63 (2011) 243–267. https://doi.org/10.1124/pr.110.002717.

[55] J. DeRuiter, GABA Receptor Ligands and the Benzodiazepines, Princ. Drug Action 2. Fall (2004) 1–27.

[56] L. Porterfield, Principles of drug action., AD Nurse. 3 (1988) 11–12. https://doi.org/10.4135/9781483326818.n3.

[57] G. Greco, E. Novellino, C. Silipo, A. Vittoria, Study of Benzodiazepines Receptor Sites Using a Combined QSAR-CoMFA Approach, Quant. Struct. Relationships. 11 (1992) 461–477. https://doi.org/10.1002/qsar.2660110403.

GENERAL CONCLUSION

As part of this dissertation work, we carried out two investigations whose findings contribute to a better understanding of the general chemistry of benzodiazepines and how they act to treat psychiatric and neurological pathologies.

Our research provided, first, an original investigation into the chemistry of six benzodiazepine basic rings. The data reported here provide chemists with a knowledge base of the structural parameters, electronic properties, as well as global, local, and long-range reactive behaviors that they need to know for modifying these six basic rings through adding/removing substituents. This contributes to ameliorating the structures of benzodiazepine drugs while regulating their main therapeutic actions or assisting to manufacture new derivatives with various therapeutic applications. Computation of global reactivity indices, including electronegativity, global hardness, global softness, electrophilicity, and nucleophilicity allowed us to describe and classify the global reactivities of considered rings. Also, statistical testing identified the density-based method Hirshfeld as the most efficient population scheme for estimating net atomic charges. As recognized, Hirshfeld ensures producing non-negative condensed Fukui functions. Accordingly, the dual descriptor 3D-mapped surfaces clearly illustrate the selectivity of local sites against attack by nucleophilic and electrophilic reactants. Indeed, the delocalization of nitrogen atoms along diazepine rings leads to significant differences in local reactivities and this may account for the substantial structural diversity of the benzodiazepine family. On the other hand, elucidating how these basic structures respond to distant reactants is crucial to acquainting all aspects of the reactivity. The repulsive and attractive electrostatic interactions are long-range in comparison to the charge-transfer effects characterized by the Fukui functions and the derived reactivity indices. To this end, a quantitative MEP analysis was performed to predict long-range intermolecular interactions. The ESP-mapped vdW surfaces illustrate well electron-deficient and electron-rich sites, thus demonstrating significant differences in the responses of local sites to long-distance nucleophilic and electrophilic attacks. Furthermore, molecular docking simulation permits incorporating the basic structures into a biological environment, and to follow the interactions between their atoms and the various types of residues. These latter act as electrophiles and nucleophiles. Therefore, docking can confirm MEP results, also offering accurate data on binding modes, binding interactions, and binding affinities with the four target sites. It is crucial to emphasize that although the binding affinities of all the structures under study are convergent, the binding modes show noticeable distinctions.

The second investigation falls within the framework of researches aimed at elucidating the binding mechanism by which a dataset of [^{3}H]diazepam derivatives allosterically modulates GABA$_A$ receptor $\alpha_1\beta_2\gamma_2$ subtypes, as well as, developing a robust predictive model that incorporates the most relevant structural requirements that contribute in improving the response of GABA$_A$/$\alpha_1\beta_2\gamma_2$ to benzodiazepines. This model can be used to predict the biological response of new derivatives or existing derivatives with an unknown biological response. Even after the discovery of the three TMD-binding interfaces, the well-known ECD-binding interface remains the main target for most of our benzodiazepines dataset, nevertheless, the orientation of certain derivatives to modulate first the TMD-incorporated binding interfaces cannot be overlooked. According to the bibliographic research, these derivatives have not been the subject of previous activity investigations, and this finding may stimulate future researches into their physiological actions. Moreover, the binding affinities-based screening selected Ro12-6377 and proflazepam as the best modulators for all four binding interfaces, simultaneously. The best binding modes and binding interactions with key residues were thoroughly discussed and compared with: native diazepam, among themselves, and outcomes from the literature. Importantly, molecular docking/dynamic simulations exhibit the ability of both modulators to establish interactions with TM$_2$ helices residues, thereby inducing direct motions in the chloride-channel lining. This observation leads to hypothesize that the two modulators probably contribute to orienting the β_2:C:Leu259 side chain towards one of the two adjacent α-subunits; this rotation is critical for opening the 9' gate and expanding the diameter of the ion channel, thereby activating the pLGICs family. Certainly, this hypothesis remains to be confirmed experimentally. Also, the good pharmacokinetic properties (ADME prediction) and compliance with all drug-likeness rules were checked via in silico tools for all the dataset compounds. The improved PLS regression yielded a robust QSAR model explaining 63.2% of the variability in benzodiazepine activity. Its stability and predictive power were ensured based on internal and external validation indicators: R^2_{adj}=0.584, F=12.806; P_{value}=6.2050e-07, Q^2_{loo}=0.639, and Q^2_{F3}=0.813. As a final result of QSAR analysis, an increase in hydrophobicity at C_7 and $C_{2'}$ and a decrease in molecular lipophilicity, molecular polarity, molecular size, and net charge of N_4 contribute to the enhancement of benzodiazepine activity. Briefly, a comprehensive interpretation was given based on the incorporation of findings from the literature and molecular docking/dynamic simulations.

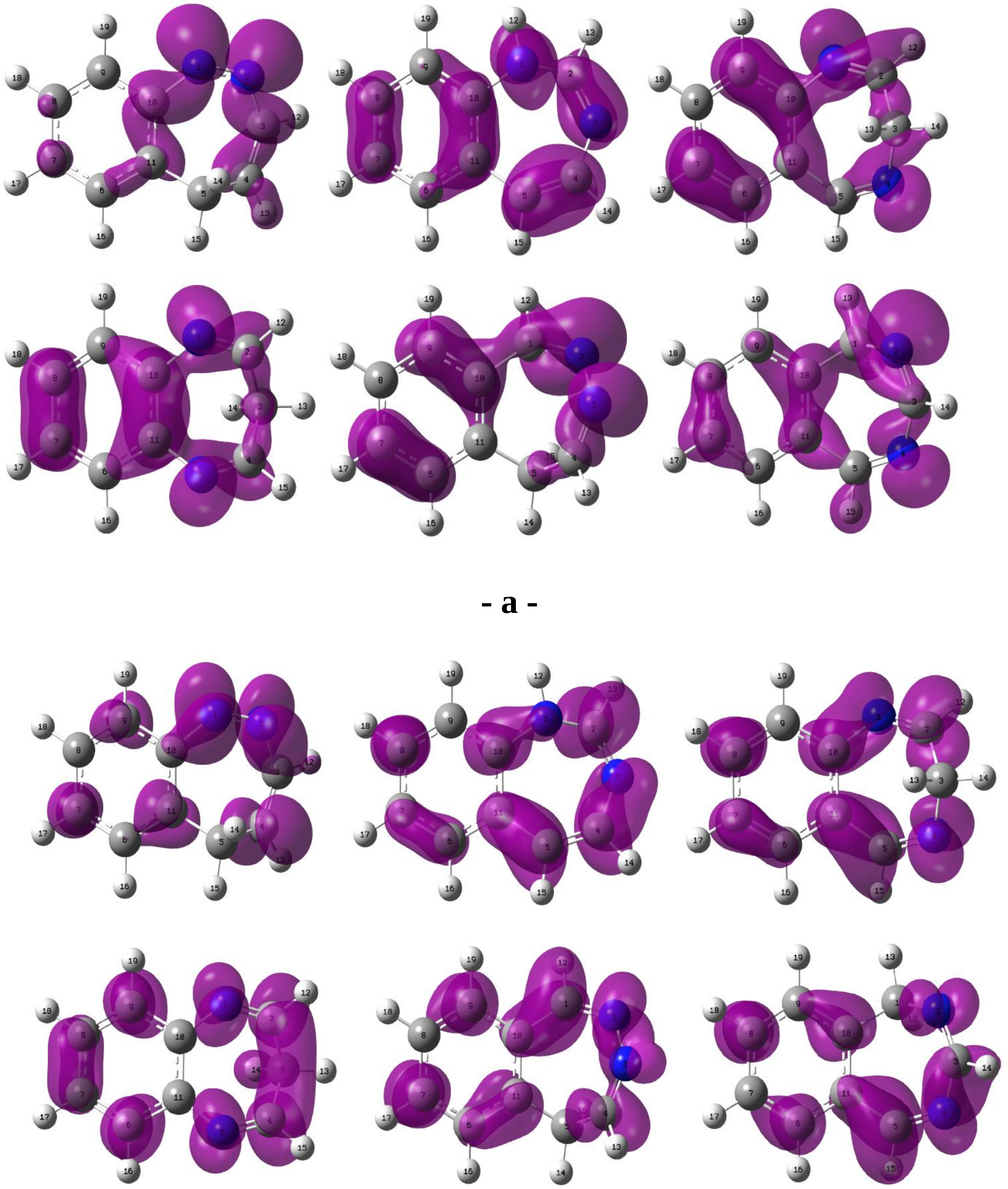
Appendices

Appendix A. Mullikan, NPA, ChelpG, and Hirshfeld Partial Charges used in the statistical test.

