الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

وزارة التعليم العالي والبحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

جامعة الشهيد حمة لخضر الوادي

Université Echahid Hamma Lakhdar -El OUED

كلية العلوم الدقيقة

Faculté de Sciences Exactes

قسم الكيمياء

Département de Chimie

Thèse en vue de l'obtention du diplôme de

DOCTORAT

Spécialité : Pharmacochimie des substances actives d'intérêt biologique

Thème :

Etude par la modélisation moléculaire, analyse in vitro et in silico de l'activité anticancer du sein de quelques N-ferrocénylméthylanilines

Présentée par : M^elle ZEGHEB Nadjiba

Soutenue le : 27/01/2022

Devant le jury composé de :

M^me KERASSA Aicha	**MCA**	**Université d'El-Oued**	**Présidente**
M^me. BOUBEKRI Cherifa	**Professeur**	**Université de Biskra**	**Rapporteur**
Mr. LANEZ Touhami	**Professeur**	**Université d'El-Oued**	**Co-rapporteur**
M^me. MEDILA Ifriqiya	**MCA**	**Université d'El-Oued**	**Examinatrice**
Mr. BECHKI Lazhar	**Professeur**	**Université de Ouargla**	**Examinateur**
Mr. LAHRACH Mokhtar	**Professeur**	**Université de Djelfa**	**Examinateur**

Année universitaire :2021/2022

People's Democratic Republic of Algeria

Ministry of Higher Education and Research Scientist

University Echahid Hamma Lakhdar -El Oued

Faculty of Exact Sciences

Department of Chemistry

Thesis in view of obtaining the diploma of

DOCTORAT

Specialty: Pharmacochemistry of active substances of biological interest

Theme:

Study by molecular modeling, in vitro and in silico analysis of the anti-breast cancer activity of some N-ferrocenylmethylanilines

Presented by: Ms. ZEGHEB Nadjiba

Graduated on: 27/07/2022

In front of the jury:

Mrs. KERASSA Aicha	MCA	University of El-Oued	President
Mrs. BOUBEKRI Cherifa	Professor	University of Biskra	Rapportor
Mr. LANEZ Touhami	Professor	University of El-Oued	Co-rapportor
Mrs. MEDILA Ifriqiya	MCA	University of El-Oued	Examiner
Mr. BECHKI Lazhar	Professor	University of Ouargla	Examiner
Mr. LAHRACH Mokhtar	Professor	University of Djelfa	Examiner

Academic year: 2021/2022

Acknowledgements

First of all, I pray to Allah, the Almighty for providing me with this opportunity and granting me the capability to proceed successfully.

At the outset, I would like to thank my supervisors Dr. BOUBEKRI Cherifa and Pr. LANEZ Touhami for providing me with the opportunity to do my doctoral research under their supervision, continuous support, encouragement throughout my research and for their beneficial comments and discussions which contributed in the making and completion of this thesis.

I wish to extend my sincere thanks to Dr.Kerasaa Aicha, MCA at the University of El Oued , for accepting to chair the dissertation of my thesis. I would also like to thank the committee members for agreeing to examine and judge my work.

I would like to express my special thanks to Dr. Tuba Tüylü Küçükkilinç, Associate Professor at Hacettepe university, Faculty of Pharmacy in Ankara, for accepting the collaboration with us to perform the In vitro part in their laboratory as well as her warm reception and good hospitality during my training. Also, I would like to thank her student Esin Oz for her help during the whole work and all the team of the laboratory.

Furthermore, I want to express my gratitude to Prof. BELAIDI Salah and Dr. KHAMOULI Saida from university of Biskra for their help and valuable advice in the QSAR part.

My special thanks and deep gratitude go to my Coach and Dear-Friend Dr. KERASSA Aicha, University of El-Oued, for her beneficial advice and her priceless support during the whole period of my research.

I would like to thank all the members of VTRS Laboratory at University of El Oued, Ali Tliba, Aicha Adaika, Hani Mimouna, Guediri Imane, Kedadra Abdellatif, Hacen Ben Amara, Elhafnaoui Lanez, Hanane Mouada, Ali Khennoufa, Messaoud Meraghni, Dr. Meriem Henni and Dr. Saad Bouchkioua, whom I had the opportunity to work with. They provided me a useful feedback and insightful comments on my work.

I also wish to thank all My School team and specially the special English teacher Raouf Boukoucha for being such a great teacher who inspired us.

Therefore, I would like to thank all the people who contributed in a way or another to the work described in this thesis.

Last, but not least, I would like to dedicate this thesis to my beloved parents who raised me with love for science and supported me in all my pursuits, my siblings and my friends for their love and continuous support – both spiritually and materially. THANK YOU

ZEGHEB Nadjiba

To my dear parents

To my brothers and sisters

To all my friends

To all those who are dear to me

Abstract

Breast cancer is the most common cancer in women in front of colorectal cancer and lung cancer. Ferrocene derivatives have proved their efficiency against breast cancer cells. In this thesis we aimed to find new ferrocene derivatives having pharmacochemistry properties as potential drugs against human breast cancer cells. A series of 29 N-ferrocenylmethylaniline derivatives (FMA) were optimized to determine the structural, crystal properties and molecular spectroscopy using DFT/ B3LYP method with 6-311 ++ G (d,p) and LanL2DZ basis sets. The Pharmacokinetics and ADME properties were also studied. The anti-proliferative activity of the studied derivatives against both hormone-dependent (MCF-7) and independent (MDA-MB 231) human breast cancer cell lines were performed using MTT test. Moreover, the binding affinity between FMA derivatives and MCF-7 was further studied by molecular docking using AutoDock 4.2 software, using both estrogen and progesterone receptors as targets. Next, another docking study was performed with aromatase and 17β-HSD1, the enzymes that catalyse the estrogen synthesis. Finally, the anti-proliferative activity of FMA derivatives against MCF-7 cell lines was investigated to predict a Quantitative Structure–Activity Relationship (QSAR) model using molecular descriptors and MLR analysis.

The FMA derivatives have shown an important cytotoxicity effect against MCF-7 cell lines, which was explained using molecular docking by the interaction of the compounds with the ERα and PR. Furthermore, the molecular docking study against aromatase shows that A4 and A25 were the best with the lowest binding free energies comparing to other compounds and ASD. However, moderate affinity was predicted against 17β-HSD1 for all compounds comparing to E2B. This indicate that these derivatives might be strong drug inhibitors of Aromatase. The best equation of regression contains HOMO, E, MR, MV, HF, POL, TPSA, qC11 and qN1 descriptors. The QSAR model indicates that these descriptors have significant relationships with the observed bioactivity. The model was validated for predictive ability inside the model using Leave One Out method (LOO), which proves that the model can be successfully applied to predict anti-proliferative activity against MCF-7 cell line.

Abbreviations list

ΔG	Free Energy
μ	Dipole moment
17β-HSD1	17-beta-hydroxysteroid dehydrogenase type 1
ADT	AutoDock Tools
ASD	Androstenedione
ATCC	American Type Culture Collection
B3LYP	Becke's three-parameter hybrid functional Lee, Yang and Parr.
BBB	Blood-Brain Barrier
DFT	Density-Functional Theory
DMEM	Dulbecco's Modified Eagle's Medium
E	total energy
EDTA	Ethylenediamine tetraacetic acid
ER	Estrogen receptors
ERα	Estrogen receptors alpha
FBS	fetal calf serum
GapE /ΔE	energetic gap
GI	GastroIntestinal
GIAO	Gauge Independent Atomic Orbital
HAc	H acceptor
HDo	H donor
HF	heat of formation
FMA	N-ferrocenylmethylaniline
HOMO	highest occupied molecular orbital,
IC50	The half maximal inhibitory concentration
IR	Infrared
K	binding constant
log P	partition coefficient octanol/water
logS	water solubility
LOO	leave –one–out cross–validation method
LUMO	Lowest unoccupied molecular orbital
MESP	The molecular electrostatic potential
MLR	multiple linear regression
MM	Molecular mechanics
MPA	Mulliken population analysis
MR	Molar refractivity
MTT	tetrazolium dye
MV	Molecular volume
MW	Molecular weight
NMR	Nuclear magnetic resonance
PBS	Phosphate Buffered Saline solution
PDB	Protein Data Bank
PDBQT	Protein data Bank, Q for charges and T for torsion
PE	predictive error
Pgp	P-glycoprotein
PLIP	Protein-Ligand Interaction Profiler web server
Pol	polarizability
PR	progesterone receptor
PRESS	predicted residual sum squares
PSE	predictive square error
q	atomic charges
Q	Factor of the quality
QSAR	Quantitative structure-activity relationship
R^2adj	Adjusted R^2

SAG	Surface area grid
SMILES	Simplified Molecular Input Line Entry System
SSY	sum of the squares of response value
TPSA	Topological polar surface area
UV-Vis	Ultraviolet visible

Tables list

Figures list

Chapter I:

Chapter II:

Chapter III:

Chapter IV:

Graphs list

Chapter III:

Table of contents

Chapter II: Molecular Geometry, Structural, electronic properties and ADME studies

Chapter III: In vitro and in silico evaluation of anti-proliferative activity against breast cancer.

General Introduction

General introduction

Although breast cancer incidence and mortality have tended to decline in recent years, thanks to early detection and increasingly appropriate treatments, breast cancer remains the leading global cause of cancer-related death among women with almost 684 996 breast cancer-related deaths in 2020 [1].

Most women with breast cancer will undergo some type of surgical excision, often combined with other treatments such as radiation therapy, chemotherapy, hormone therapy, and very recently biology therapy [2]. Surgery is meant to remove the tumor from the breast and lymph nodes. It is almost always followed by radiotherapy, which should destroy remaining breast cancer cells. it helps to reduce the risk of cancer recurrence, but it is less effective in prolonging patient survival. Biology therapy refers to the treatment with Herceptin®, a monoclonal antibody, for patients overexpressing a specific "protein" (HER2) [3, 4].

In the case of ER- breast tumours, chemotherapy is then often prescribed to kill cancer cells that could not be removed by other means, or that may have already invaded other parts of the body. However, the molecules used in chemotherapy are highly cytotoxic, and can reduce mortality by up to 20% [5, 6], its physical drawbacks include nausea, fatigue, nerve damage, anemia, hair loss, and a weaker resistance to infections.

The finding of the hormone responsiveness of ER+ breast cancer cases triggered development of antiestrogens, molecules which could compete with estradiol at its receptor and counterbalance its proliferative action. Thus, for the patients with ER+ tumours, an adjuvant treatment with antiestrogens [7, 8], or more recently with aromatase inhibitors, are currently available for their treatment [9]. This hormone therapy appears as a simple and safer alternative to ablative surgery and to undiscriminating radio- and chemotherapy for patients having hormone-dependent breast cancer [7, 10, 11], but, it causes side effects like hot flashes, Vaginal dryness or discharge, tumor flare with bone pain, high calcium level in the blood, and other rare but serious side effects like uterine cancer, blood clots and strokes[12]

Since the binding of estradiol to its receptor promotes breast cancer cell proliferation (in the ER+ tumours), many molecules targeting this protein have been synthesized to counteract the estradiol action. Ferrocene derivatives have proved their efficiency against hormone-dependent breast cancer cells (MCF7). It was found that on MCF-7 cells, the effects of hydroxyferrocifens are quite similar to that of hydroxytamoxifen, slightly more potent. Moreover, it was clarified that the presence of a ferrocene group is necessary, but not sufficient, to generate the anti-proliferative effect [13-15].

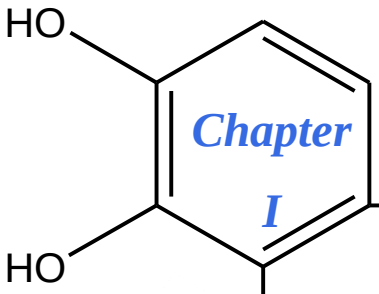
As part of this doctorate thesis, we aim to develop new non-steroidal organometallic compounds for the treatment of breast cancer. First, a study on the molecular structure, spectroscopic and electronic properties of a series of 29 ferrocenylmethylaniline derivatives was carried out using Density Functional Theory (DFT)/B3LYP method with 6-311 ++G(d,p) and LanL2DZ basis sets. The compounds were then tested on breast cancer cell lines (MCF7 and MDA-MB-231) to investigate their anti-proliferative activity using MTT essay. Moreover, molecular docking study was carried out against estrogen and progesterone receptors to understand the efficacy of the investigated compounds against the hormone dependent cells. Next, another docking study was performed with aromatase and 17β-HSD1, the enzymes that catalyse the estrogen synthesis. Finally, the anti-proliferative activity of studied derivatives against MCF-7 cell lines was investigated to predict a Quantitative Structure–Activity Relationship (QSAR) model using molecular descriptors and MLR analysis and the generated model was validated for predictive ability inside the model using Leave One Out method (LOO)

The computational work and the molecular docking study were carried out in the VTRS laboratory of El Oued University, and experimental research work was carried out at Department of Biochemistry, Faculty of Pharmacy at Hacettepe university, Turkey. The manuscript is structured in four chapters, namely:

- The first chapter is a bibliographic overview on ferrocene and its derivatives, breast cancer pathology and treatments, theoretical background on the computational used methods contains the main concepts and definitions related to computational methods was also presented.
- The second chapter contains molecular geometry, structural and electronic property studies on ferrocenylmethylaniline derivatives, where we present the results of a comparative study of two basis (6-311G++(d,p) and LanL2DZ) used in the calculation. Furthermore, the ADME and drug-likeness properties were predicted.
- The third chapter is an *in vitro* and *in silico* evaluation of anti-proliferative activity against breast cancer.
- The last chapter is a development of Quantitative structure–activity relationship (QSAR) model for the anti-proliferative affect against MCF-7 cell lines.

References

1. Organization, W.H. Cancer fact sheets. 2018; Available from: http://gco.iarc.fr/today/fact-sheets-cancers.
2. Society, A.C., Breast cancer facts & figures 2005–2006. 2005, American Cancer Society Atlanta, GA.
3. Hudis, C.A., Trastuzumab—mechanism of action and use in clinical practice. New England Journal of Medicine, 2007. **357**(1): p. 39-51.
4. Ismael, G., et al., Trastuzumab (herceptin) for early-stage breast cancer. Hematology/oncology clinics of North America, 2007. **21**(2): p. 239-256.
5. Hortobagyi, G.N. and A.U. Buzdar, Current status of adjuvant systemic therapy for primary breast cancer: progress and controversy. CA: a cancer journal for clinicians, 1995. **45**(4): p. 199-226.
6. Group, E.B.C.T.C., Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. The Lancet, 2005. **365**(9472): p. 1687-1717.
7. Jordan, V.C., Antiestrogens and selective estrogen receptor modulators as multifunctional medicines. 2. Clinical considerations and new agents. Journal of medicinal chemistry, 2003. **46**(7): p. 1081-1111.
8. Singh, M.N., et al., Tamoxifen: important considerations of a multi-functional compound with organ-specific properties. Cancer treatment reviews, 2007. **33**(2): p. 91-100.
9. Dixon, J.M., Exemestane and aromatase inhibitors in the management of advanced breast cancer. Expert Opin Pharmacother, 2004. **5**(2): p. 307-16.
10. Meegan, M. and D. Lloyd, Advances in the science of estrogen receptor modulation. Current medicinal chemistry, 2003. **10**(3): p. 181-210.
11. Magarian, R., et al., The medicinal chemistry of nonsteroidal antiestrogens: a review. Curr. Med. Chem, 1994. **1**: p. 61-104.
12. society, a.c., treating breast cancer. 2017.
13. Vessieres, A., et al., Modification of the estrogenic properties of diphenols by the incorporation of ferrocene. Generation of antiproliferative effects in vitro. Journal of medicinal chemistry, 2005. **48**(12): p. 3937-3940.
14. Hillard, E., et al., A series of unconjugated ferrocenyl phenols: prospects as anticancer agents. ChemMedChem: Chemistry Enabling Drug Discovery, 2006. **1**(5): p. 551-559.
15. Nguyen, A., Synthèse et évaluation biologique de nouveaux dérivés organométalliques du tamoxifène pour le diagnostic et le traitement du cancer du sein. 2007, Paris 6.



Chapter I

Bibliographic Overview

I.1. Ferrocene and its Derivatives

I.1.1. Introduction

The discovery of ferrocene made by Pauson and Kealy in 1951, [1] and the correction of its sandwich structure by Wilkinson and Woodward [2] marked a great revolution in the field of organic and organometallic chemistry.

In recent years, the applications of ferrocene and its derivatives in the industry have become vast, and researchers have studied their synthesis and their properties.

This first chapter is devoted to a bibliographic review of ferrocene and its derivatives: a history of its discovery and its most important physical and chemical properties and its application.

I.1.2. The discovery of ferrocene

In 1951, Kealy and Pauson published for the first time in a famous article [1] in the journal Nature, the synthesis of bis (cyclopentadienyl) iron or ferrocene. Kealy and Pauson have attempted to prepare dihydrofulvalene by oxidation of a cyclopentadienyl Grignard reagent. The compound obtained from this reaction was an orange thermally stable product. The structure proposed by Pauson for this compound has a resonance form in which iron is bound to cyclopentadienyl by a sigma bond with a canonical ionic formula as shown in figure I.1.

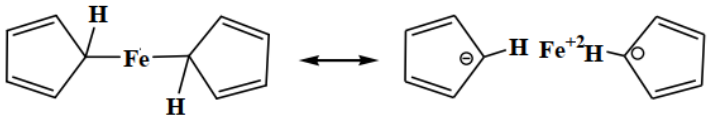


Figure I.1: The resonance structure proposed by Kealy and Pauson in 1951

The correction of the ferrocene structure was done by the English chemist Geoffrey Wilkinson and the German physicist Ernst Fischer, who played an important role in the correction of the structure proposed by Pauson, and the discovery of the exact nature of the carbon-iron bond in bis (cyclopentadienyl) iron, called ferrocene.

The correct structure of ferrocene and its aromaticity has been published rapidly in the journal of the American Chemical Society in the form of a communication [2]. In 1952 E. O. Fischer independently started to study the structure of ferrocene in his laboratory at the Technische Hochschule in Munich. His conclusions were based on X-ray crystallography data, from which he concluded that the molecule should consist of an iron (II) atom located between the two cyclopentadienes as ligands. Fischer and his colleagues immediately synthesized the ferrocenium cation, and began to explore similar molecules, such as cobaltocene[3]. The structures proposed by Wilkinson and his collaborators were truly revolutionary.

Orgel began to describe the molecular orbitals of this new structure, which has already been determined by Dunitz[4]. The title of his publication contained the sandwiched description [5].

Thus, the discovery of the ferrocene sandwich structure was the starting point for the rapid expansion of organometallic chemistry and opened up a whole new field in organometallic chemistry, which led to the Nobel Prize for Wilkinson and Fischer[6].

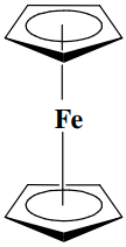


Figure I.2. Structure of Ferrocene.