Atom	6-31G	6-311 G	6-31^{+} G (d)	6-311^{+} G (d)	6-31^{++} G (d,p)	6-311^{++} G (d,p)	6-31G	6-311 G	6-31^{+} G (d)	6-311^{+} G (d)	6-31^{++} G (d,p)	6-311^{++} G (d,p)
	Mullikan						**NPA**					
1 N	-0.398	-0.367	-0.228	0.004	-0.188	-0.027	-0.448	-0.461	-0.462	-0.472	-0.462	-0.470
2 C	0.045	0.029	0.026	-0.003	0.054	0.059	0.100	0.157	0.126	0.174	0.131	0.175
3 C	-0.226	-0.375	-0.407	-0.637	-0.205	-0.426	-0.395	-0.297	-0.381	-0.290	-0.365	-0.294
4 N	-0.350	-0.303	-0.189	-0.015	-0.186	-0.103	-0.437	-0.440	-0.460	-0.459	-0.456	-0.456
5 C	-0.041	-0.004	-0.349	-0.690	-0.495	-0.506	0.056	0.103	0.088	0.128	0.091	0.126
6 C	-0.181	-0.160	-0.781	-0.438	-0.212	-0.267	-0.207	-0.169	-0.201	-0.165	-0.198	-0.165
7 C	-0.131	-0.166	-0.048	-0.220	-0.122	-0.294	-0.244	-0.203	-0.246	-0.208	-0.241	-0.209
8 C	-0.136	-0.132	-0.334	-0.335	-0.084	-0.330	-0.226	-0.183	-0.222	-0.181	-0.217	-0.182
9 C	-0.121	-0.157	-0.619	-0.267	-0.444	-0.360	-0.231	-0.194	-0.233	-0.199	-0.232	-0.200
10 C	0.100	0.089	-0.169	-0.915	-0.507	-0.779	0.128	0.132	0.131	0.141	0.135	0.143
11 C	0.115	0.000	1.388	1.485	1.123	1.421	-0.144	-0.140	-0.156	-0.150	-0.151	-0.147
12 H	0.176	0.201	0.212	0.242	0.162	0.217	0.241	0.193	0.236	0.189	0.228	0.188
13 H	0.184	0.227	0.237	0.274	0.202	0.182	0.259	0.214	0.251	0.211	0.245	0.211
14 H	0.170	0.207	0.216	0.269	0.172	0.197	0.267	0.215	0.263	0.212	0.249	0.212
15 H	0.168	0.198	0.206	0.240	0.159	0.213	0.242	0.196	0.236	0.191	0.229	0.191
16 H	0.162	0.184	0.213	0.258	0.131	0.192	0.258	0.218	0.256	0.218	0.252	0.218
17 H	0.155	0.177	0.207	0.250	0.135	0.198	0.259	0.218	0.258	0.219	0.253	0.219
18 H	0.156	0.177	0.207	0.250	0.143	0.196	0.259	0.218	0.258	0.218	0.253	0.219
19 H	0.152	0.173	0.212	0.247	0.161	0.218	0.262	0.223	0.259	0.222	0.256	0.221
	ChelpG						**Hirshfeld**					
1 N	-0.659	-0.696	-0.651	-0.649	-0.650	-0.653	-0.167	-0.166	-0.167	-0.164	-0.167	-0.163
2 C	0.268	0.288	0.224	0.227	0.225	0.242	0.056	0.057	0.063	0.062	0.064	0.064
3 C	0.412	0.445	0.547	0.524	0.551	0.515	-0.018	-0.019	-0.019	-0.021	-0.019	-0.020
4 N	-0.709	-0.725	-0.748	-0.736	-0.750	-0.736	-0.189	-0.185	-0.191	-0.187	-0.190	-0.185
5 C	0.477	0.458	0.462	0.447	0.464	0.458	0.032	0.032	0.039	0.038	0.040	0.039
6 C	-0.159	-0.178	-0.178	-0.190	-0.177	-0.177	-0.035	-0.034	-0.033	-0.034	-0.032	-0.032
7 C	-0.103	-0.091	-0.086	-0.088	-0.080	-0.085	-0.041	-0.041	-0.042	-0.043	-0.041	-0.041
8 C	-0.042	-0.055	-0.053	-0.069	-0.056	-0.074	-0.033	-0.032	-0.032	-0.033	-0.031	-0.031
9 C	-0.266	-0.269	-0.252	-0.255	-0.245	-0.229	-0.049	-0.049	-0.051	-0.051	-0.050	-0.049
10 C	0.488	0.508	0.474	0.471	0.468	0.463	0.036	0.034	0.039	0.037	0.040	0.038
11 C	-0.168	-0.149	-0.124	-0.115	-0.125	-0.125	-0.025	-0.024	-0.020	-0.021	-0.019	-0.019
12 H	0.045	0.045	0.054	0.056	0.053	0.051	0.059	0.059	0.056	0.057	0.056	0.055
13 H	-0.035	-0.046	-0.082	-0.075	-0.082	-0.073	0.046	0.045	0.045	0.045	0.044	0.044
14 H	-0.034	-0.041	-0.065	-0.057	-0.066	-0.057	0.049	0.050	0.045	0.046	0.045	0.045
15 H	-0.020	-0.006	-0.009	-0.003	-0.009	-0.007	0.051	0.052	0.049	0.050	0.049	0.048
16 H	0.139	0.144	0.139	0.145	0.137	0.138	0.058	0.057	0.056	0.057	0.055	0.054
17 H	0.112	0.110	0.104	0.109	0.101	0.105	0.058	0.057	0.056	0.057	0.055	0.054
18 H	0.111	0.115	0.110	0.118	0.109	0.115	0.059	0.058	0.058	0.059	0.056	0.056
19 H	0.141	0.142	0.134	0.140	0.131	0.130	0.053	0.050	0.048	0.049	0.046	0.046

Appendix B. The bond lengths, bond angles, and dihedral angles calculated in water, using, DFT (Ub3lyp)/6-311^{++}G(d,p)/Hirshfeld/PCM model/Gaussian W09 software.

system	5*H*-1,2-bpdz	1*H*-1,3-bpdz	3*H*-1,4-bpdz	3*H*-1,5-bpdz	5*H*-2,3-bpdz	1*H*-2,4-bpdz
Bond Length (A°)						
1-2	1.263	1.392	1.276	1.278	1.294	1.470
2-3	1.408	1.279	1.509	1.508	1.383	1.276
3-4	1.342	1.403	1.468	1.508	1.279	1.400
4-5	1.503	1.345	1.281	1.278	1.507	1.289
5-11	1.509	1.468	1.469	1.402	1.511	1.464
11-6	1.397	1.402	1.409	1.410	1.397	1.404
6-7	1.393	1.395	1.385	1.383	1.393	1.390
7-8	1.397	1.391	1.401	1.403	1.398	1.397
8-9	1.388	1.396	1.385	1.383	1.389	1.394
9-10	1.405	1.394	1.407	1.410	1.405	1.395
10-11	1.411	1.409	1.419	1.423	1.412	1.412
10-1	1.422	1.429	1.403	1.402	1.466	1.508
Bond Angle (°)						
10-1-2	124.534	123.236	122.027	121.790	129.543	110.744
1-2-3	124.619	128.798	124.382	124.103	120.943	117.115
2-3-4	126.815	122.508	109.533	105.232	121.709	130.372
3-4-5	121.910	129.591	116.732	124.103	124.800	122.214
4-5-11	107.613	127.853	128.586	121.790	106.517	128.385
11-6-7	121.068	121.965	121.344	121.563	120.501	120.392
6-7-8	120.070	119.281	119.519	119.723	120.270	119.690
7-8-9	119.653	119.822	120.127	119.723	119.789	120.325
8-9-10	120.584	120.761	121.058	121.563	120.498	120.650
1-10-9	114.049	118.242	116.614	116.231	119.676	122.150
1-10-11	126.016	121.526	124.013	124.824	120.718	118.767
9-10-11	119.796	120.197	118.957	118.634	119.505	119.045
5-11-6	122.087	119.186	117.735	116.231	122.555	119.124
5-11-10	119.189	122.834	123.336	124.824	118.076	121.034
6-11-10	118.714	117.961	118.828	118.634	119.367	119.834
Dihedral Angles (°)						
10-1-2-3	-1.941	-48.360	-3.526	3.309	-9.725	65.764
1-2-3-4	-41.680	-1.721	70.545	-66.577	51.130	2.408
2-3-4-5	4.371	29.382	-65.391	66.577	-3.827	-43.474
3-4-5-11	59.751	-1.834	3.152	-3.309	-67.406	2.598
11-6-7-8	-0.793	0.189	2.783	-3.287	0.214	1.195
6-7-8-9	0.107	0.277	1.030	0.000	-0.533	0.537
7-8-9-10	2.370	-1.070	-3.208	3.287	-1.011	-0.577
8-9-10-11	-4.153	1.395	1.572	-3.253	2.843	-1.096
8-9-10-1	179.877	-176.478	174.464	-177.161	179.219	176.623
2-1-10-11	41.170	47.717	-38.468	38.445	-36.503	-66.431
2-1-10-9	-143.155	-134.440	149.038	-148.070	147.166	115.844
7-6-11-10	-0.976	0.125	-4.358	3.253	1.622	-2.875
7-6-11-5	177.823	178.613	179.174	177.161	-177.743	178.170
4-5-11-10	-61.114	-24.718	33.322	-38.444	61.602	33.308
4-5-11-6	120.092	156.871	-150.381	148.070	-119.025	-147.749
1-10-11-5	0.027	-1.536	6.081	0.000	-0.071	3.941
1-10-11-6	178.861	176.893	-170.178	173.341	-179.465	-174.994
9-10-11-6	3.412	-0.908	2.152	0.000	-3.127	2.802
9-10-11-5	-1.941	-48.360	-3.526	3.309	-9.725	65.764



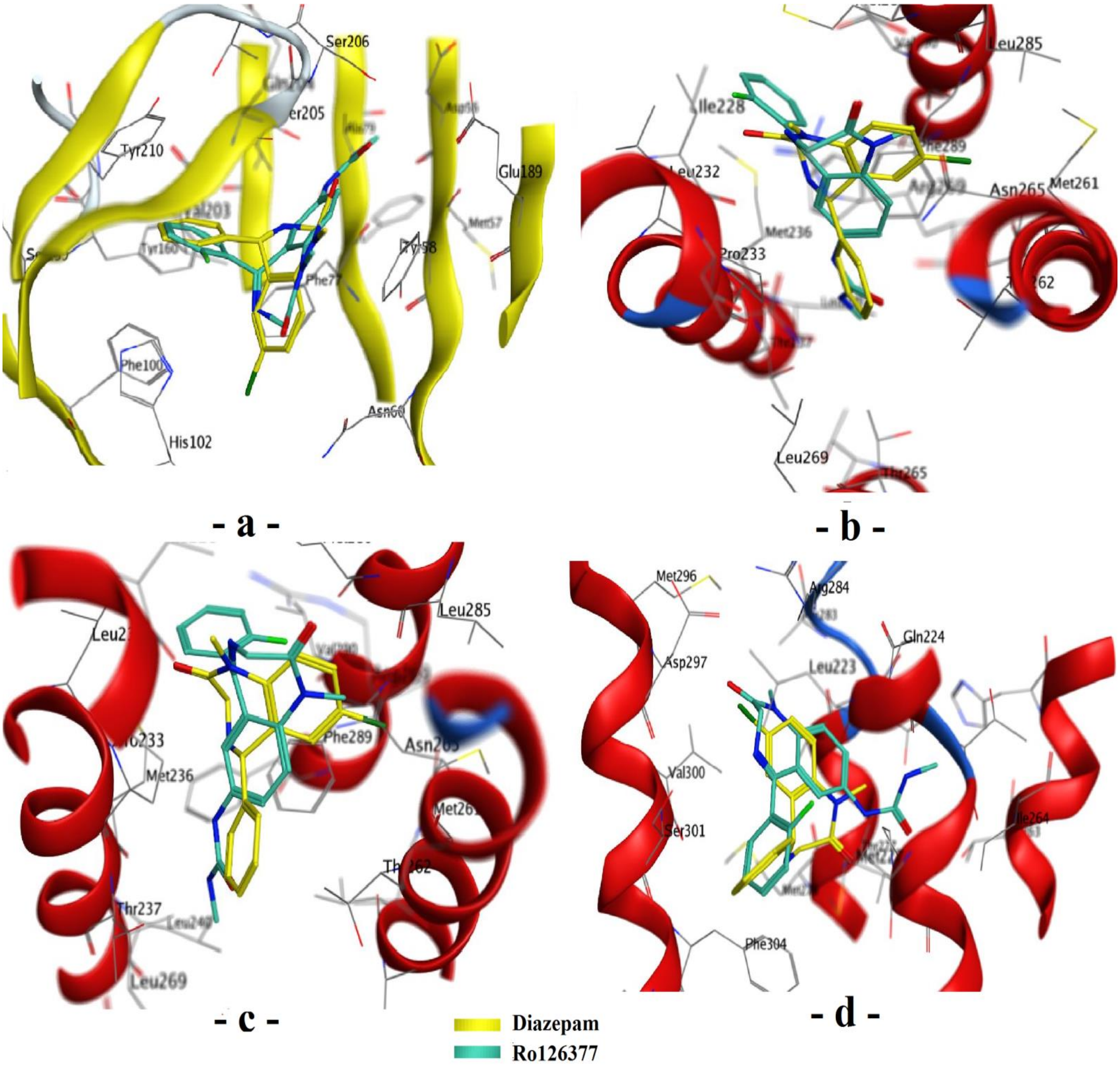
Appendix C. Condensed Fukui functions 3D-mapped surface, calculated in water using DFT (Ub3lyp)/6-311^{++}G(d,p)/Hirshfeld/PCM model/Gaussian W09 software, visualized in term of FMO theory using gaussview 5.0.8, Isovalue MO = 0.0015 a.u, purple color indicate the regions of electrophilic attacks **a,** Negative condensed Fukui function $f^{-}(r)$. **b,** Positive condensed Fukui function $f^{+}(r)$.

Appendix D. The docking scoring energy of the Co-crystallized DZP structure and the studies ligands for the four binding interfaces.

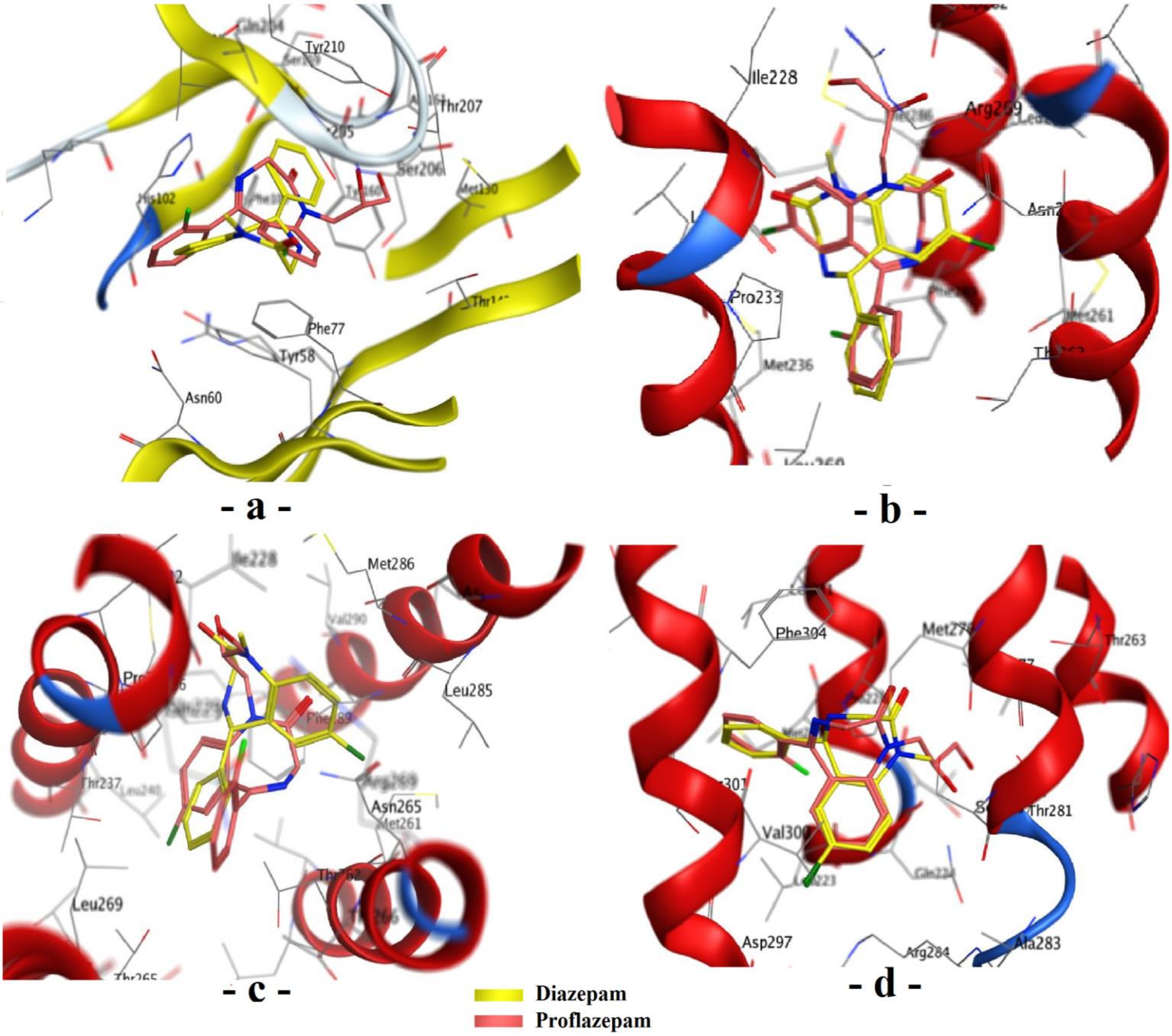
Ligand Nam	**$ECD\ \alpha_1^+/\gamma_2^-$ interface**	**$TMD\ \beta_2^+(A)/\alpha_1^-(B)$ interface**	**$TMD\ \beta_2^+(C)/\alpha_1^-(D)$ interface**	**$TMD\ \gamma_2^+/\beta_2^-$ interface**
Co-crystallized DZP	-7,003	-6,159	-6,261	-6,347
Ro05-4318/ Ro05-3418	-6,957	-6,421	-6,056	-6,442
Ro05-3072	-6,420	-6,252	-5,982	-6,137
Ro05-4528	-7,100	-6,606	-6,443	-6,539
Ro05-2921	-6,185	-6,181	-5,978	-5,685
Ro20-7736	-7,216	-6,765	-6,413	-6,887
Ro05-4619	-6,613	-6,371	-6,107	-6,145
Ro20-5397	-6,694	-6,467	-6,697	-6,398
Ro05-3061	-6,261	-5,965	-5,815	-5,896
Ro20-2533	-7,162	-6,303	-6,417	-6,618
Ro20-2541	-7,101	-6,733	-6,558	-6,523
Ro20-5747	-6,979	-6,251	-6,403	-6,515
Ro05-4336	-6,144	-6,258	-5,975	-5,948
Ro20-3053	-7,090	-6,444	-6,216	-6,829
Triflunordazepam	-6,935	-5,968	-6,091	-6,495
Diazepam	-7,040	-6,529	-6,278	-6,592
Ro07-5220	-7,021	-6,075	-5,463	-6,531
Ro14-3074	-7,029	-6,548	-6,178	-6,412
Flunitrazepam	-7,540	-6,470	-5,821	-6,600
Ro05-3590	-7,535	-6,473	-6,377	-6,276
Norflurazepam	-6,543	-5,899	-5,731	-6,270
Delorazepam	-6,785	-6,365	-5,904	-6,118
Clonazepam	-7,554	-6,615	-6,350	-6,000
Fonazepam	-7,329	-6,334	-5,984	-6,387
Ro05-6822	-6,838	-6,402	-6,097	-6,228
Ro05-4865	-6,837	-6,173	-6,027	-6,086
Ro05-6820	-6,182	-5,980	-5,884	-5,896
Nordazepam	-6,550	-5,905	-5,925	-6,191
Ro07-3953	-6,568	-5,916	-5,779	-6,318
Difludiazepam	-6,755	-6,490	-6,431	-6,555
Ro07-5193	-6,726	-6,329	-6,165	-6,260
Ro22-3294	-6,699	-5,838	-5,698	-6,091
Nitrazepam	-7,143	-6,406	-5,918	-6,493
Methylclonazepam	-7,831	-6,531	-6,447	-6,906
7-Aminoflunitrazepam	-6,965	-6,357	-6,188	-6,459
Ro12-6377	-8,013	-7,409	-7,265	-7,546
Halazepam	-6,858	-6,791	-6,280	-6,615
Pinazepam	-7,151	-7,146	-6,848	-6,990
Prazepam	-7,067	-6,961	-6,435	-6,421
Motrazepam	-7,556	-7,087	-5,919	-6,382
Ro20-1310	-6,631	-6,156	-6,606	-5,425
Ro07-2750	-6,938	-7,152	-6,456	-6,918
Ro08-9013	-7,580	-7,050	-7,455	-7,070
Proflazepam	-7,807	-7,170	-7,077	-6,972

Appendix D. Continued

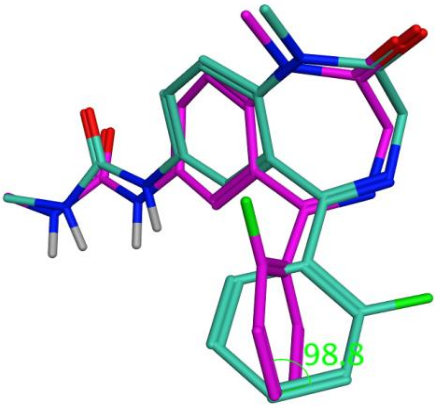
Ro22-4683	-6,859	-5,566	-5,802	-5,827
Ro11-4878	-6,729	-6,312	-6,141	-6,427
Meclonazepam	-7,856	-6,821	-6,774	-6,307
Ro11-6896	-7,785	-6,688	-5,836	-6,351
L48	-7,296	-6,801	-6,620	-6,017
Temazepam	-7,208	-6,523	-6,469	-6,810
L50	-7,221	-6,830	-6,675	-6,522
L51	-6,497	-6,116	-5,645	-6,568
Ro07-4419	-6,118	-6,102	-5,912	-5,829
Ro05-4520	-6,779	-6,485	-6,433	-6,167
Ro05-4608	-6,787	-6,373	-6,612	-6,377
Ro05-3546	-6,046	-6,199	-6,066	-5,951
Ro13-0699	-6,298	-6,336	-6,411	-6,579
Ro07-6198	-6,517	-6,020	-5,989	-6,178
Ro20-8895	-6,669	-6,299	-5,985	-6,126
Ro13-0593	-6,884	-6,517	-6,260	-6,460
L60	-6,535	-6,168	-6,056	-6,100
Ro22-6762	-6,676	-6,014	-6,419	-6,471
Ro20-8065	-6,675	-6,218	-5,874	-6,609
Ro20-8552	-6,702	-6,212	-6,277	-6,704
L64	-6,795	-6,173	-5,908	-6,288
L65	-6,834	-6,230	-5,926	-6,394



Appendix E. Superposition of Ro12-6377 to the DZP-bound structures at the interfaces of: **a**, ECD α_1^+/γ_2^-. **b**, TMD β_2^+(A)/α_1^-(B). **c**, TMD β_2^+(C)/α_1^-(D), and **d**, TMD γ_2^+/β_2^-.