I.1.3. Properties of Ferrocene

Ferrocene is an orange crystalline solid with a molecular weight of 186.04 $g.mol^{-1}$, boiling point 249 °C and a melting point of 173-174 °C. ferrocene decomposes at 400°C, its solubility in water is about 0.1 mg/ml at 21 °C, 100 mg/ml in DMSO at 19.5°C. Moreover, ferrocene is soluble in dilute nitric acid, benzene, ether, concentrated sulfuric acid but insoluble in concentrated hydrochloric acid. At room temperature, ferrocene is the most stable of metallocenes, with a smell of camphor. Studies have shown that it is sensitive to prolonged exposure to air and light[7].

Infrared spectroscopy of ferrocene is relatively simple because of its symmetrical structure. It has an absorption band at 3075 cm^{-1} equivalents to the stretching of the aromatic C-H bond. There are only four apparent bands: two at 811 and 1002 cm^{-1} are equivalent to the flexion vibration of C-H, and one at 1108 cm^{-1} is equivalent to the asymmetric vibration of the pentadienyl ring. The absorption band at 1411 cm^{-1} is equivalent to the asymmetric stretching of C-C of unsubstituted cyclopentadienyl. It has been observed that mono-substitution has no effect on the positions of these bands, although their intensity has decreased.

The UV-Vis spectrum of ferrocene shows two main wavelengths in 324 and 441nm which indicate that the transition is of the type of (d-d) [8].

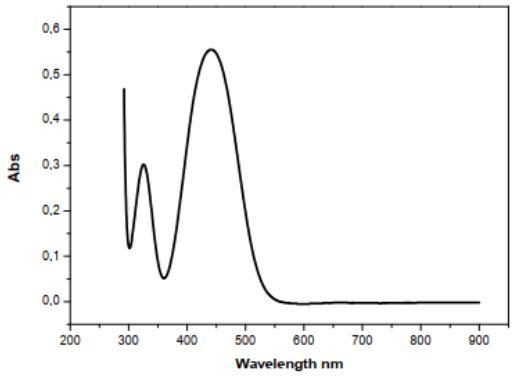


Figure I.3. UV-VIS spectrum of a diluted ethanol solution of ferrocene

The 1H NMR spectrum of the ferrocene molecule is very simple, with a single chemical shift corresponding to the ten protons of the two cycles of cyclopentadiene at 4.15 ppm. Likewise, the ^{13}C NMR spectrum which shows a single chemical shift at 68 ppm correspond to the ten carbon atoms of the two cyclopentadienyl rings.

The average length of the C-C bond in the two pentadienyl rings of ferrocene is 1.389Å, a value which is very close to that of benzene (1.395Å), the Fe-C length is 2.03Å and the C-H bonds have an average length of 1.389Å[9].

I.1.4. Substitution reaction

Substitution reactions are considered to be the most important reaction of ferrocene, whereas, the substituted ferrocenes can be synthetized directly or indirectly. In the case of indirect synthesis, ferrocene is the starting compound for the preparation of various ferrocenic derivatives. While the direct synthesis, the substituted ferrocene is prepared by cyclopentadienyl reactions substituted with iron in simple ways such as reagent Grignard reactions, organometallic compounds, lithium or sodium reagents [10, 11].

The quaternary salt of ferrocinylmethyltrimethylammonium iodide is the main structure of most derivatives based on substitution reactions, which facilitate the nucleophilic substitution include alkoside ions, cyanide, Grignard reagents, carbon reagents, amines[12-14].

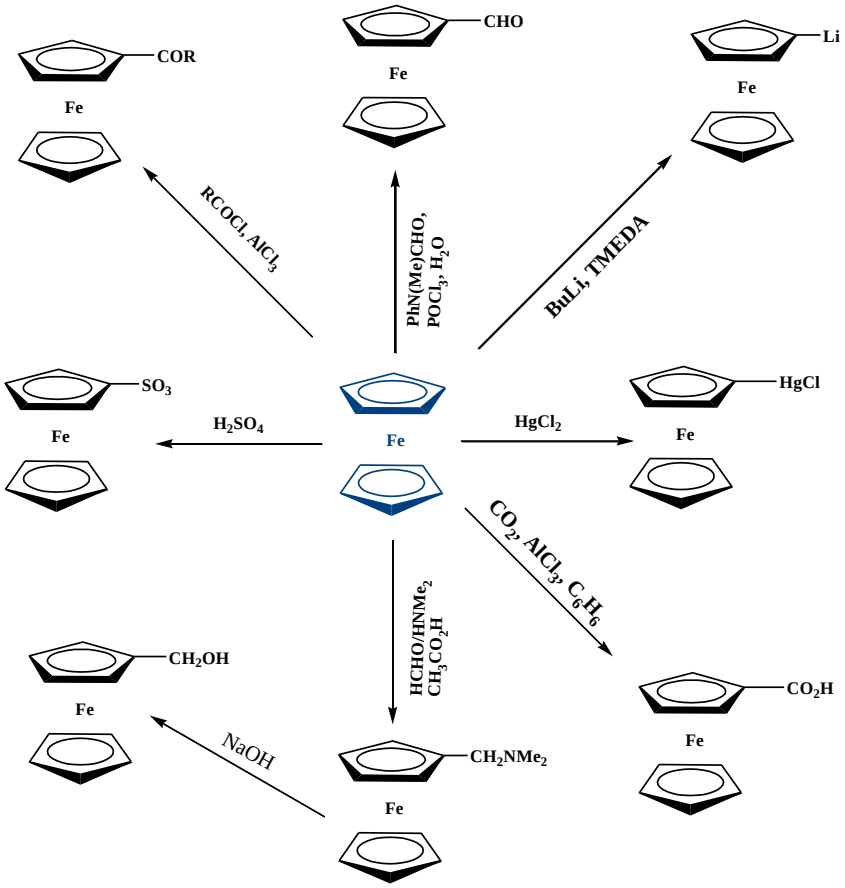


Figure I.4. The important substitution reactions of ferrocene.

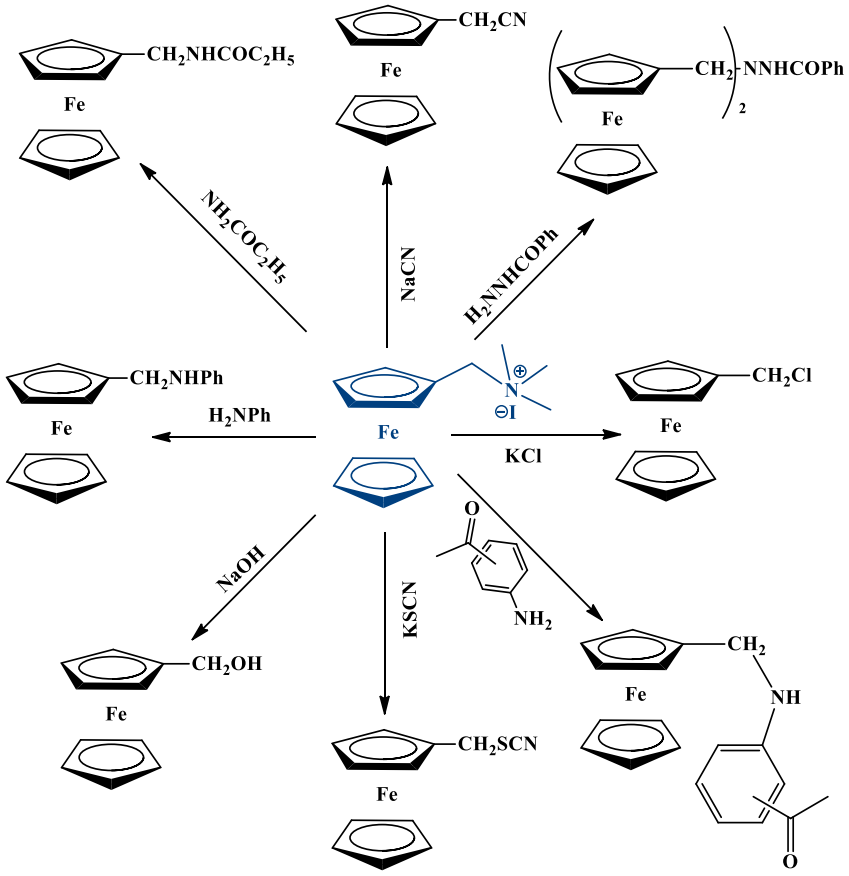


Figure I. 5. The important substitution reactions of the quaternary salt.

I.1.5. Ferrocene applications

I.1.5.1. Ferrocene derivatives in medicinal chemistry

Ferrocene derivatives have attracted significant interests in medicinal chemistry as anticancer[15, 16], anti-oxidant[17], antibacterial[18], antifungal[19] and antiparasitic drug candidates[20].

The relative stability of ferrocene in biological media has encouraged their vectorisation with bioactive compounds. Edwards *et al.* synthesized ferrocenyl antibiotics, which were active against penicillin resistant bacteria in the 1970s. In vivo toxicology studies on ferrocene derivatives disclosed low levels of toxicity, despite liver-related problems. Developed in the former USSR for the curing of iron-deficiency anemia, a sodium salt of *o*-carboxybenzoyl ferrocene is well tolerated for oral administration, and has also be prescribed for gum diseases.

In the 1990s, the two most well-known derivatives were discovered, ferroquine and ferrocifens, which have been studied extensively for the treatment of malaria and breast cancer, respectively. In 1997, Brocard and co-workers inserted a ferrocenyl group into the side chain of the chloroquine and it has been reported that the resulting compound ferroquine is much safer and more effective in mice, as well as non- mutagenic. It is not only active against chloroquine-sensitive bacteria, but also against chloroquine-resistant strains[18, 21-24].

Antiproliferative activity of ferrocifens and hydroxyferrocifens was evaluated on MCF-7 cells, which are hormone-dependent breast cancer cells having an important concentration of ER, and on MDA-MB-231 cells, which are classified as hormone-independent breast cancer cells because devoid of ERα. It was found that on MCF-7 cells, the effects of hydroxyferrocifens are quite similar to that of hydroxytamoxifen, slightly more potent at a concentration of 0.1 µM, and definitely superior at 1 µM. Furthermore, the novel feature of ferrocifens lies in the strong cytotoxic effect observed on MDA-MB231 cells, which are hormone-independent breast cancer cells. In these cells without ERα, OH-tamoxifen has no effect, while ferrocifens have a remarkable antiproliferative effect with IC50 = 0.5 µM[25].

The major developments in the medicinal applications of ferrocene and its derivatives, from its discovery and biological testing to their clinical testing, are summarized in figure I. 6.

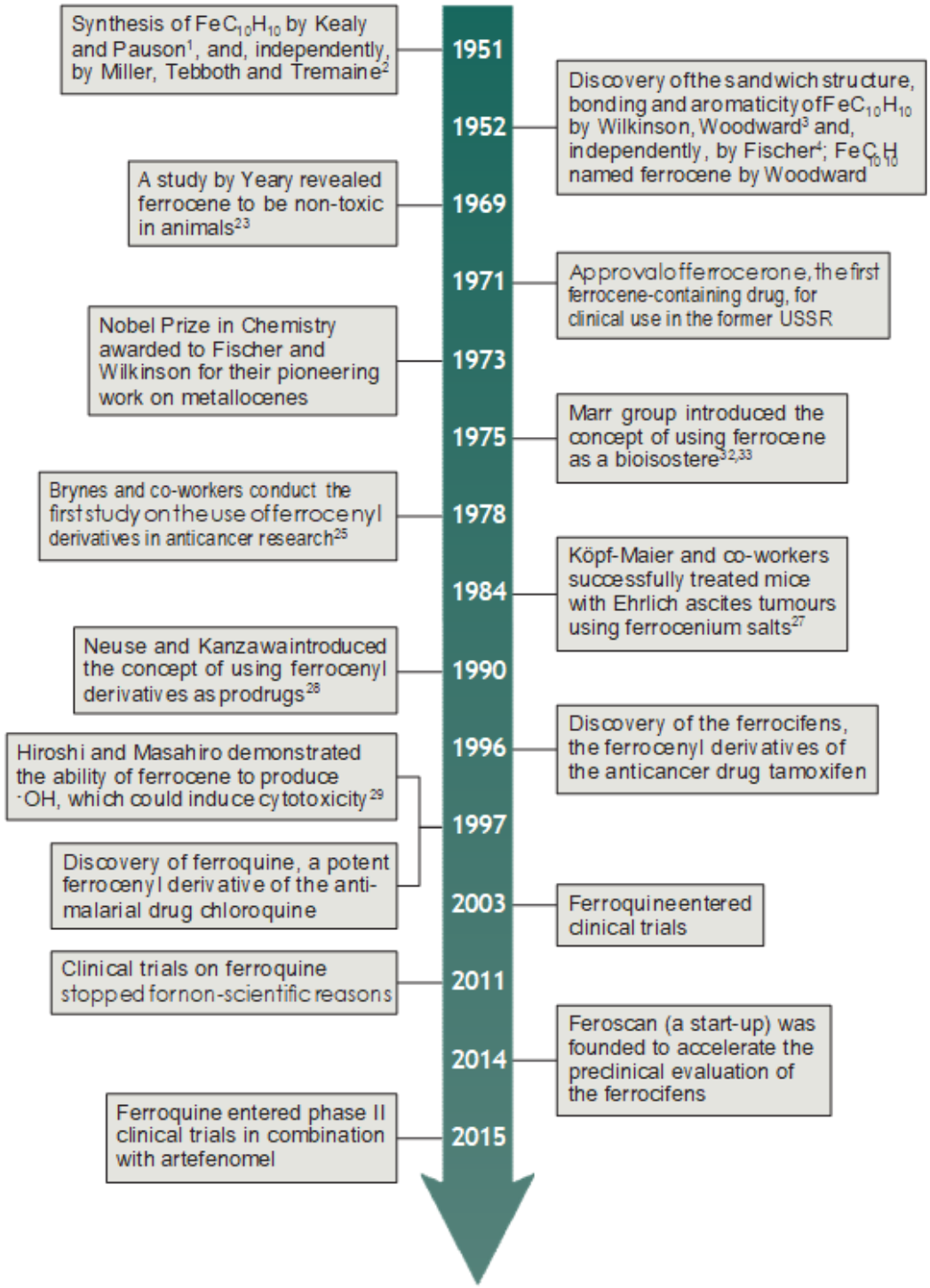


Figure I. 6. Ferrocene derivatives in medicinal applications[24].

I.1.5.2. Other ferrocene applications

The applications of ferrocene in chemistry are indeed vast. New complex catalysts using ferrocene derivatives are produced daily. From polymers to catalysts to chiral compounds, ferrocene has attracted the interest of many chemists. The possibilities of using this compound are unlimited; so, there are many areas of research on ferrocene that have not yet been explored. Ferrocene has proven to be one of the most useful and intriguing compounds in organometallic chemistry and has brought a great change to the chemical community.

Here are some of the uses of ferrocene in the field of industrial chemistry[9]:

- It can be used as a protective factor in polyethylene.
- Some ferrocene derivatives can be used in the protection of the thin layers of photoelectrochemical poles.
- The addition of ferrocene fixes the polymerization of the cellulosic material and gives it a certain rigidity, against heat and air.
- Some ferrocene derivatives are used as insecticides.
- Ferrocene is used in the improvement of thermal stability of plastic and rubber.

I.2. Breast cancer

I.2.1. Generality

The breast refers to the organ containing the mammary gland (Figure I.7). The gland is surrounded by a large adipose connective tissue, and is supported by Cooper's ligaments. It is composed of lobules, and galactophores ducts. The lobules produce the milk during breastfeeding and the ducts transport the milk to the nipple. The mammary tissue undergoes variations during the life (embryo, puberty, reproduction, menopause) under the hormonal influence. At birth, the mammary gland is unfinished and remains at rest until puberty. At this stage, the mammary gland develops under the influence of two hormones: estrogens, which stimulate the growth and ramifications of the canal system and the proliferation of adipose tissue, and progesterone which works with estrogens to differentiate the lobular structures. The first pregnancy leads to the maturation of the mammary gland and an increase in the secretory activity of the lobules which regresses 3 to 4 months after stopping lactation. At menopause, involution of the mammary gland increases with atrophy of lobular cells[26].

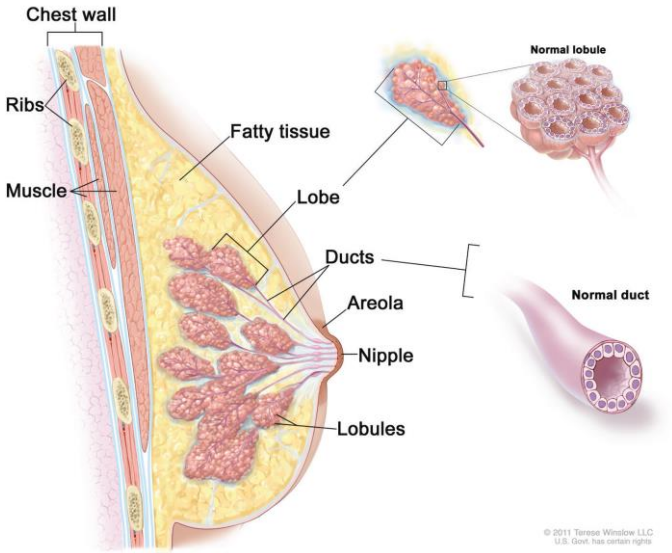
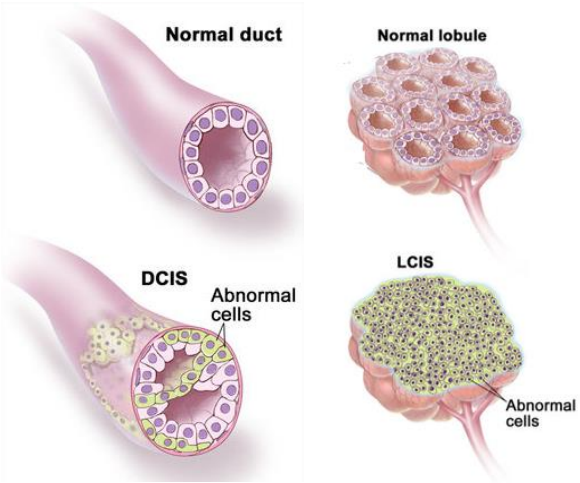


Figure I. 7. Anatomy of female breast[27].

Breast cancer is the most common cancer in women, in front of colorectal cancer and lung cancer, with more than 2 million new cases worldwide in 2020. According to the World Health Organization, breast cancer is the leading global cause of cancer-related death among women (6.9 %) with almost 684 996 breast cancer-related deaths in 2020 [28]. Although breast cancer incidence and mortality have tended to decline in recent years, thanks to early detection and increasingly appropriate treatments, breast cancer remains the leading cause of cancer death in women [29].

I.2.2. Types of breast cancer

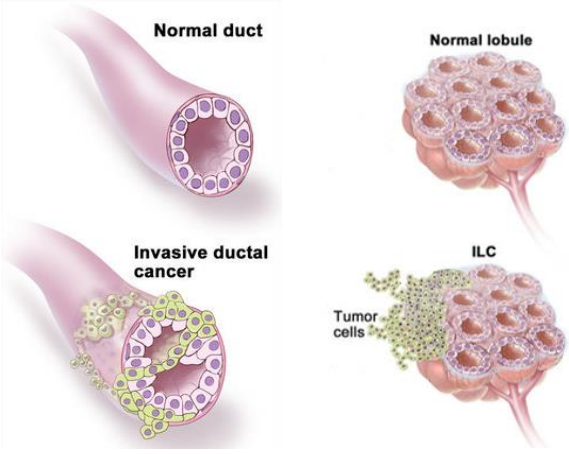
Breast cancer is a malignant tumor of the mammary gland. Classically, there are two categories of breast carcinoma categorized according to their level of tumor infiltration[30], non-infiltrating or in situ carcinomas, and invasive carcinomas. In addition, there are several histological types: ductal cancers, which affect the milk ducts, and lobular cancers that affects lobules. These are the two most common histological types, however there are other types of rare lesions[26].



(a) (b)

Figure I. 8. In situ carcinoma in ductal (a) and lobular (b)[27].

If the cancer in situ is not treated in time, the tumor cells can cross the basal membrane of the canal or lobule, spread in the axillary lymph nodes and then throughout the body and form metastases: the cancer becomes invasive.



(a) (b)

Figure I. 9. Invasive carcinoma in ductal (a) and lobular (b)[27].

Ductal cancer accounts for 75% of invasive cancers, however, lobular cancer represents only 5 to 10% of infiltrating cancers and other even rarer forms exist such as tubular, cribrous, medullary, mucinous or papillary carcinomas.

I.2.3. Risk factors

Risk factors can be intrinsic, that means, specific to the individual (gender, age, genetic inheritance, etc.) which cannot be modified. There are also extrinsic risk factors (environment, lifestyle, include decisions about having children and taking medicines that contain hormones.) that can sometimes be used to reduce the risk of cancer [31, 32].

- **Gender:** Being a woman greatly increases the risk of breast cancer. In fact, less than 1% of breast cancers affect men.
- **Age:** As you get older, the risk of breast cancer increases. Most breast cancers are found in women in menopause (55 and more).
- **Genetic inheritance and family:** About 5% to 10% of breast cancer cases are thought to be inherited, it means that they result directly from mutations passed on from a parent, also having close blood relatives with breast cancer makes a higher risk.
- **Hormonal exposure:** Several studies have shown that the risk of breast cancer increases by 10 to 20% by early puberty (before 12 years) and 20% by late menopause (after 55 years) [33-35]. Nulliparity or a late first pregnancy is also thought to be responsible for prolonged exposure to estrogen-promoting breast cancer [36]. Breastfeeding and multiparity, on the other hand, are thought to be protective factors and would reduce the risk of breast cancer [37, 38].
- **Lifestyle:** related to personal behaviors, such as diet, alcohol, obesity, and exercise. Other lifestyle-related risk factors include decisions about having children, breastfeeding, and taking medicines that contain hormones.

I.2.4. Treating Breast Cancer

According to the American Cancer Society [39] the treatment of breast cancer can be local treatments where they treat the tumor without affecting the rest of the body by surgery and radiation, or systematic treatment which can reach cancer cells almost anywhere in the body. Depending on the type of breast cancer, different types of drug treatment might be used, including chemotherapy, hormone therapy, targeted therapy and immunotherapy. In this part we are focusing on the targeted therapy and hormone therapy.

I.2.4.1. Targeted therapy (anti-HER2)

Approximately 15 to 20% of breast cancers show amplification of the human epidermal growth factor gene (HER2) or overexpression of the protein. The HER2 protein is located on the surface of breast cells and is therefore involved in breast growth. In some cases of breast cancer, the HER2 protein is overexpressed on the surface of cancer cells, referred to as HER2 positive tumours.

Before the advent of targeted anti-HER2 therapies, women with early-stage HER2-positive breast cancer presented a worse prognosis than those with HER2-negative tumors, with shorter relapse-free survival and high incidence metastases and increased mortality. The main advantage of targeted therapy is their targeted action, which limits side effects [40, 41].

❖ **Lapatinib (Tykerb)**

Lapatinib is a small molecule, inhibitor of the HER2 receptor tyrosine kinase site in its intracellular portion. It is used to treat advanced breast cancer, and might be used with certain chemotherapy drugs or hormone therapy drugs[42].

❖ **Trastuzumab (Herceptin)**

It is a humanized monoclonal antibody targeting the extracellular domain of the HER2 receptor. It has demonstrated significant efficacy in combination with metastatic phase chemotherapy. It can be used to treat both early-stage and advanced breast cancer.

❖ **Pertuzumab (Perjeta)**

It is a humanized monoclonal antibody targeting the HER2 receptor on an epitope different from trastuzumab. It prevents the dimerization of the HER2 and HER3 receptors, which are heterodimers particularly active in HER2 signaling. It can be given with trastuzumab and chemo, either before or after surgery to treat early-stage breast cancer, or to treat advanced breast cancer[41].

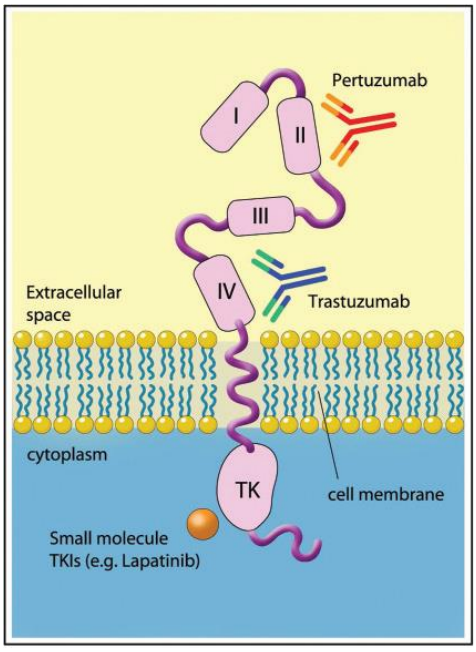


Figure I. 10. The HER-2 receptor and binding sites for Trastuzumab, Pertuzumab and Lapatinib[43].

The side effects of these medicines are often mild, but some can be serious. Some patients develop heart damage while or after treatment with trastuzumab or pertuzumab. This can lead to congestive heart failure. Lapatinib and the combination of pertuzumab with trastuzumab can

cause severe diarrhea. Lapatinib can also cause hand-foot syndrome, in which the hands and feet become sore and red, and may blister and peel[39].

I.2.4.2. Hormone therapy

Hormone sensitive tumours, i.e. those that express estrogen and/or progesterone receptors, account for about 70% of breast cancers. In these patients, hormones are involved in stimulating tumor growth and an indication for hormone therapy may be proposed to inhibit this growth of cancer cells.

The main indication for hormone therapy is to inhibit the production of estrogen or to block, by competition, its action on the receptors. In addition, the choice of hormone therapy depends particularly on the patient's menopausal status since estrogen production by the ovaries is stopped after menopause and the level of circulating estrogens then comes essentially from a conversion of androgen by aromatase. Thus, three main types of antihormonal treatments can be proposed: aromatase inhibitors, antiestrogens and pituitary gonadotropin releasing hormone analogs or Luteinizing Hormone Releasing Hormone (LH-RH)[44].

❖ **Aromatase inhibitors**

Aromatase is an enzyme mainly present in adipose tissue and allows the conversion of androgens, produced by the adrenals, into estrogens. The role of aromatase inhibitors is therefore to inhibit the action of this enzyme to prevent this conversion and therefore inhibit tumor growth. In postmenopausal patients, the main source of estrogen remains that resulting from the conversion of androgens since the ovaries no longer secrete it. Thus, aromatase inhibitors are only indicated in postmenopausal patients with hormone-sensitive tumours. The aromatase inhibitors on the market and used are the 3rd generation: Letrozole, Anastrazole and Exemestane [45-47]. Aromatase inhibitors can cause muscle pain and joint stiffness and/or pain, as well as bone thinning, sometimes leading to osteoporosis and even fractures. Also, they can very rarely cause blood clots[39].

❖ **Antiestrogens**

Antiestrogens are molecules that have a variable role depending on the site where they are used, in fact they can have both the properties of an agonist (in the uterus, bones and liver) and an antagonist (in the breast, vagina and central nervous system).

The anti-tumor action is conferred by the antagonistic properties of the molecule, which is based on competitive inhibition; these are therefore molecules that will bind to the estrogen receptors present on the tumor cells of the breast and therefore block the action of hormones on the cell. The most commonly used molecule remains tamoxifen, also known as a selective estrogen receptor modulator (SERM). It can be used to treat women with breast cancer who have or have not gone through menopause[48]. Tamoxifen can cause hot flashes, vaginal

dryness or discharge, tumor flare with bone pain, high calcium level in the blood, and other rare but serious side effects like uterine cancer, blood clots and strokes[39].

❖ **Luteinizing hormone-releasing hormone (LHRH) analogs**

LHRH analogs can be used only in pre-menopausal women to shut down the ovaries which are the main source of estrogen. They are molecules that aim to bind to LHRH receptors in such a way to cause hyperstimulation. The pituitary gland will then initially increase its secretion of LH, resulting in an increased secretion of estrogens (and progesterones), then the pituitary gland will no longer respond to pituitary stimulation, which will lead to a decrease or complete inhibition of estrogen secretion by the ovaries. It is therefore necessary to induce inhibition of ovarian function in order to stop the stimulation of tumour growth by estrogens[49]. The most common toxicities are mood disorders, vaginal dryness, hot flashes and osteoporosis [50].

I.3. Theoretical background of computational methods

I.3.1. Molecular modeling

I.3.1.1. Generality

The research and synthesis of new chemical and biochemical compounds are today often associated with a molecular modeling study. Computational chemistry (or molecular modeling) could be defined as the application of mathematical and theoretical principles to solve chemical problems, it exploits the laws of chemistry, physics and biology in specific computer programs to calculate structures and properties of chemical and biochemical entities (proteins, nucleic acids, molecular complexes, solids, crystals etc.). Molecular modeling aims to predict the structure and reactivity of molecules or molecule systems. The most precise molecular models use *ab initio* which mean ″first principles″ electronic structure methods, based on the principles of quantum mechanics, and are commonly very computationally expensive. However, due to the development in computer resources and storage capacity and processor performance, molecular modelling has been a rapidly progressing and expanding field, to the point that it becomes currently possible to solve relevant problems in an acceptable amount of time[51, 52]. It serves as a bridge between theory and experiment to[53]:

1. Extract results for a particular model.
2. Compare experimental results of the system.
3. Compare theoretical predictions for the model.
4. It helps in the comprehension and interpretation of experimental observations.
5. Correlate between microscopic details at atomic and molecular level and macroscopic properties.
6. Provide information that is not available from real experiments.

In the following, we are describing the most commonly used methods in chemistry for the study of organometallics as well as the molecular docking method.

I.3.1.2. Molecular calculation methods

All molecular calculation methods can be classified under three general categories, quantum mechanical methods, semi-empirical methods and molecular mechanics, they all used to determine the geometry of a molecule as well as its physicochemical properties[52].

I.3.1.2.1. Quantum mechanical methods

It provides the most accurate and consistent predictions for chemical systems. However, its methods are extremely computer intensive.

I.3.1.2.1.1. Basics of quantum mechanics

- **Schrödinger Equation**

The equations developed in quantum chemistry are based on solving the Schrödinger equation. This fundamental equation was proposed in 1925 by the physicist Erwin Schrödinger, and earned him the Nobel Prize in physics in 1933. It describes the movement of electrons and nuclei in atoms and molecules, as well as all the properties that result from them. The time-independent and the time- dependent Schrödinger equation are written in the following form (1) and (2) respectively:

$$\Psi\hat{\mathrm{H}} = E\Psi \quad \text{..... (Eq I.1)}$$

$$\frac{\partial \Psi}{\partial t} = -\frac{i}{\hbar}\hat{\mathrm{H}}\,\Psi \quad \text{..... (Eq I.2)}$$

Here the wave functions Ψ and Ψ are functions of all coordinates of the relevant system and Ψ is also a function of time. In our case of a molecular Hamiltonian $\hat{\mathrm{H}}$ is given by:

$$\hat{H} = \hat{T}_e + \hat{T}_n + \hat{V}_{ne} + \hat{V}_{ee} + \hat{V}_{nn} \ldots\ldots\ldots \text{(Eq I.3)}$$

$$\hat{\mathrm{H}} = -\frac{\hbar^2}{2}\sum_I \frac{1}{m_I}\nabla_I^2 - \frac{\hbar^2}{2m_e}\nabla_i^2 + \sum_I \sum_{J<I} \frac{Z_I Z_J e^2}{r_{IJ}} - \sum_I \sum_i \frac{Z_I e^2}{r_{iI}} + \sum_j \sum_{i>j} \frac{e^2}{r_{ij}} \quad \text{.... (Eq I.4)}$$

Where $\hat{T}_e$ and $\hat{T}_n$ are the kinetic energy terms for the electrons and the nuclei respectively; $\hat{V}_{ne}$ represents the electron-nuclei interaction potential, also $\hat{V}_{ee}$ and $\hat{V}_{nn}$ the interelectronic and inter-nuclear repulsion potentials.

- **The Born-Oppenheimer Approximations:**

One of the first approximations applied is the approximation proposed by Born and Oppenheimer in 1927. It consists in decoupling the movements of the electrons from those of the nuclei. Indeed, the electrons, having masses much smaller than those of the nuclei (mp / me

≈ 1836), move much more quickly. In the description of the electronic states, the nuclei can be considered as immobile. This conjecture leads to a simplification of the Schrödinger equation into an electronic part, whose solution provides the potential for the movement of nuclei, which will then be used to solve the Schrödinger equation for the nuclei alone. In the electronic equation, the positions of the nuclei play the role of simple parameters that are fixed. Therefore, in the Eq. I.5 below, which defines the electronic Hamiltonian ($\hat{H}_{el}$), the kinetic energy of the nuclei ($\hat{T}_n$) is not considered and the term associated with the nuclear repulsion potential ($\hat{V}_{nn}$) is a constant:

$$\hat{H}_{el} = \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} + \hat{V}_{nn} \quad \text{.....(Eq I.5)}$$

$$\hat{H} = -\frac{\hbar^2}{2m_e}\nabla_i^2 + \sum_I \sum_{J<I} \frac{Z_I Z_J e^2}{r_{IJ}} - \sum_I \sum_i \frac{Z_I e^2}{r_{iI}} + \sum_j \sum_{i>j} \frac{e^2}{r_{ij}} \quad \text{..... (Eq I.6)}$$

Thus, the Born-Oppenheimer approximation leads us to calculate the electronic energy (E_{el}) through the resolution of:

$$\hat{H}_{el}\Psi_{el} = E_{eff}\Psi_{el} \quad \text{.....(Eq I.7)}$$

For a given nuclear configuration, the resolution of the Eq I.7 gives direct access to the quantized electronic states of the studied system as well as the energy associated with these states. The function describing the electronic energy according to the positions of the nuclei constitutes the hypersurface of potential energy (SEP).

Despite the simplification introduced by the Born-Oppenheimer approximation, the resolution of the Eq I.7 It is only possible for the simplest systems (*eg*. the hydrogen atom, the H_2^+ molecule). Much of the work of theoretical chemists has been to find approximate resolution methods of this equation for the most complex systems.[54-56]

I.3.1.2.1.2. *Ab initio* methods

The codes used in molecular chemistry are based on numerous and different quantum chemistry methods that allow the resolution of the Schrödinger equation associated with the molecular Hamiltonian. Methods that do not include any empirical or semiempirical parameters in their equations, which mean that derive directly from theoretical principles, without the inclusion of experimental data, are called *ab initio* methods.

This does not imply that the solution obtained is the exact solution; they all consist of approximations to varying degrees of quantum mechanics calculations. This means that a particular approximation is rigorously defined on the first principles (quantum theory) and then

resolved with a margin of error that is known qualitatively to advance. If digital iteration methods are used, the goal is to iterate until machine accuracy is reached.

The simplest ab-initio method of electronic structure calculation is the Hartree-Fock (HF) scheme, in which the electron-electron Coulomb repulsion is not specifically taken into account. Only its average effect is included in the calculation. When the size of the base is increased, the energy and the wave function tend towards a limit called Hartree-Fock limit. Many types of calculations, known as post-Hartree-Fock methods, begin with a Hartree-Fock calculation and are then corrected for electron-election repulsion, also known as electronic correlation.

When these methods are pushed to their limits, they approach the exact solution of the non-relativistic Schrödinger equation. If one wants to get an exact agreement with experience, it is necessary to include relativistic and spin-orbit terms, both of which are important only for heavy atoms. In all these approaches, in addition to the choice of the method, it is necessary to choose an adequate basis. This base (in the mathematical sense of the term) is a set of functions, usually centered on the different atoms of the molecule, that are used to extend the molecular orbitals by the linear combination of atomic orbitals (LCAO) postulate. The ab initio methods therefore need to establish a level of application of the theory (method) and a basis.[57]

I.3.1.2.1.3. Hartree-Fock Method

The Hartree-Fock method makes it possible to approximate the Schrödinger equation, by assimilating the bi-electronic integrals into a Coulomb interaction integral (denoted by J) and an exchange integral (denoted by K). We thus define the new operator F (Fock operator):

$$\hat{F}\varphi = E\varphi \quad \ldots\ldots \textit{(Eq I.8)}$$

$$\hat{F} = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\sum_{i=1}^{N}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j\rangle i}^{N} J_{ij} - K_{ij} \quad \ldots\ldots \textit{(Eq I.9)}$$

The shapes of the integrals are known, so the equation can be solved. However, the wave function depends directly on the J and K parameters, which also depend on the wave function directly. Thus, a linear resolution of the equation is to be excluded. The method used is that of the self-consistent field SCF, which consists in solving the equation from a starting set of orbitals, called the basis set. This starting game makes it possible to calculate a first time the parameters J and K, which are reused to recalculate the orbitals, and so on until convergence of the various parameters (energy, RMSD ...)

This starting orbital game can be more or less developed according to our expectations. The larger the base, the more precise the calculation will be, and therefore longer. The orbitals are

either represented by Slater-type orbitals (STO) or by Gaussian functions ($e^{-\alpha r^2}$), representing in a different way the core orbitals and the valence orbitals.

This approximation and the self-consistent field method allow the Hartree-Fock protocol to quickly solve the Schrödinger equation. However, the previous approximations mean that the calculated energy will always be greater than the energy of the exact solutions (the difference is called correlation energy). From a physical point of view, the position of an electron at a given time is actually not independent of others. It is said that the electrons are correlated, and this can induce a very important difference between the HF energy and the exact energy of a system [58, 59].

I.3.1.2.1.4. Density-Functional Theory (DFT)

The density functional theory (DFT) differs from the ab initio methods based on the HF equations because it is based on the notion of electronic density ρ(r) and not on that of multi-electronic wave function. This approach is based on two theorems due to Hohenberg and Kohn[60].

The first stipulates that "the total energy of a system in its fundamental state depends only on its electronic density ρ(r)" (energy is therefore a functional of the electronic density). Any property of the ground state, including energy, can be written as a functional electronic density.

$$E = E\left[\rho(r)\right] \quad \ldots\ldots\ldots \textit{(Eq I.10)}$$

The second theorem is the analog of the variational principle applied to the density and shows that the density is stationary for the ground state.