Appendix F. Superpositions of proflazepam to the DZP-bound structures at the interfaces of: **a**, ECD α_1^+/γ_2^-. **b**, TMD $\beta_2^+(A)/\alpha_1^-(B)$. **c**, TMD $\beta_2^+(C)/\alpha_1^-(D)$, and **d**, TMD γ_2^+/β_2^-.



Appendix G. Superposition of Ro12-6377 binding modes at the two TMD $\beta_2^+(A)/\alpha_1^-(B)$ and TMD $\beta_2^+(C)/\alpha_1^-(D)$ interfaces, the purple color designed the Ro12-6377-bound structure at $\beta_2^+(C)/\alpha_1^-(D)$ interface.

Appendix H. Pharmacokinetic parameters and drug-likeness prediction.

Ligand Nam	GI absorption	BBB permeant	Pgp substrate	CYP1A2 inhibitor	CYP2C19 inhibitor	CYP2C9 inhibitor	CYP2D6 inhibitor	CYP3A4 inhibitor	Lipinski	Ghose	Veber	Egan	Muegge
Ro05-4318/ Ro05-3418	High	Yes	No	Yes	No	No	No	Yes	yes	yes	yes	yes	yes
Ro05-3072	High	No	Yes	Yes	No	No	No	Yes	yes	yes	yes	yes	yes
Ro05-4528	High	Yes	No	Yes	Yes	Yes	No	Yes	yes	yes	yes	yes	yes
Ro05-2921	High	Yes	No	Yes	No	No	No	Yes	yes	yes	yes	yes	yes
Ro20-7736	High	Yes	No	Yes	No	No	No	No	yes	yes	yes	yes	yes
Ro05-4619	High	Yes	Yes	Yes	No	No	Yes	Yes	yes	yes	yes	yes	yes
Ro20-5397	High	Yes	No	Yes	No	No	No	No	yes	yes	yes	yes	yes
Ro05-3061	High	Yes	No	Yes	No	No	Yes	Yes	yes	yes	yes	yes	yes
Ro20-2533	High	Yes	No	Yes	Yes	No	Yes	Yes	yes	yes	yes	yes	yes
Ro20-2541	High	Yes	No	Yes	Yes	Yes	No	Yes	yes	yes	yes	yes	yes
Ro20-5747	High	Yes	No	Yes	Yes	No	Yes	Yes	yes	yes	yes	yes	yes
Ro05-4336	High	Yes	No	Yes	No	No	Yes	Yes	yes	yes	yes	yes	yes
Ro20-3053	High	Yes	No	Yes	Yes	No	No	Yes	yes	yes	yes	yes	yes
Triflunordazepam	High	Yes	No	Yes	Yes	No	Yes	No	yes	yes	yes	yes	yes
Diazepam	High	Yes	No	Yes	Yes	Yes	Yes	Yes	yes	yes	yes	yes	yes
Ro07-5220	High	Yes	No	Yes	Yes	Yes	No	No	yes	yes	yes	yes	yes
Ro14-3074	High	No	No	Yes	No	No	No	No	yes	yes	yes	yes	yes
Flunitrazepam	High	No	Yes	Yes	No	No	No	No	yes	yes	yes	yes	yes
Ro05-3590	High	No	Yes	Yes	No	No	No	No	yes	yes	yes	yes	yes

Appendix H. Continued

Norflurazepam	High	Yes	No	Yes	Yes	No	Yes	Yes	yes	yes	yes	yes	yes
Delorazepam	High	Yes	No	Yes	Yes	Yes	Yes	Yes	yes	yes	yes	yes	yes
Clonazepam	High	No	Yes	Yes	No	No	No	No	yes	yes	yes	yes	yes
Fonazepam	High	No	Yes	Yes	No	No	No	No	yes	yes	yes	yes	yes
Ro05-6822	High	Yes	No	Yes	Yes	No	Yes	No	yes	yes	yes	yes	yes
Ro05-4865	High	Yes	No	Yes	Yes	No	No	No	yes	yes	yes	yes	yes
Ro05-6820	High	Yes	No	Yes	No	No	Yes	Yes	yes	yes	yes	yes	yes
Nordazepam	High	Yes	No	Yes	Yes	No	Yes	Yes	yes	yes	yes	yes	yes
Ro07-3953	High	Yes	No	Yes	Yes	No	Yes	Yes	yes	yes	yes	yes	yes
Difludiazepam	High	Yes	No	No	Yes	No	No	No	yes	yes	yes	yes	yes
Ro07-5193	High	Yes	No	Yes	Yes	Yes	Yes	Yes	yes	yes	yes	yes	yes
Ro22-3294	High	Yes	No	Yes	Yes	Yes	No	Yes	yes	yes	yes	yes	yes
Nitrazepam	High	No	Yes	Yes	No	No	No	No	yes	yes	yes	yes	yes
Methylclonazepam	High	No	Yes	Yes	No	No	No	No	yes	yes	yes	yes	yes
7-Aminoflunitrazepam	High	Yes	No	No	No	No	No	No	yes	yes	yes	yes	yes
Ro12-6377	High	Yes	Yes	No	No	No	No	No	yes	yes	yes	yes	yes
Halazepam	High	Yes	No	Yes	Yes	Yes	Yes	No	yes	yes	yes	yes	yes
Pinazepam	High	Yes	No	Yes	Yes	Yes	Yes	No	yes	yes	yes	yes	yes
Prazepam	High	Yes	No	Yes	Yes	Yes	Yes	Yes	yes	yes	yes	yes	yes
Motrazepam	High	No	Yes	Yes	No	No	No	No	yes	yes	yes	yes	yes
Ro20-1310	High	Yes	No	Yes	Yes	Yes	Yes	Yes	yes	yes	yes	yes	yes

Appendix H. Continued

Ro07-2750	High	Yes	No	Yes	Yes	No	Yes	Yes	yes	yes	yes	yes	yes
Ro08-9013	High	No	Yes	No	Yes	No	No	No	yes	yes	yes	yes	yes
Proflazepam	High	Yes	Yes	No	No	No	Yes	No	yes	yes	yes	yes	yes
Ro22-4683	High	No	Yes	Yes	No	No	No	No	yes	yes	yes	yes	yes
Ro11-4878	High	Yes	No	No	Yes	No	Yes	Yes	yes	yes	yes	yes	yes
Meclonazepam	High	No	Yes	Yes	No	No	No	No	yes	yes	yes	yes	yes
Ro11-6896	High	No	Yes	No	No	No	No	No	yes	yes	yes	yes	yes
L48	High	Yes	No	No	Yes	Yes	Yes	No	yes	yes	yes	yes	yes
Temazepam	High	Yes	No	No	No	No	No	No	yes	yes	yes	yes	yes
L50	High	Yes	No	No	Yes	Yes	No	No	yes	yes	yes	yes	yes
L51	High	Yes	No	Yes	Yes	No	Yes	Yes	yes	yes	yes	yes	yes
Ro07-4419	High	Yes	No	Yes	No	No	Yes	No	yes	yes	yes	yes	yes
Ro05-4520	High	Yes	No	Yes	Yes	No	No	No	yes	yes	yes	yes	yes
Ro05-4608	High	Yes	No	Yes	Yes	No	No	Yes	yes	yes	yes	yes	yes
Ro05-3546	High	Yes	No	Yes	Yes	No	Yes	Yes	yes	yes	yes	yes	yes
Ro13-0699	High	Yes	No	Yes	Yes	No	No	No	yes	yes	yes	yes	yes
Ro07-6198	High	Yes	No	Yes	Yes	No	Yes	Yes	yes	yes	yes	yes	yes
Ro20-8895	High	Yes	No	Yes	No	No	Yes	Yes	yes	yes	yes	yes	yes
Ro13-0593	High	Yes	No	Yes	Yes	No	Yes	No	yes	yes	yes	yes	yes
L60	High	Yes	No	Yes	Yes	Yes	No	No	yes	yes	yes	yes	yes
Ro22-6762	High	Yes	No	Yes	Yes	Yes	No	No	yes	yes	yes	yes	yes

Appendix H. Continued

Ro20-8065	High	Yes	No	Yes	Yes	Yes	Yes	Yes	yes	yes	yes	yes	yes
Ro20-8552	High	Yes	No	Yes	Yes	Yes	Yes	Yes	yes	yes	yes	yes	yes
L64	High	Yes	No	Yes	Yes	Yes	Yes	Yes	yes	yes	yes	yes	yes
L65	High	Yes	No	Yes	Yes	No	Yes	Yes	yes	yes	yes	yes	yes

Appendix I. Molecular descriptors selected for QSAR analysis.