❖ **Types of functionalities exist**

DFT calculations fall into three general categories: local density approximations (LDA), generalised gradient approximations (GGA), and ″hybrid″ combinations of DFT and Hartree-Fock terms.

- local density approximations (LDA) that depend solely on the electronic density at each point of the system and neglects any influence of the inhomogeneity of the system. This approximation is correct when the density varies sufficiently slowly[61].
- Generalised gradient approximations (GGA) functionalities that introduce the density gradient into their expression allow for the inhomogeneity of the electronic distribution. The most used functionalities are: Becke88[62], PW91[63], P86 [64] and LYP [65].
- Hybrid functional HF-DFT, which appeared recently, which include for the exchange energy a mixture Hartree-Fock and DFT while the correlation energy

remains purely DFT. These features appear to be the most reliable of the moment. Examples of hybrid methods are B3LYP and B3PW91, where B3 denotes Becke's three-parameter hybrid functional [62, 66], while 'PW91' and 'LYP' are gradient-corrected correlation functionals of Perdew and Wang [67] and, as above, Lee, Yang and Parr.

- **Success and limitations of the DFT**

Several studies have been carried out in recent years, using DFT calculations and giving good results on large chemical systems, taking into account the effects of electronic correlation. Many properties (molecular structures, vibration frequencies...) are well reproduced. However, the DFT method still suffers from several defects, including the lack of real criteria that improve the functional and molecular properties[68].

I.3.1.2.2. Semi-empirical methods

In the ab initio methods, almost all the computation time is consumed by the integral's calculations, and in order to reduce this computation time, it is necessary to simplify the Roothann equations.

A semi-empirical method is a method in which part of the computations required for the Hartree-Fock calculations are replaced by parameters adjusted to experimental values (the Hamiltonian is always parametrized by comparison with references). In general, all these methods are very accurate for families of data products close to those used for parametrization.

Semi-empirical methods considering only the electrons of the valence layer; the electrons of the inner layers are included in the nuclear core[69].

- **CNDO approximation:** (Complete Neglected of Differential Overlap) First semi-empirical method, it was proposed by Pople, Segal and Santry in 1965. This method is based on simplifying assumptions much more severe. It takes its name from the fact that all the bielectronic integrals that depend on the overlap and charge densities between different atoms are zero.
- **MNDO approximation:** (Modified Neglect of Diatomic Overlap) Proposed by Dewar and Thiel in 1977[70]. Method based on the NDDO [71] (Neglect of Diatomic Differential Overlap) approximation which consists in neglecting the differential overlap between atomic orbitals on different atoms. This method does not treat transition metals and presents difficulties for conjugated systems.
- **AM1 approximation:** (Austrin Model 1) [72] Proposed by Dewar in 1985. He attempted to correct the defects of MNDO.

- **PM3 approximation:** (Parametric Method 3) [73]Proposed by Stewart in 1989. Has many points in common with AM1.
- **PM6 approximation:** (Parametric Method 6) The most recent method proposed by James J. P. Stewart in 2007 [74] used in quantum chemistry, rewritten from the ground up with a new more precise parameterization for all major elements and transition metals. More than 9000 compounds were used to develop the new PM6 method from PM3 and PM5. This compares to only 39 compounds used at MNDO, about 200 compounds used at AM1, and about 500 compounds used in PM3.

I.3.1.2.3. Molecular mechanics (MM)

Molecular mechanics (MM) methods use the laws of classical physics for the prediction of structures and their properties, such as energy. Just as quantum mechanics methods is a powerful tool used to better understand small molecular systems and to answer some questions about the electronic structure of molecules. Molecular mechanics, on the other hand, describes molecules according to their nuclei and not their electrons. It defines a purely empirical function that takes the positions of nuclei as implicit and implicitly implicates the presence of electrons.

Molecular mechanics (figure I.11) considers a molecule as a series of masses (atoms) connected together by springs (chemical bonds). Depending on the interactions between the masses, structural deformations can take place. More particularly, for all the methods, the energy function is generally described in internal terms directly related to the covalent bonds of the atoms (bonded interactions) and in external terms reflecting the interactions between non-covalently bonded atoms (non-bonded interactions)[75].

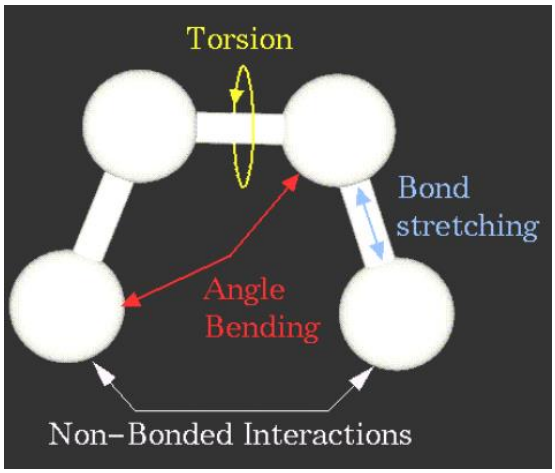


Figure I. 11. Molecular Mechanics

I.3.1.2.4. Force Field

In the functional expression of classical force fields, the total potential energy of the system (E_{tot}) is expressed as the sum of the energy of the bonded and non- bonded atoms [56].

$$E_{tot} = \underbrace{E_{stretch} + E_{bend} + E_{tors}}_{E_{bonded\ atoms}} + \underbrace{E_{VanDerWaals} + E_{electro} + \cdots}_{E_{non\text{-}bonded\ atoms}} \quad .. \ (Eq\ I.11)$$

Where:

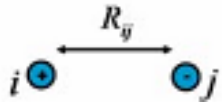
$$E_{str} = \sum_{bonds} k_r(r - r_{eq})^2 \qquad (Eq\ I.12)$$

$$E_{bend} = \sum_{bonds} k_\vartheta(\vartheta - \vartheta_{eq})^2 \qquad (Eq\ I.13)$$

$$E_{tors} = \sum_{bonds} \frac{V_n}{2}[1 + cos(n\Phi - \gamma)] \qquad (Eq\ I.14)$$

$$E_{VdW} = \sum_{i<j} \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} \qquad (Eq\ I.15)$$

$$E_{electro} = \sum_{i<j} \frac{q_i q_j}{\varepsilon R_{ij}} \qquad (Eq\ I.16)$$



There are many force fields with successive versions and many ways to define potential energy. Here are some of the best-known force fields[52, 76]:

- MMFF (Merck Molecular Force Field).
- CFF (Consistent Force Field).
- CVFF (Consistent Valence Force Field), used in the treatment of small organic molecules and biomolecules.
- GROMOS (Groningen molecular Simulation Program Package).
- AMBER (Assisted Model Building with Energy Refinement).
- CHARMm (chemistry at Harvard using molecular mechanics), specialized in the treatment of biological molecules, such as nucleic acids and proteins

I.3.1.3. The bases of atomic orbitals

Atomic orbitals are defined from basic functions that are divided into two major families.

I.3.1.3.1. Slater-Type Orbitals STO

The general expression of a function of Slater [77, 78]is given by the following relation:

$$\chi^{s}_{n',1,m}(r,\theta,\varphi) = N.r^{n-1}.e^{-\zeta r}.Y(\theta,\varphi) \quad \text{….. (Eq I.17)}$$

Where ζ is the exponent of Slater which is a positive number that depends on the atomic number and can be determined using Slater's empirical rules or by optimization. *n, l, m* are respectively the principal, secondary and magnetic quantum numbers. rA, θA, φA are the spherical coordinates that locate the electron with respect to the center A bearing the STO and *Y (θ, φ)* represents a spherical harmonic. In Cartesian coordinates.

I.3.1.3.2. Gaussian-Type Orbitals GTO

The use of Gaussian functions in quantum calculations has been proposed for the first time by Boys [79]. Their standardized general form is as follows:

$$\chi^{G}_{ijk}(x,y,z) = N.x^{i}.y^{j}.z^{k}\exp(-\alpha r^{2}) \quad \text{…… (Eq I.18)}$$

The Gaussian bases, on the other hand, have a rather poor representation of the atomic orbitals because they do not have the exact behavior at the origin (non-zero derivative) nor at the great distances (decay too fast with r) on the other hand their interest is that all the integrals involved in the calculations can be calculated explicitly without recourse to numerical integration. To compensate for this incomplete representation of atomic orbitals by Gaussian functions, linear combinations of Gaussians are used as basic functions. These functions are called gaussian contracted functions. There are a multitude of possible Gaussian bases to perform a SCF calculation the most commonly used are those developed by Pople et al.

- The simplest basis is STO-3G still called minimal base. this means that Slater-type orbitals are represented by three Gaussian functions.

- The next level developed by Pople includes split-valence bases such as 3-21G, 4-31G and 6-31G where the first number represents the number of Gaussians used to represent orbital 1s. Valence orbital are represented by two functions which are composed of the numbers of Gaussians given in the second part of the two numbers of the denomination of the basis.
- For the reduction of the valence layer, several sets of orbitals are used for each valence sub-layer. It can be split (double zeta) "CC-PVDZ" using, for example, for the atoms of the second period, two 2S orbitals and two sets of 2P orbitals. Even better bases are of triple zeta quality "CC-PVTZ", quadruple zeta "CC-PVQZ"[68].
- The Los Alamos National Laboratory basis sets, known as LANL2DZ and developed by Hay and Wadt,[80-83] have been widely used in quantum chemistry, particularly in the study of compounds or clusters containing heavy elements. These basis functions have been obtained by fitting procedure of pseudo-orbitals with Gaussian functions.

I.3.1.3.2.1. The polarization functions

They correspond to unoccupied orbitals in the ground-state atom. Therefore, they will be of the type p for the hydrogen atom, d and g for the atoms of the second period, and of the type f for transition metals. In order to increase the flexibility of the base used and take into account the deformations of the valence atomic orbitals during the formation of the molecule. Their role in the base is decisive [68].

With regard to nomenclature, in addition to the symbol of the basis that's used, the polarization is indicated by one (or two) asterisk (s) (*) or the nature of the added orbitals (p, d, f, ...) (6-31G** = 6-31G (d, p)).

I.3.1.3.2.2. Diffuse functions

In order to better describe the changes in the distribution of the long-range electronic charge, it is possible to add diffuse orbitals which represent Gaussian functions (usually p or d) of very low α exponents which decrease slowly when we move away from the nucleus. Thus, mathematical functions are needed to describe the situation where the external electron tends to move away from the nucleus. These diffuse functions are necessary to correctly describe the anions, the hydrogen bond, the weak interactions etc[55].

With regard to nomenclature, they are indicated by one (or two) sign (s) + in addition to the symbol of the basis used (6-31G++ (d, p)).

I.3.2. Molecular docking

I.3.2.1. General principle of docking protein/ligand

Molecular docking is an empirical method that can expect the affinity between two molecules. It is generally used to predict the affinity of a ligand for a protein, among others, the position and the most favorable orientation for a ligand interacting with a target protein. This tool is currently very much in demand in the search for new therapeutic molecules in biology, pharmacy and medicine. It is used to make very fast and inexpensive screens in order to identify a molecule that can have a high activity on a target protein or on the contrary to identify the target of a molecule with important biological activities. On the other hand, when the three-dimensional structure of a target protein is defined, docking makes it possible to highlight the regions and residues to be explored in order to optimize the affinity of a ligand with this target. Thus, it is also used to optimize the selectivity of a molecule between two or more proteins[84].

The general principle is, considering a target of given geometry, to try to place the ligand in interaction with the protein in the most favorable way possible. There are many docking softwares, which can be divided into two categories according to the method used:

- the simulation of trajectory.
- the adjustment of fragments (*matching*).

From a random initial position, the trajectory simulation consists in exploring several positions (with different conformations of the ligand) by evaluating the ligand/protein interaction energy. The best position is the one with the lowest energy, so that, it is the most stable ligand/protein complex. This method is used by AUTODOCK software.

The second approach, the adjustment of fragments, is based on the complementarity of forms. In this method, the ligand is broken down into several fragments and then reconstructed in the active site, trying to match the geometries and chemical functions. This method is much faster than trajectory simulation and generally allows the rapid screening of large libraries of molecules. It is used by softwares such as DOCK, SURFLEX and GOLD. The trajectory simulation method is slower but takes better account of the flexibility of the ligand and allows the exploration of larger regions[85].

I.3.2.2. Types of Molecular Docking

Docking algorithms can be separated into two major classes:

- **Rigid Docking**

In this docking procedures, the protein is considered as a rigid body, the flexibility of the ligand alone is taken into account for obtaining complexes. Depending on the method used to generate the ligand conformers and place them in the catalytic cavity of the receptor, the algorithms can be subdivided into molecular mechanics simulation algorithms and molecular

dynamics form, systematic and stochastic. In most cases, the use of algorithms considering the protein as a rigid body leads to good results, mainly when the protein has a limited flexibility[86, 87].

- **Flexible Docking**

Some proteins naturally have regions of great flexibility, undergoing considerable rearrangements in the presence of a ligand. In this case, neglecting the flexibility of the protein can influence the accuracy of the docking results, and necessitates the use of approaches that take into account the flexibility of the entire system[88]. Indirect or direct methods, where the flexibility of the protein is partially or completely taken into account, are described in the literature. Nevertheless, these methods are not often used because the gain in precision compared to the traditional algorithms is generally too small compared to the increase of the simulation time[86].

I.3.2.3. Molecular Docking protocol

Two steps are mainly used[75]:

- The first step is to download the chemical structures of the targets to be treated, for this it is necessary to go directly to the Protein Data Bank PDB (http://www.pdb.org or http://www.rcsb.org) and determine where are deposited the structures of these targets. The PDB contains several thousand protein structures obtained either by crystallography (X-rays) or by NMR. If the target is not yet deposited at the level of the Bank, and the latter contains a protein with similar sequences, homology modeling intervenes to build the 3D structure of the desired target. After downloading the target (PDB), we use a visualization software to see with which ligands the protein is co-crystallized (water, ligands, ion, ...).
- The second step concerns the structures of the ligand (s) used during molecular docking. There are two large databases of ligand chemical structures. The first represents these structures by molecular modeling computer programs, where different structures are generated by geometry optimization. These structures are governed by the laws of quantum chemistry. In the second case, they are obtained from databases like PubChem Project or other databases of structures.

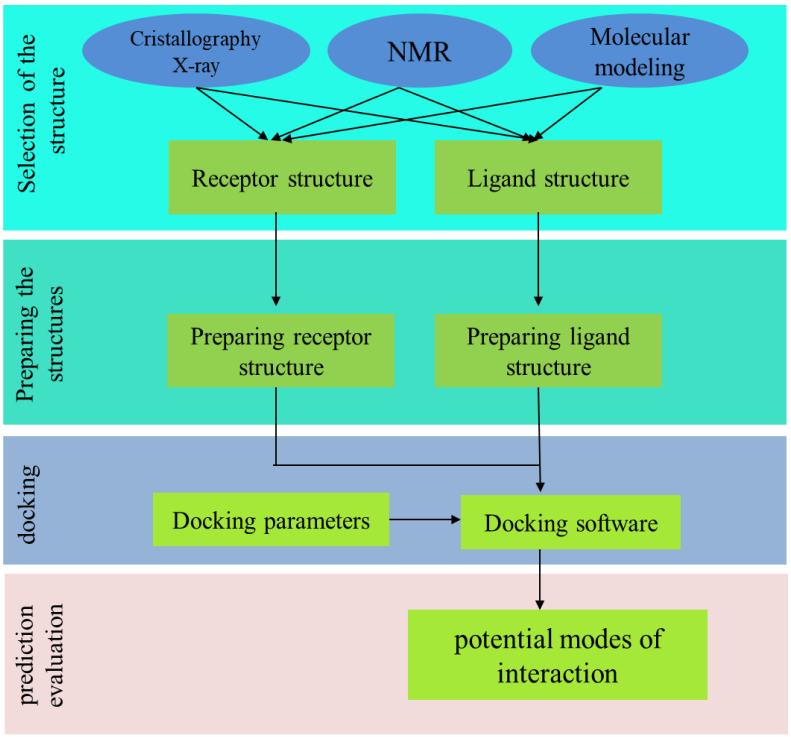


Figure I. 12. Typical steps of a molecular docking.

I.3.3. Quantitative structure-activity relationship QSAR

I.3.3.1. Generality

The Quantitative Structure-Activity Relationship (QSAR) model is a correlation study based on a statistical calculation of physico-chemical parameters with the biological activity [89]. The comparison of the biological activity of certain molecules and their structure has made it possible to establish in many cases correlations between the structural parameters and the properties of a molecule. The association of variation of the activity with the parameters makes it possible to obtain a system of equations which gives, for a given chemical series and for a defined activity, a correlation equation.

The essential interest of this equation is that it should make it possible to determine the value of the parameters corresponding to maximum activity and thus to predict the activity of molecules that have not yet been synthesized[90].

The validity of a QSAR model will therefore depend on the choices made on the parameters. It is therefore advisable to identify and quantitatively evaluate the most relevant parameters to be used according to the activity and molecule selected.

I.3.3.2. Molecular Descriptors

A molecular descriptor is a parameter (a numerical value) specific to a given chemical structure. These values can be obtained experimentally or calculated from the structure of the molecule. Calculated descriptors allow predictions to be made without having to synthesize the molecules, which is one of the objectives of QSAR.

Molecular descriptors play a fundamental role in the studies of the QSAR. They are used as independent variables to predict a dependent variable (activity or property). There is a very large number of molecular descriptors of different complexities and designs, and no strict rules have been established for the selection of suitable descriptors from the large number of available descriptors. This choice has often been based on the chemical intuition of researchers, or by bending to tradition. Many different kinds of descriptors are classified based on the effect (steric, hydrophobic, electronic, etc.) or based on its dimension that called Descriptors Block, whereas we find four types of molecular descriptors 0D, 1D, 2D and 3D. In what follows, we will present only the most commonly used molecular descriptors and those that have been used in all of our work [91, 92].

- **Constitutional Descriptors**

Constitutional descriptors can simply obtain the information from the compounds' chemical composition. Typical constitutional descriptors include molecular weight, total numbers of atoms, bonds, and rings in the molecule. These are one dimensional (1D) descriptors which can easily be obtained with knowledge of the molecular formula[56].

- **Topological Descriptors**

Topological descriptors are two-dimensional (2D) descriptors obtained from molecular topology information; they express the molecular atomic connectivity and can be determined from 2D graph representation of molecules [93].

- **Geometrical Descriptors**

They are based on the spatial arrangement of the atoms making up the molecule and are defined by the coordinates of the atomic nuclei and the size of the molecule represented. These descriptors include information on the molecular surface obtained by the Van Der Waals areas and their superposition [94]. Molecular volumes can be obtained by Van Der Waals volumes [95].

- **Physico-chemical descriptors**

The physicochemical descriptors, some of them reflect the molecular composition of the compound (number and type of atoms and bonds present in the molecule, number of rings, donor/acceptor properties of H-bond, cation, anion, etc.)[96]. Others represent the hydrophilic or lipophilic character of the molecule generally evaluated from the Octanol/water partition coefficient represented by log P [97].

❖ **Molecular Descriptors**

Geometry, topology and atom-weights assembly (GETAWAY), encode the geometrical information obtained from the molecular matrix along with the topological information obtained from the molecular graph, as well as atomic weights information[98].