Comp N°	Q_{N1}	Q_{C3}	Q_{N4}	Q_{C6}	Q_{C7}	Q_{C8}	Q_{C9}	$Q_{C2'}$	$Q_{C6'}$	π_7	$\pi_{2'}$	HA	HD	DM (debye)	Log (P)	M_R	F_T
1	-0,096	0,568	-0,608	-0,227	0,376	-0,268	-0,087	-0,102	-0,106	-1,230	0,000	3,000	1,000	6,301	1,643	79,708	1
2	-0,541	0,539	-0,618	-0,242	0,389	-0,259	-0,152	-0,091	-0,044	-1,230	0,000	3,000	2,000	6,583	1,779	76,592	1
3	-0,131	0,533	-0,590	-0,060	-0,043	-0,120	-0,091	-0,080	-0,075	-0,570	0,000	3,000	0,000	2,121	2,328	80,729	2
4	-0,568	0,518	-0,620	-0,005	-0,147	-0,059	-0,195	-0,095	-0,045	0,000	0,000	2,000	1,000	5,061	2,608	71,891	1
5	-0,083	0,596	-0,655	-0,207	0,515	-0,386	-0,007	0,351	-0,057	-1,340	0,140	4,000	2,000	6,721	2,123	82,211	2
6	-0,494	0,631	-0,698	-0,289	0,433	-0,256	-0,153	0,040	-0,110	-1,230	0,710	3,000	2,000	6,791	2,383	81,396	1
7	-0,514	0,513	-0,620	-0,003	-0,050	-0,160	-0,126	-0,068	-0,011	-0,650	0,000	3,000	1,000	3,137	2,320	78,475	2
8	-0,531	0,523	-0,613	-0,209	0,381	-0,250	-0,144	-0,083	-0,019	0,140	0,000	2,000	1,000	3,369	2,751	72,108	1
9	-0,558	0,532	-0,619	-0,127	0,018	-0,147	-0,177	-0,071	-0,024	1,020	0,000	2,000	1,000	5,662	3,566	81,533	2
10	-0,128	0,581	-0,634	-0,077	-0,043	-0,116	-0,094	0,327	-0,068	-0,570	0,140	3,000	0,000	3,218	2,471	80,945	2
11	-0,512	0,554	-0,631	-0,105	0,182	-0,217	-0,137	-0,088	-0,034	0,820	0,000	2,000	1,000	4,936	3,345	81,578	2
12	-0,543	0,540	-0,657	-0,061	-0,134	-0,044	-0,201	0,336	-0,014	0,000	0,140	2,000	1,000	5,752	2,751	72,108	1
13	-0,516	0,594	-0,670	-0,112	-0,098	-0,022	-0,191	0,344	-0,021	-0,550	0,140	3,000	1,000	3,274	2,308	82,510	2
14	-0,543	0,509	-0,614	-0,017	-0,144	-0,061	-0,188	-0,090	-0,045	0,880	0,000	2,000	1,000	1,909	3,486	77,865	2
15	-0,102	0,530	-0,601	-0,027	0,074	-0,110	-0,109	-0,081	-0,053	0,710	0,000	2,000	0,000	3,228	3,076	79,812	1
16	-0,081	0,559	-0,679	-0,091	0,115	-0,099	-0,116	0,038	-0,037	0,710	0,710	2,000	0,000	3,462	4,284	89,421	1
17	-0,498	0,561	-0,659	-0,195	0,350	-0,262	-0,105	0,353	-0,013	0,460	0,140	4,000	1,000	4,057	4,085	82,395	2

Appendix I. Continued

18	-0,111	0,550	-0,622	-0,112	0,022	-0,150	-0,072	0,340	-0,061	-0,280	0,140	5,000	2,000	3,209	2,555	81,544	2
19	-0,486	0,605	-0,673	-0,146	0,048	-0,133	-0,153	-0,070	-0,079	-0,280	0,880	5,000	3,000	3,692	3,426	84,185	3
20	-0,510	0,585	-0,665	-0,061	0,060	-0,044	-0,201	0,343	-0,019	0,710	0,140	2,000	1,000	4,140	3,355	76,912	1
21	-0,483	0,606	-0,696	-0,084	0,077	-0,054	-0,186	0,078	-0,102	0,710	0,710	2,000	1,000	3,938	3,816	81,501	1
22	-0,468	0,627	-0,696	-0,184	0,080	-0,149	-0,131	0,058	-0,117	-0,280	0,710	5,000	3,000	2,488	3,152	83,016	2
23	-0,520	0,562	-0,649	-0,116	0,022	-0,120	-0,142	0,357	-0,017	-0,280	0,140	5,000	3,000	2,378	2,691	78,428	2
24	-0,101	0,563	-0,637	-0,236	0,371	-0,249	-0,063	0,356	-0,049	0,140	0,140	2,000	0,000	4,373	2,757	75,440	1
25	-0,090	0,544	-0,597	-0,201	0,370	-0,269	-0,062	-0,072	-0,071	0,140	0,000	2,000	0,000	3,338	2,615	57,224	1
26	-0,492	0,565	-0,654	-0,251	0,393	-0,256	-0,100	0,345	-0,008	0,140	0,140	2,000	1,000	4,242	2,893	72,324	1
27	-0,527	0,549	-0,624	-0,028	0,064	-0,071	-0,184	-0,113	-0,037	0,710	0,000	3,000	3,000	3,257	3,212	76,696	1
28	-0,520	0,560	-0,689	-0,011	0,046	-0,039	-0,213	0,399	0,425	0,710	0,140	2,000	1,000	3,645	3,497	77,129	1
29	-0,108	0,469	-0,700	-0,039	0,091	-0,117	-0,088	0,355	0,381	0,710	0,140	2,000	0,000	3,646	3,361	80,245	1
30	-0,507	0,520	-0,681	-0,085	0,114	-0,090	-0,176	0,102	0,357	0,710	0,710	2,000	1,000	3,557	3,959	81,717	1
31	-0,465	0,623	-0,710	-0,092	0,105	-0,076	-0,179	0,070	0,004	0,710	0,710	2,000	1,000	3,486	4,420	86,306	1
32	-0,531	0,520	-0,615	-0,072	0,006	-0,113	-0,162	-0,084	-0,043	-0,280	0,000	5,000	3,000	1,381	2,548	78,212	2
33	-0,083	0,601	-0,663	-0,157	0,067	-0,165	-0,066	0,059	-0,145	-0,280	0,710	5,000	2,000	3,365	3,016	97,151	2
34	-0,084	0,674	-0,665	-0,262	0,408	-0,271	-0,080	0,334	-0,075	-1,230	0,140	3,000	1,000	6,893	1,786	79,924	1
35	-0,088	0,577	-0,655	-0,193	0,373	-0,269	-0,055	0,356	-0,035	-1,030	0,140	3,000	2,000	4,302	1,743	93,660	4
36	-0,145	0,462	-0,573	-0,029	0,052	-0,077	-0,134	-0,070	-0,052	0,710	0,000	2,000	0,000	4,083	4,031	85,262	3

Appendix I. Continued

37	-0,173	0,564	-0,607	-0,008	0,047	-0,091	-0,080	-0,090	-0,063	0,710	0,000	2,000	0,000	3,104	3,304	87,392	3
38	-0,204	0,494	-0,584	-0,008	0,050	-0,089	-0,114	-0,075	-0,052	0,710	0,000	2,000	0,000	3,049	3,857	91,754	3
39	-0,077	0,487	-0,569	-0,126	0,019	-0,125	-0,103	-0,057	-0,010	-0,280	0,000	6,000	2,000	1,785	2,475	87,181	4
40	-0,411	0,502	-0,587	-0,013	0,032	-0,055	-0,145	-0,053	-0,030	0,710	0,000	2,000	0,000	3,222	4,130	93,618	2
41	-0,325	0,593	-0,642	-0,042	0,050	-0,028	-0,265	0,343	-0,047	0,710	0,140	3,000	1,000	3,078	2,529	86,321	3
42	-0,221	0,609	-0,640	-0,078	0,086	-0,070	-0,202	0,368	-0,023	0,710	0,140	4,000	1,000	5,437	1,843	98,979	6
43	-0,179	0,577	-0,634	-0,057	0,069	-0,087	-0,108	0,362	-0,048	0,710	0,140	4,000	2,000	5,140	1,898	92,283	4
44	-0,451	0,613	-0,657	-0,152	0,046	-0,174	-0,026	0,028	-0,158	-0,280	0,710	5,000	2,000	4,125	4,070	99,938	3
45	-0,626	0,622	-0,736	-0,014	0,039	-0,042	-0,231	0,357	-0,021	0,710	0,140	2,000	1,000	3,539	3,923	81,406	1
46	-0,581	0,763	-0,766	-0,134	0,053	-0,128	-0,170	0,059	-0,129	-0,280	0,710	5,000	3,000	2,404	3,721	87,510	2
47	-0,213	0,718	-0,698	-0,102	0,017	-0,122	-0,126	0,359	-0,067	-0,280	0,140	5,000	2,000	3,063	3,123	86,038	2
48	-0,180	0,696	-0,660	-0,030	0,079	-0,096	-0,138	-0,069	-0,073	0,710	0,000	2,000	0,000	2,958	3,645	84,306	1
49	-0,045	0,897	-0,609	-0,012	0,050	-0,070	-0,111	-0,085	-0,078	0,710	0,000	3,000	1,000	4,751	2,787	81,010	1
50	-0,181	0,408	-0,541	-0,043	0,053	-0,077	-0,135	0,369	-0,026	0,710	0,140	2,000	0,000	6,653	4,124	84,764	1
51	-0,552	0,512	-0,647	0,058	0,069	-0,191	-0,155	-0,043	-0,016	0,560	0,000	2,000	1,000	5,677	3,635	81,974	1
52	-0,559	0,540	-0,674	-0,029	-0,142	-0,037	-0,206	0,390	0,417	0,000	0,140	2,000	1,000	5,429	2,893	72,324	1
53	-0,122	0,600	-0,648	-0,071	-0,108	-0,108	-0,086	0,319	-0,069	0,000	0,140	2,000	0,000	5,768	2,615	75,224	1
54	-0,110	0,626	-0,666	-0,123	-0,077	-0,102	-0,097	0,028	-0,170	0,000	0,710	2,000	0,000	5,557	3,076	79,812	1
55	-0,559	0,493	-0,605	0,050	-0,087	-0,122	-0,148	-0,043	-0,042	0,000	0,000	2,000	1,000	4,500	3,212	76,696	1

Appendix I. Continued

56	-0,094	0,581	-0,635	-0,011	-0,078	-0,124	-0,072	0,393	-0,025	0,000	0,140	2,000	0,000	5,583	3,219	80,028	1
57	-0,526	0,549	-0,684	-0,049	-0,122	0,127	-0,192	0,406	0,446	0,000	0,140	2,000	1,000	4,474	3,497	77,129	1
58	-0,525	0,569	-0,676	-0,026	-0,262	0,231	-0,323	0,339	-0,021	0,000	0,140	2,000	1,000	6,296	3,264	77,149	1
59	-0,116	0,549	-0,620	-0,095	-0,117	-0,077	0,059	0,380	-0,017	0,000	0,140	2,000	0,000	5,954	3,219	80,028	1
60	-0,045	0,588	-0,632	-0,059	-0,024	0,006	-0,006	0,352	-0,059	0,000	0,140	2,000	0,000	4,116	3,823	84,833	1
61	-0,087	0,535	-0,611	-0,060	0,073	0,048	-0,067	-0,104	-0,091	0,710	0,000	2,000	0,000	2,420	3,680	84,617	1
62	-0,512	0,564	-0,662	-0,063	0,048	0,101	-0,189	0,351	-0,004	0,710	0,140	2,000	1,000	3,363	3,959	81,717	1
63	-0,503	0,560	-0,668	-0,246	0,197	-0,014	-0,127	0,356	-0,006	0,560	0,140	2,000	1,000	5,148	3,868	81,954	1
64	-0,481	0,514	-0,614	-0,027	0,055	-0,077	-0,006	-0,068	-0,023	0,710	0,000	2,000	1,000	3,589	3,816	81,501	1
65	-0,462	0,506	-0,617	-0,070	0,108	-0,221	0,117	-0,101	-0,042	0,710	0,000	2,000	1,000	3,336	3,725	81,737	1

Appendix J. Hydrophobic interactions established with C_7 and $C_{2'}$ positions and electrostatic interactions established with N_4, resulting from the molecular docking of the data set at ECD α_1^+/γ_2^- interface.