❖ **Electronic Descriptors**

Electronic descriptors reflect the molecule's electronic structure, based on 3D structure and the charge distribution in the molecule[56].

Table I. 1. List of descriptors used in our work.

Descriptors	Abbreviations	Type
Molecular volume	MV	Geometrical
Molecular weight	MW	
Topological polar surface area	TPSA	Topological
Surface area grid	SAG	Physico-chemical
Polarizability	Pol	
Water solubility	LogS	
Partition coefficient octanol/water	Log P	
Molar refractivity	MR	
H donor	HDo	
H acceptor	HAc	
Total energy	E	Electronic
Atomic charges	q	
Energetic gap	GapE	
Energy of frontier orbital's HOMO,	HOMO	
Energy of frontier orbital's LUMO	LUMO	
Dipole moment	DM	
Heat of formation	HF	Thermodynamic

I.3.3.3 Biological Data

Biological parameters are measurable and quantifiable (specific enzyme or hormone concentration, phenotype distribution of particular genes in a population, etc.) acting as indices for health-related physiological assessments, disease incidence, psychological conditions, environmental exposure and its consequences, metabolic processes, drug addiction, etc.

Biological data are typically expressed on a logarithmic scale. Moreover, inverse logarithm of the activity (log 1/C) are also used to obtain mathematical values higher when structures are biologically very efficient.

The table below lists several examples of biochemical and biological data used in the QSAR analysis [99]:

Table I. 2. biochemical and biological data used in the QSAR analysis

Source of Activity	Biological Parameters
Isolated receptors	
Rate constants	Log k
Michaelis-Menten Constants	Log 1/Km
Inhibition constants	Log 1/Ki
Cellular systems	
Inhibition constants	Log 1/IC50
Cross resistance	Log CR
In vitro biological data	Log 1/C
Mutagenicity states	Log TA98
In vivo systems	
Biocencentration factor	Log BCF
In vivo reaction rates	Log I (Induction)
Pharmacodynamic rates	Log T ((total clearance)

References:

1. Kealy, T. and P. Pauson, A new type of organo-iron compound. Nature, 1951. **168**(4285): p. 1039-1040.
2. Wilkinson, G., et al., The structure of iron bis-cyclopentadienyl. Journal of the American Chemical Society, 1952. **74**(8): p. 2125-2126.
3. Fischer, E. and W. Pfab, Cyclopentadiene-metallic complex, a new type of organo-metallic compound. Z. Naturforsch. B, 1952. **7**: p. 377-379.
4. Dunitz, J. and L. Orgel, Bis-cyclopentadienyl iron: a molecular sandwich. Nature, 1953. **171**(4342): p. 121-122.
5. Woodward, R., M. Rosenblum, and M. Whiting, A new aromatic system. Journal of the American Chemical Society, 1952. **74**(13): p. 3458-3459.
6. Belgacem, T., synthèse, cyclisation, étude electrochimique et structurale de quelques N-acyl-n'alkylférrocenyl méthyle hydrazide. 2007, university Elhadj Lakhder of batna.
7. Meriem, H., Synthèse, caractérisation et évaluation in vitro de l'activité biologique et du pouvoir mutagène de quelques N-(ferrocenylmethyl) aminobenzonitrile, in chemistry departement 2017, university of Eloued.
8. Sørensen, T.J. and M.F. Nielsen, Synthesis, UV/vis spectra and electrochemical characterisation of arylthio and styryl substituted ferrocenes. Central European Journal of Chemistry, 2011. **9**(4): p. 610-618.
9. Khelef, A., Synthèse et étude du comportement anodique de quelques N-ferrocenyl-N-phenylalkanamides et N'-ferrocenyl-N'-phenylalkanehydrazides et étude structurale de leurs phases cristallines. 2014, Université Mohamed Khider Biskra.
10. Broadhead, G., J. Osgerby, and P. Pauson, 127. Ferrocene derivatives. Part V. Ferrocenealdehyde. Journal of the Chemical Society (Resumed), 1958: p. 650-656.
11. Benkeser, R.A., D. Goggin, and G. Schroll, A Route to Monosubstituted Ferrocene Compounds. Journal of the Chemical Society (Resumed), 1954. **76 (4025)**.
12. LINDSAY, J.K. and C.R. HAUSER, Aminomethylation of Ferrocene to Form N, N-Dimethylaminomethylferrocene and Its Conversion to the Corresponding Alcohol and Aldehyde1. The Journal of Organic Chemistry, 1957. **22**(4): p. 355-358.
13. Lednicer, D. and C.R. Hauser, N, N-Dimethylaminomethylferrocene Methiodide. Organic syntheses, 1960. **40**: p. 31-31.
14. Nesmeyanov, A., et al., synthesis of ferrocene derivatives by means of the iodomethylate of n, n-dimethylaminomethylferrocene. Doklady akademii nauk sssr, 1959. **124**(2): p. 331-334.

15. Top, S., et al., Ferrocenyl hydroxytamoxifen: a prototype for a new range of oestradiol receptor site-directed cytotoxics. Chemical Communications, 1996(8): p. 955-956.
16. Top, S., et al., Synthesis, biochemical properties and molecular modelling studies of organometallic specific estrogen receptor modulators (SERMs), the ferrocifens and hydroxyferrocifens: evidence for an antiproliferative effect of hydroxyferrocifens on both hormone-dependent and hormone-independent breast cancer cell lines. Chemistry–A European Journal, 2003. **9**(21): p. 5223-5236.
17. Khelef, A. and T. Lanez, In vitro assays of the antioxidant activities of ferrocene derivatives bearing amine, amide or hydrazine groups. Der Pharma Chemica, 2015. **7**(6): p. 318-323.
18. Edwards, E., R. Epton, and G. Marr, 1, 1′-Ferrocenyldiacetic Acid Anhydride and its Use in the preparation of heteroannularly substituted ferrocenyl-penicillins and-cephalosporins. Journal of Organometallic Chemistry, 1976. **122**(3): p. C49-C53.
19. Tabbì, G., et al., Water stability and cytotoxic activity relationship of a series of ferrocenium derivatives. ESR insights on the radical production during the degradation process. Journal of medicinal chemistry, 2002. **45**(26): p. 5786-5796.
20. Chavain, N., Conception, synthèse, analyse structurale et activité antipaludique de dérivés de la ferroquine: investigation des mécanismes d'action. 2008, Lille 1.
21. Nesmeyanov, A., et al., Ferrocerone. US Patent, 1971. **119356**.
22. Top, S., et al., Facile route to ferrocifen, 1-[4-(2-dimethylaminoethoxy)]-1-(phenyl-2-ferrocenyl-but-1-ene), first organometallic analogue of tamoxifen, by the McMurry reaction. Journal of Organometallic Chemistry, 1997. **541**(1): p. 355-361.
23. Patra, M. and G. Gasser, The medicinal chemistry of ferrocene and its derivatives. Nature Reviews Chemistry, 2017. **1**(9): p. 0066.
24. GÖRMEN, M., Synthesis of ferrocenophane derivatives and evaluation of their antiproliferative activity against breast and prostate cancer cells. 2010, university of Pierre et Marie Curie (Paris VI).
25. Jaouen, G., et al., Organometallics targeted to specific biological sites: The development of new therapies, in Bioorganometallics: Biomolecules, Labeling, Medicine. 2006, Wiley Online Library.
26. Cordina-Duverger, E., Facteurs de risques hormonaux et anthropométriques dans le cancer du sein de la femme: étude CECILE. 2015, UNIVERSITY OF PARIS-SUD.
27. Winslow, T. 2011; Available from: https://www.teresewinslow.com/#/breast/.
28. Organization, W.H. Cancer fact sheets. 2018; Available from: http://gco.iarc.fr/today/fact-sheets-cancers.

29. Jéhannin-Ligier, K., et al., Projection de l'incidence et de la mortalité par cancer en France métropolitaine en 2017. Rapport technique. Saint-Maurice: Santé publique France, 2017: p. 1-80.
30. Platet, N., et al., Estrogens and their receptors in breast cancer progression: a dual role in cancer proliferation and invasion. Critical reviews in oncology/hematology, 2004. **51**(1): p. 55-67.
31. Sun, Y.-S., et al., Risk factors and preventions of breast cancer. International journal of biological sciences, 2017. **13**(11): p. 1387.
32. society, a.c. breast cancer risk and prevention. 2019; Available from: https://www.cancer.org/.
33. Brinton, L.A., et al., Menstrual factors and risk of breast cancer. Cancer investigation, 1988. **6**(3): p. 245-254.
34. Hsieh, C.C., et al., Age at menarche, age at menopause, height and obesity as risk factors for breast cancer: associations and interactions in an international case-control study. International journal of cancer, 1990. **46**(5): p. 796-800.
35. Kvåle, G. and I. Heuch, Menstrual factors and breast cancer risk. Cancer, 1988. **62**(8): p. 1625-1631.
36. MacMahon, B., et al., Age at first birth and breast cancer risk. Bulletin of the World Health Organization, 1970. **43**(2): p. 209.
37. Rosero-Bixby, L., M.W. Oberle, and N.C. Lee, Reproductive history and breast cancer in a population of high fertility, Costa Rica, 1984–85. International journal of cancer, 1987. **40**(6): p. 747-754.
38. Cancer, C.G.o.H.F.i.B., Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50 302 women with breast cancer and 96 973 women without the disease. The lancet, 2002. **360**(9328): p. 187-195.
39. society, a.c., treating breast cancer. 2017.
40. Harari, D. and Y. Yarden, Molecular mechanisms underlying ErbB2/HER2 action in breast cancer. Oncogene, 2000. **19**(53): p. 6102.
41. Classe, J.-M. and P. Descamps, CANCER DU SEIN. 2007: Elsevier.
42. Baselga, J., et al., Lapatinib with trastuzumab for HER2-positive early breast cancer (NeoALTTO): a randomised, open-label, multicentre, phase 3 trial. The Lancet, 2012. **379**(9816): p. 633-640.

43. Okines, A.F. and D. Cunningham, Trastuzumab: a novel standard option for patients with HER-2-positive advanced gastric or gastro-oesophageal junction cancer. Therapeutic advances in gastroenterology, 2012. **5**(5): p. 301-318.
44. Jahanmohan, J.P., Les cancers du sein agressifs : conséquences de la ménopause chimio-induite chez les femmes jeunes atteintes d'un cancer du sein non métastatique et facteurs pronostiques de la rechute du cancer du sein triple négatif. 2019, university of clermont auvergne.
45. Goss, P.E., et al., Late extended adjuvant treatment with letrozole improves outcome in women with early-stage breast cancer who complete 5 years of tamoxifen. Journal of clinical oncology, 2008. **26**(12): p. 1948-1955.
46. Cuzick, J., et al., Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 10-year analysis of the ATAC trial. The lancet oncology, 2010. **11**(12): p. 1135-1141.
47. Regan, M.M., et al., Assessment of letrozole and tamoxifen alone and in sequence for postmenopausal women with steroid hormone receptor-positive breast cancer: the BIG 1-98 randomised clinical trial at 8· 1 years median follow-up. The lancet oncology, 2011. **12**(12): p. 1101-1108.
48. Group, E.B.C.T.C., Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. The Lancet, 2005. **365**(9472): p. 1687-1717.
49. Group, I.B.C.S., Adjuvant chemotherapy followed by goserelin versus either modality alone for premenopausal lymph node–negative breast cancer: a randomized trial. Journal of the National Cancer Institute, 2003. **95**(24): p. 1833-1846.
50. Zhang, P., et al., Effects of ovarian ablation or suppression in premenopausal breast cancer: a meta-analysis of randomized controlled trials. European Journal of Surgical Oncology (EJSO), 2017. **43**(7): p. 1161-1172.
51. Leach, A.R., Molecular modelling: principles and applications-Prentice. second edition ed. 2001, New Jersey: Prentice Hall.
52. Dorsett, H. and A. White, Overview of molecular modelling and ab initio molecular orbital methods suitable for use with energetic materials. 2000, DEFENCE SCIENCE AND TECHNOLOGY ORGANIZATION SALISBURY (AUSTRALIA).
53. BELAIDI, H., Étude par la chimie informatique, des corrélations structures-activités/propriétés dans des hétérocycles à intérêt thérapeutique. 2014, University Mohamed Khider of Biskra.

54. Griffths, D.J., Introduction to quantum mechanics. Vol. 193. 1995, USA: New Jersey: PrenticeHall, Inc 224-237.
55. Kutudila, P., Structure et réactivité des triarylbismuths : approche théorique et expérimentale. 2017, University Paris-Est.
56. Aicha, K., Etude par la modélisation moléculaire des relations structures-propriétés de quelques séries hétérocycliques bioactives. 2015, university of biskra.
57. Radhia, M., Nouvelle approche des relations structures- activités dans des molécules antibiotiques. 2015, University of Mohamed Khider Biskra.
58. Hartree, D.R. The wave mechanics of an atom with a non-Coulomb central field. Part I. Theory and methods. in Mathematical Proceedings of the Cambridge Philosophical Society. 1928. Cambridge University Press.
59. Diharce, J., Étude par modélisation moléculaire de systèmes multienzymatiques impliqués dans la biosynthèse des flavonoïdes. 2014, Universite De Nice Sophia Antipolis.
60. Hohenberg, P. and W. Kohn, Inhomogeneous electron gas. Physical review, 1964. **136**(3B): p. B864.
61. Khalil, S., Contribution à l'étude des complexes moléculaires du diclofénac : Etudes expérimentales et par modélisation moléculaire, in Département des Sciences de la Matière. 2015, University of Guelma.
62. Becke, A.D., Density-functional exchange-energy approximation with correct asymptotic behavior. Physical review A, 1988. **38**(6): p. 3098.
63. Perdew, J.P. and Y. Wang, Accurate and simple analytic representation of the electron-gas correlation energy. Physical Review B, 1992. **45**(23): p. 13244.
64. Perdew, J.P. and W. Yue, Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation. Physical review B, 1986. **33**(12): p. 8800.
65. Lee, C., W. Yang, and R.G. Parr, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Physical review B, 1988. **37**(2): p. 785.
66. Becke, A.D., Density-functional thermochemistry. III. The role of exact exchange. The Journal of Chemical Physics, 1993. **98**(7): p. 5648-5652.
67. Perdew, J.P., K. Burke, and Y. Wang, Generalized gradient approximation for the exchange-correlation hole of a many-electron system. Physical Review B, 1996. **54**(23): p. 16533.

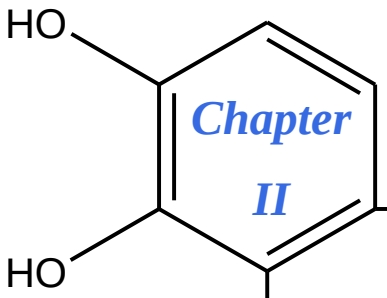
68. NABIL, B., Etude théorique et comparative de différentes familles de composés chimiques : Clusters de Bor et Carboranes apparentées, complexes organométalliques et composés organiques biologiquement actifs. 2019, University of El-Oued.
69. Fatima, B.n.L., Contribution à l'etude de la stabilité des complexes de metaux pyrazoles par modelisation moleculaire, in Department Of Chemistry. 2011, Universite Abou-Bakr Belkaid De Tlemcen.
70. Dewar, M.J. and W. Thiel, Ground states of molecules. 38. The MNDO method. Approximations and parameters. Journal of the American Chemical Society, 1977. **99**(15): p. 4899-4907.
71. Pople, J.A. and G.A. Segal, Approximate self-consistent molecular orbital theory. III. CNDO results for AB2 and AB3 systems. The Journal of Chemical Physics, 1966. **44**(9): p. 3289-3296.
72. Dewar, M.J., et al., Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. Journal of the American Chemical Society, 1985. **107**(13): p. 3902-3909.
73. Stewart, J.J., Optimization of parameters for semiempirical methods II. Applications. Journal of computational chemistry, 1989. **10**(2): p. 221-264.
74. Stewart, J.J., Optimization of parameters for semiempirical methods V: modification of NDDO approximations and application to 70 elements. Journal of Molecular modeling, 2007. **13**(12): p. 1173-1213.
75. Soufi, L.W.n., Contribution a la modélisation des interactions dans les biomolécules: Cas de la Maladie de Parkinson. 2016, university abou bekr belkaid of Tlemcen.
76. Fatima, B.n.L., Etude par modélisation moléculaire des mécanismes de complexation. 2016, University abou bekr belkaid of Tlemcen.
77. Slater, J.C., Atomic shielding constants. Physical Review, 1930. **36**(1): p. 57.
78. Slater, J., Analytic atomic wave functions. Physical Review, 1932. **42**(1): p. 33.
79. Boys, S.F., Electronic wave functions-I. A general method of calculation for the stationary states of any molecular system. Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 1950. **200**(1063): p. 542-554.
80. Hay, P.J. and W.R. Wadt, Ab initio effective core potentials for molecular calculations. Potentials for K to Au including the outermost core orbitals. The Journal of Chemical Physics, 1985. **82**(1): p. 299-310.
81. Hay, P.J. and W.R. Wadt, Ab initio effective core potentials for molecular calculations. Potentials for the transition metal atoms Sc to Hg. The Journal of Chemical Physics, 1985. **82**(1): p. 270-283.

82. Wadt, W.R. and P.J. Hay, Ab initio effective core potentials for molecular calculations. Potentials for main group elements Na to Bi. The Journal of Chemical Physics, 1985. **82**(1): p. 284-298.
83. Chiodo, S., N. Russo, and E. Sicilia, LANL2DZ basis sets recontracted in the framework of density functional theory. J Chem Phys, 2006. **125**(10): p. 104107.
84. Traore, M.D.M., Synthèse et études de modélisation moléculaire dans l'optimisation de la sélectivité de nouveaux agents antiparasitaires inspirés de produits naturels. 2016, University Grenoble Alpes.
85. Fortuné, A., Techniques de Modélisation Moléculaire appliquées à l'Etude et à l'Optimisation de Molécules Immunogènes et de Modulateurs de la Chimiorésistance. 2006, University of Joseph Fourier- Grenouble1.
86. Feher, M., Consensus scoring for protein–ligand interactions. Drug discovery today, 2006. **11**(9-10): p. 421-428.
87. Férey, N., et al., Docking de protéines en réalité virtuelle. Une approche hybride et multimodale. 2009.
88. Clark, R.D., et al., Consensus scoring for ligand/protein interactions. Journal of Molecular Graphics and Modelling, 2002. **20**(4): p. 281-295.
89. Allinger, N.L., K. Chen, and J.H. Lii, An improved force field (MM4) for saturated hydrocarbons. Journal of Computational Chemistry, 1996. **17**(5-6): p. 642-668.
90. Jones, G.B. and B.J. Chapman, π stacking effects in asymmetric synthesis. Synthesis, 1995. **1995**(05): p. 475-497.
91. Karelson, M., Molecular descriptors in QSAR/QSPR. Vol. 230. 2000: Wiley-Interscience New York.
92. Puzyn, T., J. Leszczynski, and M.T. Cronin, Recent advances in QSAR studies: methods and applications. Vol. 8. 2010: Springer Science & Business Media.
93. Randić, M., The connectivity index 25 years after. Journal of Molecular Graphics and Modelling, 2001. **20**(1): p. 19-35.
94. Labute, P., A widely applicable set of descriptors. Journal of Molecular Graphics and Modelling, 2000. **18**(4-5): p. 464-477.
95. Higo, J.I. and N. Gō, Algorithm for rapid calculation of excluded volume of large molecules. Journal of computational chemistry, 1989. **10**(3): p. 376-379.
96. Bosque, R., et al., A QSPR study of the p solute polarity parameter to estimate retention in HPLC. Journal of chemical information and computer sciences, 2003. **43**(4): p. 1240-1247.

97. Viswanadhan, V.N., et al., Atomic physicochemical parameters for three dimensional structure directed quantitative structure-activity relationships. 4. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition of certain naturally occurring nucleoside antibiotics. Journal of Chemical Information and Computer Sciences, 1989. **29**(3): p. 163-172.

98. Consonni, V., R. Todeschini, and M. Pavan, Structure/response correlations and similarity/diversity analysis by GETAWAY descriptors. 1. Theory of the novel 3D molecular descriptors. Journal of chemical information and computer sciences, 2002. **42**(3): p. 682-692.

99. Selassie, C. and R.P. Verma, History of quantitative structure–activity relationships. Burger's Medicinal Chemistry and Drug Discovery, 2003: p. 1-96.



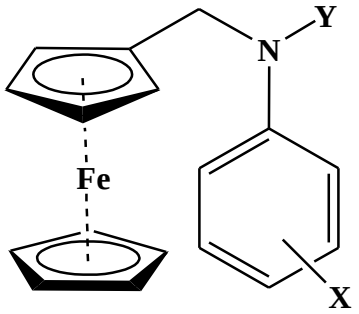
Chapter II

Molecular Geometry, Structural, Electronic properties and ADME studies

II. 1. Introduction

Ferrocene is the most well-known organometallic. First prepared in 1951, and structurally characterized in 1952, this 'sandwich' complex has an interesting history [1, 2]. Its derivatives have attracted significant interest in medicinal chemistry. The salt of ferrocenylmethyl trimethylammonium iodide plays a very important role in the preparation of a very large number of substituted N-ferrocenylmethylaniline, which are likely to exihib antimutagenic activity. In addition the amine function in this derivatives can be acylated by different acylation reagents producing a series ferrocerene amides of the N-ferrocenylmethyl-N-phenylacylamide type which are likely to possess antimutagenic activity. In this chapter, due to the lack of adequate information about the chemical properties and structural characteristics, we carry out a study of the molecular structure of 29 N-ferrocenylmethylaniline derivatives (table II.1) using Density Functional Theory (DFT) method concerning the numerous applications of new methods in quantum mechanics to study structural properties and molecules spectroscopy. DFT with different basis sets is one of the efficient and effective methods to determine structural and crystal properties of the molecules.

Table II. 1. General formula of N-ferrocenylmethylaniline derivatives.



N°:	X	Y	N°:	X	Y	N°:	X	Y
A0	H	H	**A10**	4-Cl	H	**A20**	4-COMe	H
A1	H	COMe	**A11**	3-Br	H	**A21**	H	NHCOMe
A2	H	COEt	**A12**	4-Br	H	**A22**	H	NHCOPh
A3	H	COPh	**A13**	2-CN	H	**A23**	2-NO2	COMe
A4	H	CO-Ph-2-NO2	**A14**	3-CN	H	**A24**	3-NO2	COMe
A5	2-Me	H	**A15**	4-CN	H	**A25**	4-NO2	COMe
A6	3-Me	H	**A16**	2-NO2	H	**A26**	2-CN	COMe
A7	2-Et	H	**A17**	3-NO2	H	**A27**	3-CN	COMe
A8	2-Cl	H	**A18**	4-NO2	H	**A28**	4-CN	COMe
A9	3-Cl	H	**A19**	2-COMe	H			

II.2. Computational details

Molecular modelling is a set of techniques for investigating chemical problems on a computer. The types of calculations used in this part are:

II.2.1. Molecular Geometry

It's the shape of the molecules, bond lengths, angles and dihedrals bond, which can be interpreted by the results of X-ray diffraction studies and electron diffraction or refine them.

II.2.2. Geometry Optimization

The geometry optimization process starts with a set of coordinates that describes the starting geometry, and then the computation looks for a new set of coordinates that has the geometry of the minimum energy.

II.2.3. Mulliken population analysis (MPA)

This population analysis is the oldest and one of the most commonly used. MPA splits bond orbital populations equally between the two bond atoms. This approach is quite simplified and does not take into consideration that one of bonded atoms can attract electrons markedly more than the second one. At the other hand MPA 's simplicity is also an advantage, since the approach can be easily used [3].

II.2.4. Single Point Calculations

This method calculates the energy, the wave function and other properties to a fixed simple geometry, it is usually done at the beginning of the study of a molecule to identify the geometry and symmetry of the system. It is also frequently used after an optimization of the geometry[4].

II.2.5. The molecular electrostatic potential (MESP)

MESP surface which is a plot of electrostatic potential mapped onto the iso-electron density surface [5], its importance lies in the fact that it simultaneously displays the molecular size and shape as well as positive, negative and neutral electrostatic potential regions in terms of the electrostatic surface, which explain the investigation of the molecular structure with its physiochemical property relationship [6].

II.2.6. Molecular Orbitals (HOMO-1, HOMO, LUMO, LUMO+1, ΔE)

The most important orbitals in a molecule are the frontier molecular orbitals, called highest occupied molecular orbital (HOMO) which describes the ionization energy and lowest unoccupied molecular orbital (LUMO) which describes the electron affinity in a molecule. These orbitals determine the way the molecule interacts with other species including the interactions between a ligand and its receptor [7]. While, electrophilic and nucleophilic attack will most likely occur at atoms where the coefficients of the corresponding atomic orbitals in HOMO and LUMO, respectively, are large [8].

Sometimes, even, the first excited state is degenerate the terms HOMO and LUMO cannot give clear distinctions. Thus, we can also use the second highest occupied and second lowest

unoccupied molecular orbitals (HOMO -1 and LUMO +1, respectively), whereas, these orbitals are lower in energy than HOMO and LUMO.

II.2.7. Chemical Reactivity:

For instance, knowing where the electrons are concentrated (nucleophilic sites) and where they want to go (electrophilic sites) allows us to anticipate where various kinds of reagents would attack a molecule [9].

II.2.8. IR and NMR spectra:

These can be computed, and if the molecule is unknown, someone trying to make it knows what to look for. Enable the calculation of infrared stretching and bending absorption frequencies and it is a lot of fun to view animations of these types of motions in molecules. The vibrational frequency of a two-atom system is proportional to the square root of the force constant (the second derivative of the energy with respect to the interatomic distance) divided by the reduced mass of the system (which depends on the masses of the two atoms) [9].

- **Gaussian calculations**

First, all structures were optimized by Gaussian 09 and viewed by GaussView 5.0 using DFT/B3LYP method and the two basis sets 6-311G ++ (d, p) and LanL2DZ. Then the geometric parameters like bond lengths and bond angles between its constituent atoms, were compared with the experimental values obtained from X-ray mesurments (obtained from scientific literature) [10-14]. The harmonies between calculated and experimental structural parameters were studied and the regression coefficients (R2) were calculated using Excel from Microsoft office tools. Moreover, Mulliken atomic charges of constituent atoms and MESP of the studied compounds were calculated using the same computational methods with the same bases set and were compared to each other. Also, the optimized structures of the studied compounds were used to calculate the ^{1}H and ^{13}C NMR spectra using DFT/B3LYP and the gauge independent atomic orbital (GIAO) methods with 6-311G++(d,p) and LanL2DZ basis set and chloroform as solvent. The chemical shifts of the compounds were computed with reference to TMS B3LYP / 6-311G++ (d, p) GIAO. The obtained values were compared to each other and then to the experimental chemical shifts. The vibrational frequencies were obtained as well from the optimized structures, and compared with the experimental frequencies, the vibrational assignments were observed and explained as symmetric, asymmetric, stretching, in plane bending and/or out of plane bending. Finally, HOMO and LUMO energies, as well as the energetic gaps (HOMO-LUMO) were calculated using single point energy.

II.2.9. ADME and Pharmacokinetics properties

In order to develop drugs with desired properties and optimal dosing regimens, it is very essential to determine the pharmacokinetic properties of these drugs, including their ADME. Pharmacokinetics is the field that studies the four processes of ADME. It answers whether a drug is able to get to the site of action. ADME refers to **Absorption**, **Distribution**, **Metabolism** (biotransformation), and **Elimination,** once a drug enters the body, the body starts to process these four aspects of a drug [15].

Nowadays, drug developers are facing significant difficulties to develop more efficient as well as cost-effective drugs as compared to current treatments, therefore, a need to optimize physicochemical properties including their bioactivities and the ADME properties as well have arisen in order to develop drugs with the least adverse effects. Other properties that are taken into account in the prediction of the behaviors of newly-developed drugs are related to the size of doses and their frequencies as well. These properties include drugs' bioavailability, oral absorption, clearance, volume of distribution, as well as the penetration through the blood-brain barrier (BBB). The metabolic stability of drugs is also considered as a crucial issue that drug developers need to pay particular attention to, since the success or failure of the development process depends greatly on the metabolism profile of these drugs [16].

The ADME and Pharmacokinetics properties of FMA derivatives were predicted using SwissADME web server. The drug-like nature including the Lipinski rule of five, Ghose, Veber, Egan and Muegge rules, were predicted for all compounds by entering the SMILES of the compounds as well as the bioavailability score and the pharmacokinetics properties.

II.3. Results and discussion

II.3.1. Geometric parameters

The parameters for optimizing the molecular geometry (bond distances and angles) of A0 are tabulated in table II.2 and the optimized structure with the numbering arrangement are shown in figure II.1. Since the crystal structure of A0 is not available, the optimized structure can only be compared with other similar systems A1 for which the crystal structures have been solved.

The comparision of the calculated geometric parameters with the X-ray data shows a good agreement, with a coefficient regression R^2 equal to 0.980 for interatomic distances with both basis sets and 0.930 and 0.940 for angles with 6-311G++ (d,p) and LanL2DZ respectively (find in appendix n°1 the distribution graphs for the geometric parameters).

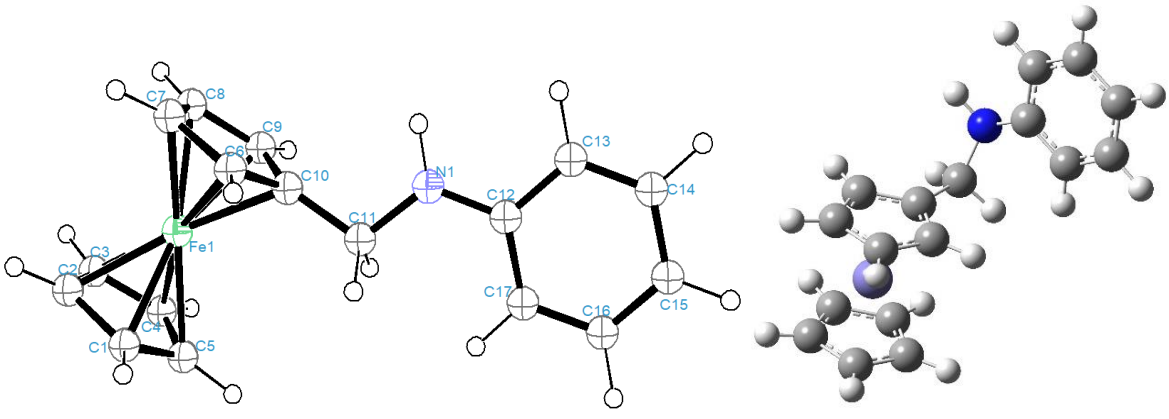
The Fe-C bond lengths were equal to 2.08 for ferrocene [17], also Mauricio. Y et al have found that this bond is equal to 2.08 for 2-ferrocenyl-1,8-naphthyridine complex [18]. In this

work, we calculated the Fe-C bond in around 2.07Å (2.11Å) using 6-311G++ (d,p) (LanL2DZ), while the experimental vary from 1.99 to 2.08 Å with max difference equal to 0.08 Å (0.13 Å) for 6-311G++ (d,p) (LanL2DZ). The C-C bond lengths were calculated in the range of 1.388 – 1.508 Å (1.401-1.505 Å) using 6-311G++ (d,p) (LanL2DZ), which fit with the experimental values between 1.52 and 1.31 Å with max differences equal to 0.11 and 0.13 for 6-311G++ (d,p) and LanL2DZ, respectively. Similar to this results, Mauricio. Y et al [18] have calculated this bond at 1.416 Å for a ferrocenic complex, and another study has shown that the C-C bond of ferrocene is equal to 1.39 Å [17]. The interatomic distances shows better agreement with other compounds (appendix n° 1) where max differences around 0.06 for the C-C bonds lengths and coefficient regressions R^2 from 0.97 to 0.99 with both bases sets.

Moreover, the C-N bonds were in the range of 1.39 -1.46 Å for both basis set, which is in good agreement with the X-ray values with max difference equal to 0.04.

The expected value for bond angle in benzene ring is 120° [19, 20]. Likewise, in this work the C-C-C angles for benzene ring are from 117° to 121°. Moreover, the bond angles C6-C10-C11of the studied compounds were calculated 127°(126°) using 6-311G++ (d,p) (LanL2DZ) and in experimental result 126°, whereas, the same angle were equal to 125.8° for 2-ferrocenyl-1,8-naphthyridine [18].

As seen from tables II.2 and the other tables of geometric parameters in appendix n° 1, most of the optimized bond lengths are slightly longer than the experimental values, and the bond angles are slightly different from experimental ones because the molecular states are different during experimental and theoretical processes.



(a) (b)

Figure II. 1 **(a)** The ORTEP view of A0 shows atomic labeling scheme and 40% probability level displacement ellipsoids. **(b)** DFT optimized structure of A0 by GaussView.

Table II. 2. Selected optimized geometry parameters of compound A0 by DFT/B3LYP method.

		X-ray	a	b			X-ray	a	b
Bond distances(Å)									
C1	C2	1,37	1,425	1,443	C8	C7	1,38	1,425	1,444
C1	C5	1,43	1,426	1,443	C8	Fe1	2,03	2,076	2,117
C1	Fe1	2,05	2,078	2,121	C9	Fe1	2,088	2,071	2,115
C2	C3	1,48	1,426	1,443	C10	C9	1,47	1,433	1,446
C2	Fe1	2,056	2,078	2,119	C10	C11	1,52	1,508	1,506
C3	C4	1,44	1,426	1,443	C10	Fe1	2,002	2,075	2,119
C3	Fe1	2,04	2,077	2,119	C13	C12	1,372	1,408	1,424
C4	Fe1	2,03	2,078	2,121	C14	C13	1,376	1,388	1,401
C5	C4	1,31	1,426	1,443	C14	C15	1,366	1,397	1,411
C5	Fe1	1,99	2,078	2,12	C15	C16	1,368	1,392	1,407
C6	C7	1,34	1,425	1,442	C15	C14	1,38	1,397	1,411
C6	C10	1,4	1,431	1,446	C17	C16	1,385	1,394	1,407
C6	Fe1	2,03	2,074	2,114	C17	C12	1,374	1,405	1,421
C7	Fe1	2,03	2,076	2,117	N1	C12	1,435	1,396	1,393
C8	C9	1,4	1,425	1,441	N1	C11	1,479	1,465	1,468
Bond angles (°)									
C9	Fe1	C6	68,28	67,56	66,93	H1	C1	C5	125,98	126,01	126,01
C9	Fe1	C10	41,06	40,43	39,95	H1	C1	Fe1	125,81	124,85	125,28
C9	Fe1	C8	40,21	40,21	39,83	C2	C1	C5	108	108,01	107,97
C9	Fe1	C7	67,68	67,51	66,93	C2	C1	Fe1	70,33	69,94	70,06
C9	Fe1	C3	130,1	123,53	123,92	C5	C1	Fe1	69,41	69,93	70,09
C9	Fe1	C1	128,95	159,57	159,92	H2	C2	C1	126,14	126,02	126,02
C9	Fe1	C4	116,55	108,72	109,39	H2	C2	C3	126,25	125,93	125,99
C9	Fe1	C2	167,78	158,96	158,94	H2	C2	Fe1	126	124,71	125,05
C9	Fe1	C5	107,14	123,97	124,59	C1	C2	C3	107,61	108,04	107,98
C6	Fe1	C10	40,68	40,35	39,95	C1	C2	Fe1	69,36	69,94	70,15
C6	Fe1	C8	68,21	67,55	66,95	C3	C2	Fe1	70,04	69,91	70,1
C6	Fe1	C7	40,85	40,16	39,85	H3	C3	C4	125,6	126	125,96
C6	Fe1	C3	132	158,8	158,95	H3	C3	C2	125,54	126,01	125,97
C6	Fe1	C1	116,25	108,58	109,36	H3	C3	Fe1	126,4	124,8	124,97
C6	Fe1	C4	170,27	159,64	159,89	C4	C3	C2	108,87	107,98	108,07
C6	Fe1	C2	109,34	123,36	123,88	C4	C3	Fe1	69,79	69,94	70,18
C6	Fe1	C5	148,28	123,92	124,57	C2	C3	Fe1	69,83	69,94	70,1
C10	Fe1	C8	68,57	67,93	67,1	H4	C4	C3	126,05	126,06	125,98
C10	Fe1	C7	68,53	67,83	67,1	H4	C4	C5	126,14	125,94	126,1
C10	Fe1	C3	168,82	159,36	159,49	H4	C4	Fe1	126,12	124,79	125,26
C10	Fe1	C1	106,72	123,58	124,43	C3	C4	C5	107,8	108	107,92