ECD α_1^+/γ_2^- interface								
Ligand N°	**Posit [a]**	**category**	**Type of interactions**	**From**	**From chemistry**	**To**	**To chemistry**	**Dist [b] (A°)**
L12	4	Hydrogen Bond	C-H Bond	D:HIS102:HE1	H-Donor	:*0:N	H-Acceptor	2,82
L14	7	Hydrophobic	Pi-Alkyl	E:PHE77	Pi-Orbitals	:*0:C	Alkyl	5,49
L15	7	Hydrophobic	Pi-Alkyl	E:TYR58	Pi-Orbitals	:*0:CL	Alkyl	5,37
L19	2’	Hydrophobic	Pi-Alkyl	D:TYR160	Pi-Orbitals	:*0:C	Alkyl	4,65
L20	7	Hydrophobic	Alkyl	E:ALA79	Alkyl	:*0:CL	Alkyl	4,16
L21	7	Hydrophobic	Pi-Alkyl	D:HIS102	Pi-Orbitals	:*0:CL	Alkyl	4,36
L26	4	Hydrogen Bond	C-H Bond	D:HIS102:HE1	H-Donor	:*0:N	H-Acceptor	2,80
L27	7	Hydrophobic	Alkyl	E:ALA79	Alkyl	:*0:CL	Alkyl	4,25
L28	7	Hydrophobic	Pi-Alkyl	D:HIS102	Pi-Orbitals	:*0:CL	Alkyl	4,41
L29	7	Hydrophobic	Pi-Alkyl	E:TYR58	Pi-Orbitals	:*0:CL	Alkyl	5,32
L30	4	Hydrogen Bond	Carbon Hydrogen Bond	D:HIS102:HE1	H-Donor	:*0:N	H-Acceptor	2,81
	7	Hydrophobic	Pi-Alkyl	E:TYR58	Pi-Orbitals	:*0:CL	Alkyl	4,59
L35	7	Hydrophobic	Alkyl	E:ALA79	Alkyl	:*0:C	Alkyl	3,55
L36	7	Hydrophobic	Pi-Alkyl	E:TYR58	Pi-Orbitals	:*0:CL	Alkyl	5,24
L38	4	Hydrogen Bond	Carbon Hydrogen Bond	D:HIS102:HE1	H-Donor	:*0:N	H-Acceptor	2,60
	7	Hydrophobic	Pi-Alkyl	E:TYR58	Pi-Orbitals	:*0:CL	Alkyl	3,47
L42	4	Hydrogen Bond	C-H Bond	D:HIS102:HE1	H-Donor	:*0:N	H-Acceptor	2,69
	7	Hydrophobic	Pi-Alkyl	E:TYR58	Pi-Orbitals	:*0:CL	Alkyl	4,82
L45	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	E:MET130	Alkyl	4,11
L48	7	Hydrophobic	Pi-Alkyl	E:TYR58	Pi-Orbitals	:*0:CL	Alkyl	5,32
L49	7	Hydrophobic	Pi-Alkyl	E:TYR58	Pi-Orbitals	:*0:CL	Alkyl	5,38
L51	7	Hydrophobic	Pi-Alkyl	E:TYR58	Pi-Orbitals	:*0:C	Alkyl	5,02
	7	Hydrophobic	Pi-Alkyl	E:PHE77	Pi-Orbitals	:*0:C	Alkyl	4,96
L52	4	Hydrogen Bond	C-H Bond	D:HIS102:HE1	H-Donor	:*0:N	H-Acceptor	2,81
L61	7	Hydrophobic	Pi-Alkyl	E:TYR58	Pi-Orbitals	:*0:CL	Alkyl	4,58
L62	7	Hydrophobic	Pi-Alkyl	D:HIS102	Pi-Orbitals	:*0:CL	Alkyl	4,43
L63	7	Hydrophobic	Pi-Alkyl	E:TYR58	Pi-Orbitals	:*0:C	Alkyl	4,31
L64	7	Hydrophobic	Alkyl	E:ALA79	Alkyl	:*0:CL	Alkyl	4,48

a: position of interaction, **b**: distance of interaction, ***0:** ligand structure

Appendix K. Hydrophobic interactions established with C_7 and $C_{2'}$ positions and electrostatic interactions established with N_4, resulting from the molecular docking of the data set at TMD $\beta_2^+(A)/\alpha_1^-(B)$ interface.

TMD $\beta_2^+(A)/\alpha_1^-(B)$ interface								
Ligand N°	**Posit [a]**	**category**	**Type of interactions**	**From**	**From chemistry**	**To**	**To chemistry**	**Dist [b] (A°)**
L2	4	Hydrogen Bond	Carbon Hydrogen Bond	A:MET286:HA	H-Donor	:*0:N	H-Acceptor	2,74
L4	4	Hydrogen Bond	Conventional Hydrogen Bond	A:ASN265:HD21	H-Donor	:*0:N	H-Acceptor	2,34
L6	4	Hydrogen Bond	Conventional Hydrogen Bond	A:ASN265:HD21	H-Donor	:*0:N	H-Acceptor	2,77
L8	4	Hydrogen Bond	Conventional Hydrogen Bond	A:ASN265:HD21	H-Donor	:*0:N	H-Acceptor	2,33
L9	7	Hydrophobic	Alkyl	:*0:C	Alkyl	A:MET286	Alkyl	4,71
	7	Hydrophobic	Alkyl	:*0:C	Alkyl	B:ILE228	Alkyl	4,13
L11	7	Hydrophobic	Alkyl	:*0:C	Alkyl	A:MET286	Alkyl	4,17
	7	Hydrophobic	Alkyl	:*0:C	Alkyl	B:ILE228	Alkyl	3,95
L12	4	Hydrogen Bond	Conventional Hydrogen Bond	A:ASN265:HD21	H-Donor	:*0:N	H-Acceptor	2,31
L14	7	Hydrophobic	Alkyl	:*0:C	Alkyl	B:PRO233	Alkyl	4,08
L15	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:PRO233	Alkyl	4,63
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:MET236	Alkyl	5,30
	7	Hydrophobic	Pi-Alkyl	A:PHE289	Pi-Orbitals	:*0:CL	Alkyl	4,69
L16	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:LEU232	Alkyl	4,27
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:MET236	Alkyl	3,84
L18	4	Hydrogen Bond	Carbon Hydrogen Bond	B:PRO233:HA	H-Donor	:*0:N	H-Acceptor	2,94
L19	4	Hydrogen Bond	Carbon Hydrogen Bond	B:PRO233:HA	H-Donor	:*0:N	H-Acceptor	2,97
	2’	Hydrophobic	Alkyl	:*0:C	Alkyl	B:PRO233	Alkyl	4,25
L20	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	A:MET286	Alkyl	4,56
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:ILE228	Alkyl	4,06
L21	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:LEU232	Alkyl	4,04
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:MET236	Alkyl	3,89
L26	4	Hydrogen Bond	Conventional Hydrogen Bond	A:ASN265:HD21	H-Donor	:*0:N	H-Acceptor	2,30
L27	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	A:VAL290	Alkyl	5,02
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:MET236	Alkyl	4,11
	7	Hydrophobic	Pi-Alkyl	A:PHE289	Pi-Orbitals	:*0:CL	Alkyl	5,12
L28	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	A:VAL290	Alkyl	4,99
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:MET236	Alkyl	4,13
	7	Hydrophobic	Pi-Alkyl	A:PHE289	Pi-Orbitals	:*0:CL	Alkyl	5,10
L29	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:PRO233	Alkyl	4,69
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:MET236	Alkyl	5,26
	7	Hydrophobic	Pi-Alkyl	A:PHE289	Pi-Orbitals	:*0:CL	Alkyl	4,73

Appendix K. Continued

L30	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:LEU232	Alkyl	4,04
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:MET236	Alkyl	3,93
L31	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:PRO233	Alkyl	4,23
L35	7	Hydrophobic	Alkyl	:*0:C	Alkyl	A:VAL258	Alkyl	5,37
	7	Hydrophobic	Alkyl	:*0:C	Alkyl	B:LEU240	Alkyl	5,04
	7	Hydrophobic	Pi-Alkyl	A:PHE289	Pi-Orbitals	:*0:C	Alkyl	5,30
L36	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:LEU232	Alkyl	4,09
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:MET236	Alkyl	3,77
L37	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:LEU232	Alkyl	4,22
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:MET236	Alkyl	3,71
L38	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:MET236	Alkyl	3,80
L39	4	Hydrogen Bond	Carbon Hydrogen Bond	A:MET286:HA	H-Donor	:*0:N	H-Acceptor	2,56
L40	4	Hydrogen Bond	Conventional Hydrogen Bond	A:ASN265:HD21	H-Donor	:*0:N	H-Acceptor	2,18
	7	Hydrophobic	Pi-Alkyl	A:PHE289	Pi-Orbitals	:*0:CL	Alkyl	5,35
L42	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	A:MET286	Alkyl	4,55
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:ILE228	Alkyl	4,15
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:MET236	Alkyl	3,73
L44	4	Hydrogen Bond	Carbon Hydrogen Bond	B:PRO233:HA	H-Donor	:*0:N	H-Acceptor	2,99
L45	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:PRO233	Alkyl	4,62
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:MET236	Alkyl	5,37
	7	Hydrophobic	Pi-Alkyl	A:PHE289	Pi-Orbitals	:*0:CL	Alkyl	4,60
L48	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:PRO233	Alkyl	4,54
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:MET236	Alkyl	5,47
	7	Hydrophobic	Pi-Alkyl	A:PHE289	Pi-Orbitals	:*0:CL	Alkyl	4,57
L49	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:PRO233	Alkyl	4,48
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:MET236	Alkyl	5,37
	7	Hydrophobic	Pi-Alkyl	A:PHE289	Pi-Orbitals	:*0:CL	Alkyl	4,31
L51	7	Hydrophobic	Pi-Alkyl	A:PHE289	Pi-Orbitals	:*0:C	Alkyl	3,74
L52	4	Hydrogen Bond	Conventional Hydrogen Bond	A:ASN265:HD21	H-Donor	:*0:N	H-Acceptor	2,30
L53	4	Hydrogen Bond	Conventional Hydrogen Bond	A:ASN265:HD21	H-Donor	:*0:N	H-Acceptor	2,28
L54	4	Hydrogen Bond	Conventional Hydrogen Bond	A:ASN265:HD21	H-Donor	:*0:N	H-Acceptor	2,26
L57	4	Hydrogen Bond	Conventional Hydrogen Bond	A:ASN265:HD21	H-Donor	:*0:N	H-Acceptor	2,84

Appendix K. Continued

L58	4	Hydrogen Bond	Carbon Hydrogen Bond	A:MET286:HA	H-Donor	:*0:N	H-Acceptor	2,82
L62	7	Hydrophobic	Pi-Alkyl	A:PHE289	Pi-Orbitals	:*0:CL	Alkyl	3,76
L63	7	Hydrophobic	Pi-Alkyl	A:PHE289	Pi-Orbitals	:*0:C	Alkyl	3,76
L64	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:PRO233	Alkyl	4,43
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:MET236	Alkyl	5,39
	7	Hydrophobic	Pi-Alkyl	A:PHE289	Pi-Orbitals	:*0:CL	Alkyl	4,23
L65	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:PRO233	Alkyl	4,47
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	B:MET236	Alkyl	5,20
	7	Hydrophobic	Pi-Alkyl	A:PHE289	Pi-Orbitals	:*0:CL	Alkyl	4,24

a: position of interaction, **b**: distance of interaction, ***0:** ligand structure

Appendix L. Hydrophobic interactions established with C_7 and $C_{2'}$ positions and electrostatic interactions established with N_4, resulting from the molecular docking of the data set at TMD $\beta_2^+(C)/\alpha_1^-(D)$ interface.