C10	Fe1	C4	148,32	123,76	124,43	C3	C4	Fe1	70,21	69,92	70,04
C10	Fe1	C2	129,41	159,08	159,47	C5	C4	Fe1	69,23	69,92	70,07
C10	Fe1	C5	114,86	108,6	109,65	H5	C5	C1	126,09	126,04	125,98
C8	Fe1	C7	40,08	40,16	39,87	H5	C5	C4	126,19	125,98	125,96
C8	Fe1	C3	119,06	108,15	108,69	H5	C5	Fe1	125,44	125,07	125,53
C8	Fe1	C1	167,64	158,85	158,94	C1	C5	C4	107,72	107,97	108,06
C8	Fe1	C4	109,49	123,49	123,91	C1	C5	Fe1	69,73	69,93	70,11
C8	Fe1	C2	151,25	123,18	123,51	C4	C5	Fe1	70,3	69,91	70,15
C8	Fe1	C5	129,53	159,37	159,52	H6	C6	C7	126,06	125,87	126,22
C7	Fe1	C3	111,35	123,16	123,53	H6	C6	C10	125,88	125,72	125,45
C7	Fe1	C1	150,28	123,44	123,89	H6	C6	Fe1	125,94	125,21	125,42
C7	Fe1	C4	131,46	158,89	158,96	C7	C6	C10	108,06	108,4	108,33
C7	Fe1	C2	118,96	108,15	108,65	C7	C6	Fe1	70,04	70	70,2
C7	Fe1	C5	168,53	159,29	159,52	C10	C6	Fe1	69,61	69,86	70,22
C3	Fe1	C1	67,5	67,47	66,81	H7	C7	C6	126,01	125,91	126,02
C3	Fe1	C4	40	40,14	39,79	H7	C7	C8	126,01	125,99	126,05
C3	Fe1	C2	40,12	40,15	39,81	H7	C7	Fe1	126,57	124,78	125,23
C3	Fe1	C5	67,61	67,48	66,82	C6	C7	C8	108,09	108,09	107,93
C1	Fe1	C4	68,1	67,47	66,83	C6	C7	Fe1	69,11	69,85	69,95
C1	Fe1	C2	40,31	40,12	39,79	C8	C7	Fe1	69,97	69,92	70,06
C1	Fe1	C5	40,86	40,14	39,8	H8	C8	C7	125,91	126,08	125,99
C4	Fe1	C2	67,81	67,45	66,84	H8	C8	C9	126,01	126,05	126,07
C4	Fe1	C5	40,46	40,16	39,78	H8	C8	Fe1	126,45	124,69	125,27
C2	Fe1	C5	68,19	67,44	66,79	C7	C8	C9	108,08	107,85	107,94
H1	C1	C2	126,01	125,97	126,02	C7	C8	Fe1	69,95	69,93	70,06
C9	C8	Fe1	69,17	69,7	69,98	H11b	C11	H11a	107,71	106,14	106,75
H9	C9	C8	125,56	125,92	126,23	H11b	C11	N1	108,93	108,73	109,96
H9	C9	C10	125,55	125,59	125,43	H11b	C11	C10	108,99	110,67	109,76
H9	C9	Fe1	125,81	124,48	125,4	H11a	C11	N1	108,97	106,91	109,87
C8	C9	C10	108,88	108,48	108,34	H11a	C11	C10	108,97	109,33	109,77
C8	C9	Fe1	70,62	70,09	70,19	N1	C11	C10	113,13	114,67	110,66
C10	C9	Fe1	69,65	69,94	70,2	N1	C12	C17	120,76	122,87	121,83
C6	C10	C9	106,88	107,19	107,46	N1	C12	C13	119,81	119,1	119,81
C6	C10	C11	126,28	127,33	126,26	C17	C12	C13	119,6	117,95	118,37
C6	C10	Fe1	69,71	69,79	69,83	H13	C13	C14	120,23	119,76	120,06
C9	C10	C11	126,81	125,42	126,28	H13	C13	C12	120,26	119,18	119,34
C9	C10	Fe1	69,28	69,63	69,85	C14	C13	C12	119,51	121,06	120,6
C11	C10	Fe1	124,61	127,89	126,6	H14	C14	C13	119,57	119,27	119,22
C15	C16	C17	120,29	121,17	121,09	H14	C14	C15	119,57	120,03	119,92
H17	C17	C16	119,79	119,04	119,43	C13	C14	C15	120,85	120,71	120,86

H17	C17	C12	119,6	120,46	120,26	H15	C15	C16	120,39	120,72	120,66
C16	C17	C12	120,61	120,5	120,3	H15	C15	C14	120,3	120,66	120,56
H12	N1	C12	-	112,9	118,76	C16	C15	C14	119,31	118,62	118,78
H12	N1	C11	-	112,89	117,36	H16	C16	C15	119,9	119,89	119,89
C12	N1	C11	116,9	123,47	123,88	H16	C16	C17	119,81	118,94	119,02

a : 6-311G++ (d, p) , b: LanL2DZ

II.3.2. Atomic charges

The Mulliken analysis is one of the most widely used population analysis method because it is the oldest and the simplest one. It plays an essential role in determining the atoms' vibrational properties. Indeed, it figures out which atom is electron donor and which one is electron acceptor. So it can determine the dipole moment and polarizability of the molecule [21, 22]. Mulliken atomic charges between constituent atoms of A0 shows that most of carbon atoms have negative charges, while the iron and nitrogen atoms have positive charges using 6-311G++ (d,p), in contrast, the LanL2dz basis set shows that all atoms have negative charges. However, the MESP map below (figure. II. 2) shows coincidence with the 6-311G++ (d,p) basis set. The Mulliken atomic charges of the other compounds are presented in appendix n°1.

Table II. 3. Mulliken atomic charges for A0 by DFT/B3LYP method with 6-311G++ (d,p) and LanL2DZ basis sets.

Atoms	6-311G++(d,p)	LanL2dz
Fe	0.561	-0.154
C1	-0.204	-0.232
C2	-0.227	-0.236
C3	-0.197	-0.236
C4	-0.247	-0.232
C5	-0.185	-0.245
C6	-0.352	-0.353
C7	-0.385	-0.229
C8	-0.259	-0.229
C9	-0.095	-0.353
C10	0.662	-0.305
C11	-1.140	-0.444
C12	-0.324	0.436
C13	-0.124	-0.412
C14	-0.272	-0.187
C15	-0.373	-0.265
C16	-0.232	-0.218
C17	0.446	-0.344
N1	0.206	-0.446

II.3.3. Molecular electrostatic potential map (MESP)

The MESP simultaneously exhibits molecular size, shape as well as positive, negative and neutral electrostatic potential regions in terms of color grading and is very useful in research of molecular structure with its physicochemical property relationship. The MESP generally show that the maximum positive region which preferred site for nucleophilic attack symptoms as blue color, while the maximum negative region which preferred site for electrophilic attack indications as red color. The green regions signify electrically neutral regions. Potential decreases in this order: red < orange < yellow < green < blue.

The MESP surface map and contour map of A0 (Fig. II 2) show the one region characterized by blue color (positive electrostatic potential) around the hydrogen binded to the nitrogen atom which explain the ability for a nucleophilic attack on these positions. The MESP of the other compounds are presented in appendix n°1.

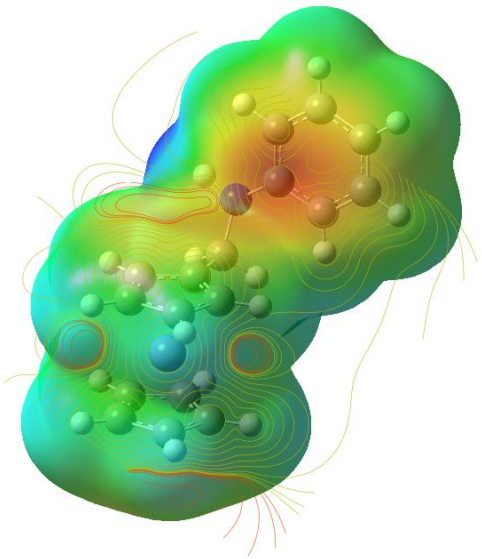


Figure II. 2. MESP Surface and contour map for A0.

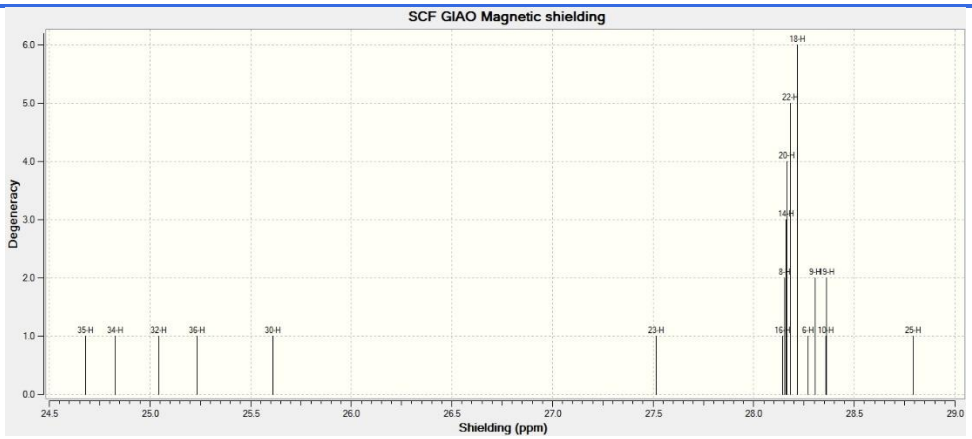
II.3.4. NMR Spectrum Analysis

The optimized structure was used to calculate the ^{1}H and ^{13}C NMR spectra. it is apparent from ^{1}H NMR that the experimental chemical shifts of the ferrocenic part of A0 are around 4.00 ppm, whereas, the chemical shifts of both computational basis 6-311G++(d,p) and LanL2DZ are between 2.86 and 3.71 ppm where the chemical shifts obtained from LanL2dz basis set are lower. On the other hand, the chemical shifts for the rest of A0 are in the range of 4.01 - 6.81 ppm in experimental and 2.06 - 7.20 ppm in both computational basis. It is clear that the experimental and calculated data are very close to each other, where the values of 6-311G++ (d, p) are closer.

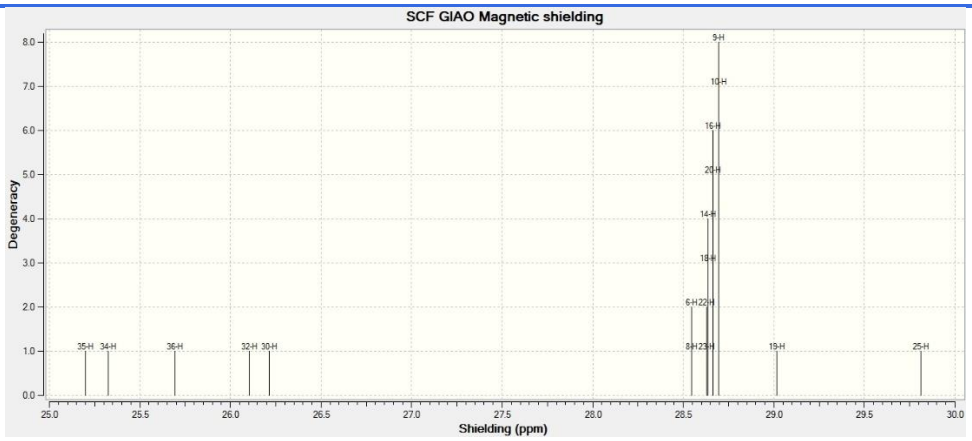
The ^{13}C NMR of A0 shows that the experimental chemical shifts of ferrocenic part are around 68 ppm, while the calculated ones are in the range of 59.80 - 73.41. The aromatic carbon

atoms generally give rise to signals in the range of 100–150 ppm. The aromatic carbons of A0 observed in the expected range from 104.12 to 153.56 ppm. The least chemical shifts are for C11 between 39.74 and 44.39 ppm. It can be concluded that there is a good agreement between the experimental and the computed data. The ^{1}H and ^{13}C NMR spectra of the other compounds are presented in appendix n°1.

^{1}H NMR by 6-311G++(d,p)



NMR ^{1}H by LanL2dz



Experimental NMR ^{1}H

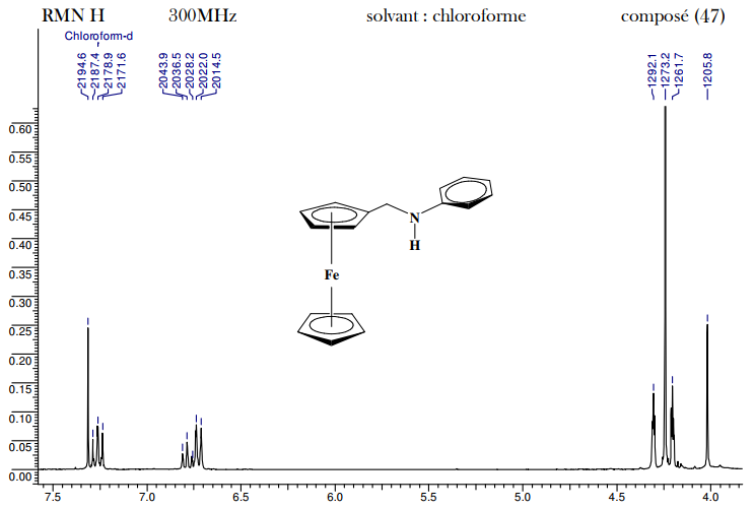
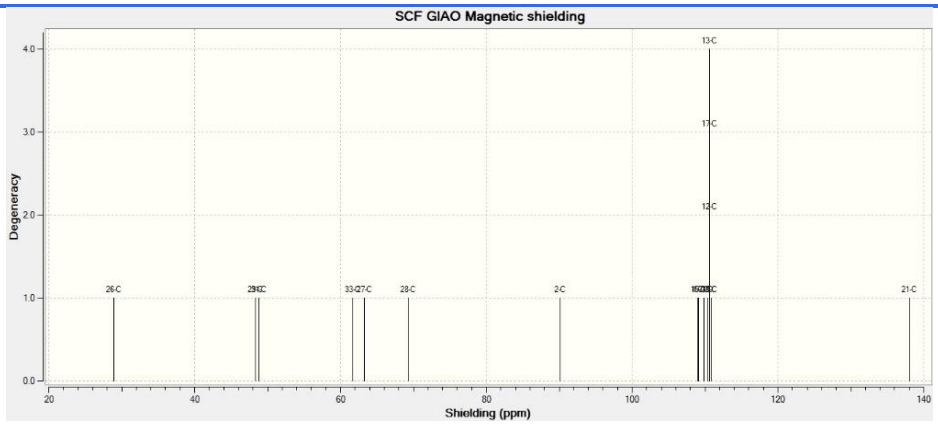
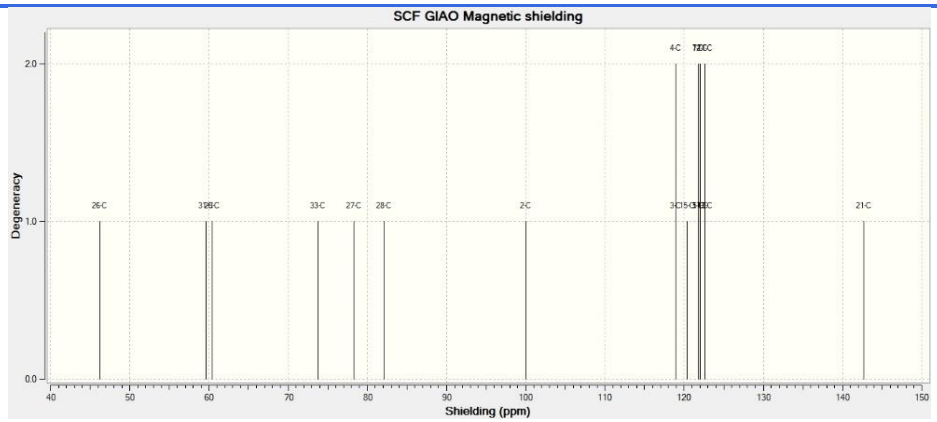


Figure II. 3. Optimized and experimental NMR ^{1}H Spectra of A0

NMR ^{13}C by 6-311G++(d,p)



NMR ^{13}C LanL2dz



Experimental NMR ^{13}C

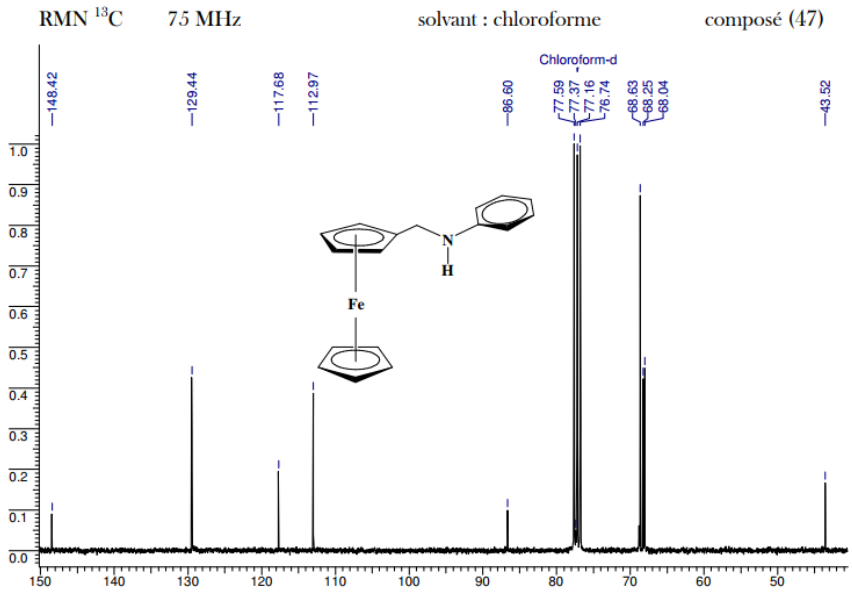


Figure II. 4. Optimized and experimental NMR ^{13}C Spectra of A0

Table II. 4. Calculated ^{1}H NMR and ^{13}C NMR chemical shift (ppm) of A0 $CDCl_3$ by experimental and DFT/B3LYP methods

NMR H^1				NMR C^{13}			
Atom	Exp.	a	b	Atom	Exp.	A	b
H6	4.20	3.71	3.33	C1	68.04	71.84	60.40
H9	4.20	3.59	3.33	C2	68.04	71.84	59.80
H7	4.24	3.52	3.22	C3	68.04	71.84	59.80
H8	4.24	3.59	3.22	C4	68.04	71.84	60.40
H5	4.24	3.52	2.86	C5	68.04	73.33	62.05
H1	4.24	3.71	3.22	C6	68.25	73.41	63.55
H2	4.24	3.71	3.22	C7	68.25	72.56	60.70
H3	4.24	3.71	3.22	C8	68.25	71.55	60.70
H4	4.24	3.71	3.22	C9	68.63	72.14	63.55
H11a	4.30	4.36	3.22	C10	86.60	92.33	82.48
H11b	4.30	3.71	3.22	C11	43.52	44.39	39.74
H12	4.01	3.08	2.06	C12	148.42	153.56	136.23
H17	6.73	6.83	5.77	C13	117.68	119.19	104.12
H16	6.81	7.20	6.68	C14	129.44	134.08	122.03
H15	6.78	6.64	6.18	C15	129.44	120.73	108.72
H14	6.76	7.05	6.55	C16	129.44	133.60	122.81
H13	6.71	6.27	5.66	C17	117.68	113.20	100.34

a : 6-311G++ (d, p) , b: LanL2DZ

II.3.5. Vibrational analysis

A0 has 102 vibration modes. All those vibrational modes are active in IR (figure II.5). The most important stretching and bending vibrations along with their probable assignments of title molecules are given in table II. 5, All of these vibrations were measured by the computational methods of DFT / B3LYP with 6-311G++(d,p) and LanL2DZ basis sets. The computational frequencies were compared to experimental frequencies[11].

It can be seen from the results that the simulated frequencies by Gaussian 09 software were higher than experimental frequencies. So these frequencies should be multiplied by Scalling factor to be close to the experimental frequencies. The FT-IR spectra of the other compounds are presented in appendix n°1.