TMD $\beta_2^+(C)/\alpha_1^-(D)$ interface								
Ligand N°	**Posit [a]**	**category**	**Type of interactions**	**From**	**From chemistry**	**To**	**To chemistry**	**Dist [b] (A°)**
L2	4	Hydrogen Bond	Carbon Hydrogen Bond	D:PRO233:HA	H-Donor	:*0:N	H-Acceptor	2,61
L4	4	Hydrogen Bond	Conventional Hydrogen Bond	C:ASN265:HD22	H-Donor	:*0:N	H-Acceptor	2,09
L6	4	Hydrogen Bond	Carbon Hydrogen Bond	D:PRO233:HA	H-Donor	:*0:N	H-Acceptor	2,64
L7	4	Hydrogen Bond	Carbon Hydrogen Bond	D:PRO233:HA	H-Donor	:*0:N	H-Acceptor	2,58
L8	4	Hydrogen Bond	Conventional Hydrogen Bond	C:ASN265:HD22	H-Donor	:*0:N	H-Acceptor	2,07
L9	7	Hydrophobic	Alkyl	:*0:C	Alkyl	C:LEU285	Alkyl	5,03
	7	Hydrophobic	Pi-Alkyl	C:PHE289	Pi-Orbitals	:*0:C	Alkyl	4,94
L11	7	Hydrophobic	Alkyl	:*0:C	Alkyl	C:ARG269	Alkyl	4,41
L12	4	Hydrogen Bond	Conventional Hydrogen Bond	C:ASN265:HD22	H-Donor	:*0:N	H-Acceptor	2,06
L13	7	Hydrogen Bond	Carbon Hydrogen Bond	C:MET286:HA	H-Donor	:*0:O	H-Acceptor	2,80
L14	7	Hydrophobic	Alkyl	:*0:C	Alkyl	D:PRO233	Alkyl	4,85
L15	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:PRO233	Alkyl	4,71
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:MET236	Alkyl	5,19
	7	Hydrophobic	Pi-Alkyl	C:PHE289	Pi-Orbitals	:*0:CL	Alkyl	4,62
L16	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:MET236	Alkyl	5,37
	7	Hydrophobic	Pi-Alkyl	C:PHE289	Pi-Orbitals	:*0:CL	Alkyl	4,80
L18	4	Hydrogen Bond	Carbon Hydrogen Bond	C:MET286:HA	H-Donor	:*0:N	H-Acceptor	2,54

Appendix L. Continued

L19	4	Halogen	Halogen (Fluorine)	:*0:N	Halogen Acceptor	:*0:F	Halogen	2,83
	4	Halogen	Halogen (Fluorine)	:*0:N	Halogen Acceptor	:*0:F	Halogen	2,91
	2’	Hydrophobic	Alkyl	:*0:C	Alkyl	D:PRO233	Alkyl	4,56
	2’	Hydrophobic	Alkyl	:*0:C	Alkyl	D:MET236	Alkyl	5,07
	2’	Hydrophobic	Pi-Alkyl	C:PHE289	Pi-Orbitals	:*0:C	Alkyl	4,58
L20	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:PRO233	Alkyl	4,94
L21	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:PRO233	Alkyl	4,66
L26	4	Hydrogen Bond	Carbon Hydrogen Bond	D:PRO233:HA	H-Donor	:*0:N	H-Acceptor	2,64
L28	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:PRO233	Alkyl	4,93
L29	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:PRO233	Alkyl	4,75
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:MET236	Alkyl	5,38
	7	Hydrophobic	Pi-Alkyl	C:PHE289	Pi-Orbitals	:*0:CL	Alkyl	4,72
L30	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:PRO233	Alkyl	4,97
L31	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:PRO233	Alkyl	4,92
L35	7	Hydrophobic	Alkyl	:*0:C	Alkyl	D:LEU240	Alkyl	4,63
L36	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:PRO233	Alkyl	4,75
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:MET236	Alkyl	5,11
	7	Hydrophobic	Pi-Alkyl	C:PHE289	Pi-Orbitals	:*0:CL	Alkyl	4,68
L37	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:PRO233	Alkyl	4,75
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:MET236	Alkyl	5,27
	7	Hydrophobic	Pi-Alkyl	C:PHE289	Pi-Orbitals	:*0:CL	Alkyl	4,41
L38	4	Hydrogen Bond	Conventional Hydrogen Bond	C:ASN265:HD22	H-Donor	:*0:N	H-Acceptor	1,93
L40	4	Hydrogen Bond	Conventional Hydrogen Bond	C:ASN265:HD22	H-Donor	:*0:N	H-Acceptor	1,89
L42	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	C:LEU285	Alkyl	4,59
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	C:MET286	Alkyl	4,00
L45	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:PRO233	Alkyl	4,69
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:MET236	Alkyl	5,22
	7	Hydrophobic	Pi-Alkyl	C:PHE289	Pi-Orbitals	:*0:CL	Alkyl	4,56
L48	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:PRO233	Alkyl	4,66
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:MET236	Alkyl	5,22
	7	Hydrophobic	Pi-Alkyl	C:PHE289	Pi-Orbitals	:*0:CL	Alkyl	4,67
L49	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:PRO233	Alkyl	4,67
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:MET236	Alkyl	5,20
	7	Hydrophobic	Pi-Alkyl	C:PHE289	Pi-Orbitals	:*0:CL	Alkyl	4,64
L51	7	Hydrophobic	Alkyl	:*0:C	Alkyl	C:MET286	Alkyl	4,64
L52	4	Hydrogen Bond	Conventional Hydrogen Bond	C:ASN265:HD22	H-Donor	:*0:N	H-Acceptor	2,04
L54	4	Hydrogen Bond	Carbon Hydrogen Bond	D:PRO233:HA	H-Donor	:*0:N	H-Acceptor	2,84
	2’	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:PRO233	Alkyl	5,20
	2’	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:MET236	Alkyl	5,47
	2’	Hydrophobic	Pi-Alkyl	C:PHE289	Pi-Orbitals	:*0:CL	Alkyl	4,52
L57	4	Hydrogen Bond	Conventional Hydrogen Bond	C:ASN265:HD22	H-Donor	:*0:N	H-Acceptor	2,06
L58	4	Hydrogen Bond	Conventional Hydrogen Bond	C:ASN265:HD22	H-Donor	:*0:N	H-Acceptor	2,12

Appendix L. Continued

L60	4	Hydrogen Bond	Carbon Hydrogen Bond	C:THR262:HA	H-Donor	:*0:N	H-Acceptor	2,70
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:LEU232	Alkyl	4,01
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:MET236	Alkyl	4,21
L61	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:PRO233	Alkyl	4,70
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:LEU269	Alkyl	5,22
	7	Hydrophobic	Pi-Alkyl	C:PHE289	Pi-Orbitals	:*0:CL	Alkyl	4,68
L62	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:PRO233	Alkyl	4,71
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:MET236	Alkyl	5,24
	7	Hydrophobic	Pi-Alkyl	C:PHE289	Pi-Orbitals	:*0:CL	Alkyl	4,64
L63	7	Hydrophobic	Alkyl	:*0:C	Alkyl	C:LEU285	Alkyl	5,20
	7	Hydrophobic	Pi-Alkyl	C:PHE289	Pi-Orbitals	:*0:C	Alkyl	4,97
L64	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:PRO233	Alkyl	4,71
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:MET236	Alkyl	5,09
	7	Hydrophobic	Pi-Alkyl	C:PHE289	Pi-Orbitals	:*0:CL	Alkyl	4,41
L65	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:PRO233	Alkyl	4,70
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	D:MET236	Alkyl	5,09
	7	Hydrophobic	Pi-Alkyl	C:PHE289	Pi-Orbitals	:*0:CL	Alkyl	4,42

a: position of interaction, **b**: distance of interaction,***0:** ligand structure

Appendix M. Hydrophobic interactions established with C_7 and $C_{2'}$ positions and electrostatic interactions established with N_4, resulting from the molecular docking of the data set at TMD γ_2^+/β_2^- interface.

TMD γ_2^+/β_2^- interface								
Ligand N°	**Posit [a]**	**category**	**Type of interactions**	**From**	**From chemistry**	**To**	**To chemistry**	**Dist [b] (A°)**
L9	7	Hydrophobic	Alkyl	:*0:C	Alkyl	A:PRO228	Alkyl	4,01
L11	7	Hydrophobic	Alkyl	:*0:C	Alkyl	A:PRO228	Alkyl	4,16
L14	7	Hydrophobic	Alkyl	:*0:C	Alkyl	E:VAL300	Alkyl	4,15
L15	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	A:LEU223	Alkyl	4,73
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	E:VAL300	Alkyl	4,26
L16	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	E:ARG284	Alkyl	3,73
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	E:VAL300	Alkyl	3,99
L19	4	Hydrogen Bond	Carbon Hydrogen Bond	E:SER301:HA	H-Donor	:*0:N	H-Acceptor	3,05
	2’	Hydrophobic	Alkyl	:*0:C	Alkyl	A:PRO228	Alkyl	4,56
	2’	Hydrophobic	Alkyl	:*0:C	Alkyl	A:LEU231	Alkyl	4,98
	2 ‘	Hydrophobic	Pi-Alkyl	E:PHE304	Pi-Orbitals	:*0:C	Alkyl	4,09
L20	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	A:PRO228	Alkyl	4,07
L21	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	E:MET276	Alkyl	4,94
	7	Hydrophobic	Pi-Alkyl	E:PHE304	Pi-Orbitals	:*0:CL	Alkyl	4,40
L27	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	E:MET276	Alkyl	5,14
	7	Hydrophobic	Pi-Alkyl	E:PHE304	Pi-Orbitals	:*0:CL	Alkyl	4,43
L28	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	A:PRO228	Alkyl	4,11
L29	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	A:LEU223	Alkyl	4,44
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	E:VAL300	Alkyl	4,51
L30	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	E:MET276	Alkyl	4,93
	7	Hydrophobic	Pi-Alkyl	E:PHE304	Pi-Orbitals	:*0:CL	Alkyl	4,40
L31	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	E:MET276	Alkyl	4,82
	7	Hydrophobic	Pi-Alkyl	E:PHE304	Pi-Orbitals	:*0:CL	Alkyl	4,42
L35	7	Hydrophobic	Alkyl	:*0:C	Alkyl	A:PRO228	Alkyl	4,75
	7	Hydrophobic	Alkyl	:*0:C	Alkyl	A:ILE264	Alkyl	3,81

Appendix M. Continued

L36	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	E:VAL300	Alkyl	4,18
L37	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	A:LEU223	Alkyl	4,57
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	E:VAL300	Alkyl	4,42
L38	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	A:LEU223	Alkyl	4,45
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	E:VAL300	Alkyl	4,58
L40	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	E:MET276	Alkyl	5,40
	7	Hydrophobic	Pi-Alkyl	E:PHE304	Pi-Orbitals	:*0:CL	Alkyl	4,37
L42	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	A:LEU223	Alkyl	4,40
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	E:VAL300	Alkyl	4,56
L45	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	A:PRO228	Alkyl	4,13
L48	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	A:LEU223	Alkyl	4,73
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	E:VAL300	Alkyl	4,29
L49	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	A:LEU223	Alkyl	4,71
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	E:VAL300	Alkyl	4,29
L51	7	Hydrophobic	Alkyl	:*0:C	Alkyl	E:VAL300	Alkyl	4,22
L61	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	A:LEU223	Alkyl	4,80
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	E:VAL300	Alkyl	4,20
L62	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	E:MET276	Alkyl	5,18
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	A:PRO228	Alkyl	4,44
L63	7	Hydrophobic	Alkyl	:*0:C	Alkyl	E:MET276	Alkyl	5,28
	7	Hydrophobic	Pi-Alkyl	E:PHE304	Pi-Orbitals	:*0:C	Alkyl	4,49
L64	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	A:LEU223	Alkyl	4,59
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	E:VAL300	Alkyl	4,35
L65	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	A:LEU223	Alkyl	4,60
	7	Hydrophobic	Alkyl	:*0:CL	Alkyl	E:VAL300	Alkyl	4,38

a: position of interaction, **b**: distance of interaction, ***0:** ligand structure

Orbital: Electron. J. Chem. **2021**, 13(4), 301-315

the electronic journal of chemistry

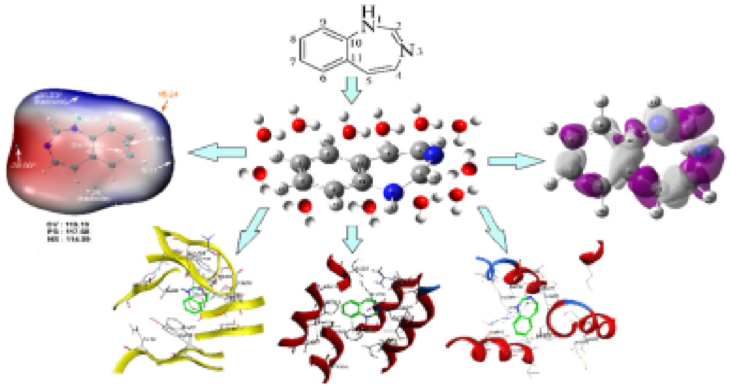
Full Paper | http://dx.doi.org/10.17807/orbital.v13i4.1607

Combined Conceptual-DFT, Quantitative MEP Analysis, and Molecular Docking Study of Benzodiazepine Analogs

Rachida Djebaili [a], Nadjib Melkemi [a], Samir Kenouche [b], Ismail Daoud [c], Mohammed Bouachrine [d,e], Halima Hazhazi [a], and Toufik Salah [a]

In the present work, a combined approach based on conceptual-DFT formalism and molecular docking simulations were performed to investigate the chemical reactivity of six Benzodiazepine analogs. Chemical reactivity descriptors derived from the conceptual DFT were determined and discussed to explain the global and local reactivity of the six studied analogs. Also, long-range interactions were studied using the quantitative analyses of molecular electrostatic potential (MEP) on van der Waals surface to identify the nucleophilic and electrophilic sites. Moreover, a statistical analysis was performed to assess the robustness of atomic charges to the basis set. The results revealed that Hirshfeld population analysis (HPA) was the most efficient for this purpose. Molecular docking simulations were performed to predict the binding affinities of the issued molecules and estimate the binding poses into four binding sites, three of them were recently discovered, located in $GABA_A$ receptor.

Graphical abstract



Keywords

Benzodiazepine
Binding site
Dual descriptor
Molecular docking
Quantitative MEP analysis
$GABA_A$

Article history

Received 16 March 2021
Revised 09 August 2021
Accepted 30 august 2021
Available online 26 September 2021

Handling Editor: Arlan Gonçalves

1. Introduction

The term benzodiazepines (BDZ) refer to bicyclic heterocyclic compounds based on a benzene nucleus fused to a diazepine ring. The delocalization of nitrogen atoms in the diazepine ring divided the benzodiazepines into six basic rings: 5*H*-1,2-bdz, 1*H*-1,3-bdz, 3*H*-1,4-bdz, 3*H*-1,5-bdz, 5*H*-2,3-bdz, and 1*H*-2,4-bdz (Figure 1) [1]. The classical BDZ based on

---

[a] Group of Computational and Pharmaceutical Chemistry, LMCE Laboratory, University Mohamed Khider, 07000 Biskra, Algeria. [b] Group of Modeling of Chemical Systems using Quantum Calculations, Applied Chemistry Laboratory (LCA). University M. Khider of Biskra, 07000 Biskra, Algeria. [c] University Abou-Bakr Belkaid-Faculty of Science- Department of Chemistry- Laboratory of Natural Substances and Bioactive (LASNABIO)-Tlemcen-Algeria. [d] MCNS Laboratory, Faculty of Science, Moulay Ismail University of Meknes, Morocco. [e] EST Khenifra, Sultan Moulay Slimane University, Morocco. *Corresponding author. E-mail: **bouachrine@gmail.com**

Structural Chemistry
https://doi.org/10.1007/s11224-022-02029-4

ORIGINAL RESEARCH

Investigation of [^{3}H]diazepam derivatives as allosteric modulators of GABAA receptor $\alpha_1\beta_2\gamma_2$ subtypes: combination of molecular docking/dynamic simulations, pharmacokinetics/drug-likeness prediction, and QSAR analysis

Rachida Djebaili[1] · **Samir Kenouche**[2] · **Ismail Daoud**[3,4] · **Nadjib Melkemi**[1] · **Ahlem Belkadi**[1] · **Fouzia Mesli**[3]

Received: 9 March 2022 / Accepted: 2 August 2022

Abstract

In this paper, a data set of [^{3}H] diazepam derivatives was analyzed using various computational methods: molecular docking/dynamic simulations, and QSAR analysis. The main aims of these studies are to understand the binding mechanisms by which benzodiazepines allosterically modulate GABA$_A$ receptor $\alpha_1\beta_2\gamma_2$ subtypes, from inducing neuronal inhibition at lower doses to the anesthetic effect at higher doses, and also, to define the structural requirements that contribute to improving the response of GABA$_A$/$\alpha_1\beta_2\gamma_2$ receptor to benzodiazepine drugs. The results of the molecular docking study allowed selecting Ro12-6377 and proflazepam as the best modulators for the four binding sites simultaneously. Subsequently, the stability of the selected complexes was investigated by performing molecular dynamics simulation. The latter confirmed the features of both modulators to exert direct effects on the chloride-channel lining residues. Pharmacokinetics and drug-likeness profile were assessed through in silico tool. Furthermore, a QSAR analysis was conducted using an improved vemolecular dynamics simulations proposed byrsion of PLS regression. The goodness of fit and the predictive power of the resulting PLS model were estimated according to internal and external validation parameters: $R^2 = 0.632$, $R^2_{\mathrm{adj}} = 0.584$, $F = 12.806$; p-value = 6.2050e − 07, $Q^2_{\mathrm{loo}} = 0.639$, and $Q^2_{\mathrm{F3}} = 0.813$. Clearly, the obtained results ensure the predictive ability of the developed QSAR model for the design of new high-potency benzodiazepine drugs.

Keywords Benzodiazepine · GABA$_A$ receptor · Extracellular domain · Transmembrane domain · Chloride channel · TM$_2$ helix

Introduction

Gamma-aminobutyric acid type A (GABA$_A$) are fast-acting ionotropic receptors that belong to the superfamily of Cys loop-type ligand-gated ion channels (PLGICs) [1]. This type of GABA receptor is found at 20–50% of brain synapses [2].GABA$_A$Rs assemble into pentameric isoforms consisting of five heteromeric subunits that surround a selective chloride-conducting pore. The general architecture of each mature subunit is constructed by the sequence of 450 amino acid residues [3]. In total, 200–250 amino acids contribute in the extracellular domain (ECD) to form a long hydrophilic N-terminal α-helix followed by ten β-strands folded into a β-sandwich containing the Cys-loops. A total of 85–255 amino acids contribute in the transmembrane domain (TMD) to form four membrane α-helices (named from TM$_1$ to TM$_4$) connected by three loops (short intracellular loop links TM$_1$-TM$_2$, short extracellular loop links TM$_2$-TM$_3$, and long cytoplasmic loop links TM$_3$-TM$_4$) and terminating with one small extracellular C-terminal [4]. The five TM$_2$α helices were assembled in the center to form the chloride-channel liner, with a possible contribution fromTM$_1$ [5]. Upon activation, GABA$_A$Rs allow negative chloride ions to flow through

---

✉ Ismail Daoud
i.daoud@univ-biskra.dz

[1] Group of Computational and Pharmaceutical Chemistry, LMCE Laboratory, University of Mohamed Khider, 07000 Biskra, Algeria

[2] Group of Modeling of Chemical Systems Using Quantum Calculations, Applied Chemistry Laboratory (LCA), University of Mohamed Khider, 07000 Biskra, Algeria

[3] Laboratory of Natural and Bioactive Substances, LASNABIO, University of Abou-BakrBelkaid, Tlemcen, Algeria

[4] Present address: Department of Matter Sciences, University of Mohamed Khider Biskra, BP 145 RP, 07000 Biskra, Algeria