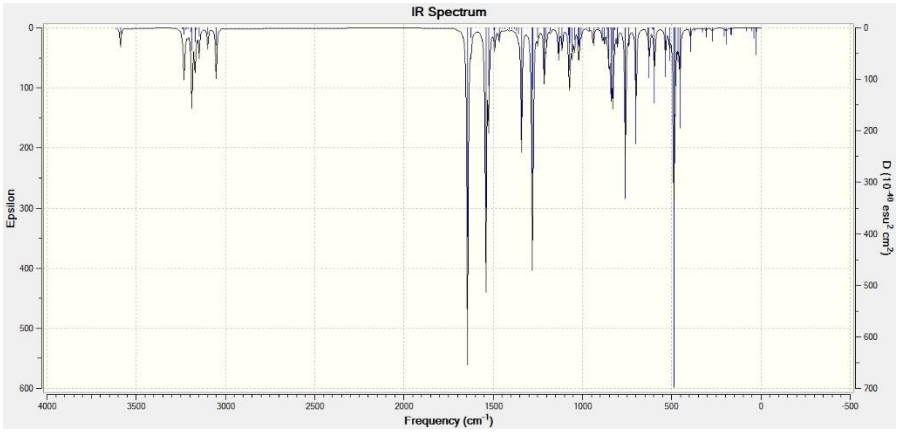


Figure II. 5.Theoretical IR spectrum of A0.

- **C – H vibrations**

Aromatic compounds commonly exhibit multiple weak bands in the region of 3100–3000 cm^{-1} due to aromatic C–H stretching vibration [23], in this derivative there is one benzene and two cyclopentadienyl (Cp) rings as well, which appeared as symmetric and asymmetric stretching vibrational frequency in the range of 2972 -2916 cm^{-1} with 6-311G++(d,p) and 3053 – 2976 cm^{-1} with LanL2dz basis sets, as it's presented in the table II.5 . The methylene bridges belonged to C_{11} of A0 was observed as symmetric stretching vibration at 2807 cm^{-1} (2790cm^{-1}) for 6-311G++ (d,p) (LanL2dz).

The molecule has bending C-H as well. The methylene bridge on C11 was observed as out of plane bending vibration at 1417 and 429cm^{-1} for 6-311G++ (d,p) and LanL2dz basis sets. The aromatic and cyclopentadienyl C-H in and out of plane bending vibrational frequency appeared in the range of 1211- 698cm^{-1} with both basis sets.

The experimental bands for C-H stretching vibration appeared in the range of 3230 – 3051 cm^{-1}.

- **C=C and C-C vibrations**

The carbon-carbon stretching vibrations of aromatic ring are expected in the range of 1650 -1200 cm^{-1} [24]. In A0 derivative, C–C bonds of aromatic rings had the strongest symetric stretching vibrational frequency in the frequency region of 1512 cm^{-1} that was obtained using 6-311G ++ (d, p), likewise, in 1545 cm^{-1} using LanL2dz. Indeed, it can be seen in experimental spectra that this bond is located in 1602 cm^{-1}, which agreed with the calculated result. These bands match to C=C stretching vibrations.

Moreover, the C-C bonds correspond to meduim frequencies of carbon-carbon stretching vibrations, which were calculated in the range of 1136 – 986 cm^{-1}. Furthermore, it shows medium in plane bending vibration in the range of 1403 – 1234 cm^{-1}. However, the experimental spectra shows similar result at 1100 cm^{-1}.

- **C-N vibrations**

It can be seen from the table II.5 the band at 1403 (1360 cm^{-1}) using 6-311G++(d,p) (LanL2dz) are assigned to C–N medium stretching vibration. Another similar strong band appeared only at LanL2dz in the region 1271 cm^{-1}.

The theoretically predicted scaled values show excellent agreement with experimental data.

- **Fe-Cp vibrations**

The Fe-Cp stretching vibrational frequency are generally observed in the region of 700 - 400 cm^{-1} [25, 26]. As it is expected, the Fe-Cp stretching vibrations appeared in the range of 448 - 401 cm^{-1}.

- **N-H vibrations**

As it is shown in table II.5, the NH bond appeared as weak stretching vibration at 3300 and 3400cm^{-1} with 6-311G++(d, p) and LanL2dz basis sets, respectively. Moreover, it appeared as out of plane bending at 1417 and 548 cm^{-1} using 6-311G++(d, p) and 1429 and 298 using LanL2dz basis set.

Table II. 5. The observed FT-IR of A0 and its calculated wavenumbers (in cm^{-1}) with their probable assignments

Experimental frequency (cm^{-1})	**Calculated frequencies**				**Vibrational Assignments**
	6-311G++(d,p)		**Lan2dz**		
FT-IR	unscaled	scaled	unscaled	scaled	
3405	3587	3300	3656	3400	υ_w (NH) sy
	3230	2972	3283	3053	υ_w ($C_1H_1C_2H_2C_3H_3C_4H_4C_5H_5C_6H_6C_7H_7$ $C_8H_8C_9H_9$)sy
3090	3187	2932	3220	2995	υ_m ($C_{13}H_{13}C_{14}H_{14}C_{15}H_{15}C_{16}H_{16}C_{17}H_{17}$)asy
2840	3170	2916	3200	2976	υ_m ($C_{13}H_{13}C_{14}H_{14}C_{15}H_{15}C_{16}H_{16}C_{17}H_{17}$)asy
2208	3051	2807	3000	2790	υ_w ($C_{11}H_{11a}H_{11b}$)sy
1602	1643	1512	1661	1545	υ_s($C_{13}C_{14}$ C_{15} C_{16})sy
	1540	1417	1537	1429	α_s($C_{11}H_{11a}H_{11b}$)asy+ α_s (NH) asy
	1525	1403	1462	1360	α_m($C_{14}C_{15}C_{17}C_{12}$)asy+ υ_m (C_{12}N) sy
1100	1341	1234	1395	1297	α_m($C_{14}C_{15}C_{17}C_{12}$)asy+ α_m ($C_{11}H_{11a}H_{11b}$)asy
	-	-	1367	1271	π_s ($C_{16}H_{16}C_{17}H_{17}$) + υ_s (C_{12}N) sy
	1280	1178	-	-	π_s ($C_1H_1C_2H_2C_3H_3C_4H_4C_5H_5C_6H_6C_7H_7$ $C_8H_8C_9H_9$)asy+ α_s ($C_{11}H_{11a}H_{11b}$)sy
	1276	1174	1302	1211	π_m ($C_6H_6C_7H_7$ $C_8H_8C_9H_9$)asy
	1213	1116	1221	1136	υ_m(C_6C_{10} C_9)asy+ α_m($C_{11}H_{11a}H_{11b}$)sy
	1072	986	1095	1018	υ_m(C_6C_{10} C_9)asy+ π_m ($C_2H_2C_3H_3C_4H_4C_5H_5$)sy
999	829	763	816	759	α_m ($C_1H_1C_2H_2C_3H_3C_4H_4C_5H_5C_6H_6C_7H_7$ $C_8H_8C_9H_9$)asy
	759	698	783	728	α_m ($C_{13}H_{13}C_{14}H_{14}C_{15}H_{15}C_{16}H_{16}C_{17}H_{17}$)asy
	596	548	320	298	α_m (NH)asy
	487	448	444	413	υ_s ($C_1FeC_2FeC_3FeC_4FeC_5FeC_6Fe$ $C_7FeC_8FeC_9Fe$)sy
	479	441	443	412	υ_m ($C_1FeC_2FeC_3FeC_4FeC_5FeC_6Fe$ $C_7FeC_8FeC_9Fe$)asy
	452	416	431	401	υ_m ($C_1FeC_2FeC_3FeC_4FeC_5FeC_6Fe$ $C_7FeC_8FeC_9Fe$)asy

asy – asymmetric, Sy – symmetric, υ – stretching, π- in plane bending, α- out of plane bending, s - strong, m - medium. The scaling factor is equal to 0.92 for 6-311G++(d,p) and 0.93 for LanL2DZ.

II.3.6. Frontier molecular orbitals

The lowest-lying unoccupied molecular orbitals (LUMO) and the highest occupied molecular orbitals (HOMO) are known as frontier molecular orbitals (FMO). The HOMO represents the capacity to give an electron, while LUMO as an electron acceptor represents the capacity to acquire an electron. The energy gap between HOMO and LUMO decides the chemical reactivity, kinetic stability and, optical polarizability and chemical hardness–softness of a molecule [36, 37]. In the current study, the HOMO, LUMO and the energy gap between them (ΔE) were calculated at both level of theory 6-311G++ (d,p) and LanL2DZ. The energies and the contour diagrams of these molecular orbitals are shown in figure II.6.

It can be seen from figure II.6 that the HOMOs mainly concentrate on the ferrocenic part of the molecule, however, the LUMOs are distributed over the aromatic ring. From the energy gap (ΔE) of the studied molecule, A25 is predicted to be the most reactive with ΔE energy gap equal to 2.612 and 2.334eV with 6-311G++(d,p) and LanL2DZ, respectively. The HOMO and LUMO energies and thier energetic gaps of the other compounds are presented in appendix n°1.

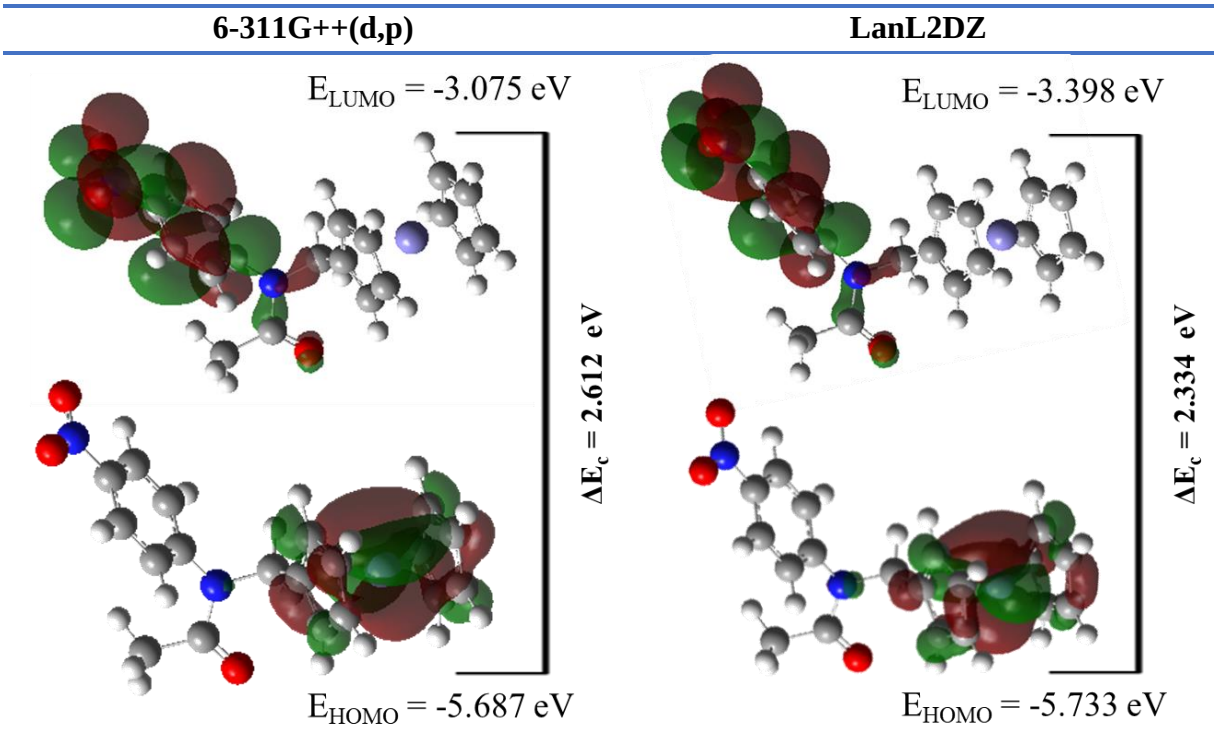


Figure II. 6. Frontier molecular orbitals of A25 by DFT/B3LYP methods with 6-311G++(d, p) and LanL2DZ basis sets .

II.3.7. ADME and Pharmacokinetics properties

SwissADME online web tool provided by the Swiss Institute of Bioinformatics (SIB) was utilized for the computation of the ADME parameters, pharmacokinetic properties and drug-like nature of the compounds A0-A28 to assure that they are not only promising candidates in terms of biological efficacy but also from the pharmacokinetic aspects.

II.3.7.1. Drug-likeness properties

Among all 117 different routes of administration for a drug molecule approved by the US Food and Drug Administration (FDA), oral administration is the most popular one. A patient finds it easier to open the lid of a bottle and take the tablets/capsules rather than taking an injection intravenously. Orally administered medications are ingested in the mouth, after which it travels through a long journey through the human body until drugs reach the interaction site and exerts its pharmacological effect [15].

The results have shown that most of tested compounds fulfil the drug-likeness characteristics as defined by the major pharmaceutical companies; Lipinski's (Pfizer) [27], Ghose's (Amgen)[28], Veber's (GSK) [29] , Egan's (Pharmacia) [30] and Muegge's (Bayer) [31] filters. However, some of the compounds have shown 1 violation (MLOGP>4.15) with Lipinski rule of five, as well as other violations was predicted in Muegge rule (XLOGP3>5, Heteroatoms<2). The oral bioavailability score was predicted to be 0.55 for all compounds, which indicate a good oral administration of the studied compounds.

Table II. 6 . Drug-likeness and bioavailability of the studied compounds.

Comp.	Lipinski #violations	Ghose #violations	Veber #violations	Egan #violations	Muegge #violations	Bioavailability Score
A0	0	0	0	0	1	0.55
A1	0	0	0	0	0	0.55
A2	0	0	0	0	0	0.55
A3	1	0	0	0	1	0.55
A4	0	0	0	0	1	0.55
A5	0	0	0	0	2	0.55
A6	0	0	0	0	2	0.55
A7	1	0	0	0	2	0.55
A8	1	0	0	0	2	0.55
A9	1	0	0	0	2	0.55
A10	1	0	0	0	2	0.55
A11	1	0	0	0	2	0.55
A12	1	0	0	0	2	0.55
A13	0	0	0	0	1	0.55
A14	0	0	0	0	0	0.55
A15	0	0	0	0	0	0.55
A16	0	0	0	0	1	0.55
A17	0	0	0	0	0	0.55
A18	0	0	0	0	0	0.55
A19	0	0	0	0	0	0.55
A20	0	0	0	0	0	0.55
A21	0	0	0	0	0	0.55
A22	0	0	0	0	0	0.55
A23	0	0	0	0	0	0.55
A24	0	0	0	0	0	0.55
A25	0	0	0	0	0	0.55
A26	0	0	0	0	0	0.55
A27	0	0	0	0	0	0.55
A28	0	0	0	0	0	0.55

II.3.7.2. Pharmacokinetic properties

The pharmacokinetic properties are very important when predicting the behaviors of newly-developed drugs including blood-brain barrier (BBB), gastrointestinal absorbtion (GI absorbtion), P-glycoprotein substrate (P-pg substrate) as well as drug metabolism.

Permeable drugs primarily cross biological barriers including the intestinal epithelial and the BBB (blood-brain barrier) by the mechanism of passive diffusion, where substances are transported by the effect of a concentration gradient. Basically, there are two types of passive diffusion, one is the paracellular transport while the other is the transcellular transport mechanism; other drugs are being transported by either the carrier-mediated or the P-pg mediated transport.

The process of metabolism is a very complex process that involves various enzymatic activities that vary among individuals due to different genetic factors. The cytochrome P450 (CYP) is considered to be the most influential enzyme in the drug metabolism, which led to the development of many models such as QSAR for the prediction of the metabolism of molecules by the CYP enzyme [16, 32].

All compounds showed high gastrointestinal absorption (GI absoption) with BBB permeability which is due to their optimal physicochemical properties situated in the suitable physicochemical properties range for oral bioavailability. Moreover, some of the compounds are not P-glycoprotein substrates, which indicate that they are not susceptible to the efflux mechanism performed by this transporter which is used by many tumor cell lines as a drug-resistance mechanism [32].

In the family of CYP enzymes, the CYP3A4 was the most important enzyme on account of metabolizing 50% of all drugs by itself and the CYP2C9 enzyme mainly metabolizes several clinically used drugs such as celecoxib and diclofenac [33]. It can be seen from table II. 7 that all compounds were predicted to not be CYP2C9, CYP2D6 and CYP3A4 inhibitors, however, some of them were predicted to be CYP1A2 and CYP2C19 inhibitors.

Table II. 7. ADME prediction of the studied compounds.

Comps.	GI absorption	BBB permeant	Pgp substrate	CYP1A2 inhibitor	CYP2C19 inhibitor	CYP2C9 inhibitor	CYP2D6 inhibitor	CYP3A4 inhibitor
A0	High	Yes	Yes	No	No	No	No	No
A1	High	Yes	Yes	No	No	No	No	No
A2	High	Yes	Yes	No	No	No	No	No
A3	High	Yes	Yes	No	No	No	No	No
A4	High	No	No	No	No	Yes	No	Yes
A5	High	Yes	Yes	No	No	No	No	No
A6	High	Yes	Yes	No	No	No	No	No
A7	High	Yes	Yes	No	No	No	No	No
A8	High	Yes	Yes	Yes	No	No	No	No
A9	High	Yes	Yes	Yes	No	No	No	No
A10	High	Yes	Yes	Yes	No	No	No	No
A11	High	Yes	Yes	Yes	No	No	No	No
A12	High	Yes	Yes	Yes	No	No	No	No
A13	High	Yes	Yes	Yes	No	No	No	No
A14	High	Yes	Yes	Yes	No	No	No	No
A15	High	Yes	Yes	Yes	No	No	No	No
A16	High	Yes	No	Yes	No	No	No	Yes
A17	High	Yes	No	Yes	No	No	No	No

A18	High	Yes	No	Yes	No	No	No	No
A19	High	Yes	Yes	No	No	No	No	No
A20	High	Yes	Yes	No	No	No	No	No
A21	High	Yes	Yes	No	Yes	No	No	No
A22	High	Yes	Yes	No	Yes	No	No	No
A23	High	Yes	No	Yes	No	No	No	Yes
A24	High	Yes	No	Yes	No	No	No	Yes
A25	High	Yes	No	Yes	No	No	No	Yes
A26	High	Yes	Yes	No	No	No	No	No
A27	High	Yes	Yes	No	No	No	No	No
A28	High	Yes	Yes	No	No	No	No	No

II.4. Conclusion

A computational work in chemistry have not been carried out to examine the structural and spectroscopic properties of ferrocenylmethylanilines derivatives. Thus, the structure of the title molecules were studied using the theoretical methods of DFT / B3LYP with 6-311G++(d,p) and LanL2DZ basis sets.

After simulating the molecular structure of the candidate drugs with Gaussian 09 software and obtaining bond lengths and angles between its constituent atoms, a comparison with the experimental XRD data has shown a good agreement.

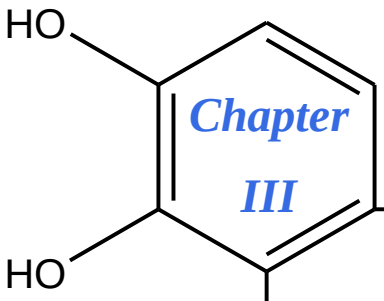
A vibrational analysis has been studied and functional groups were determined using both experimental and theoretical methods. Computational ^{1}H and ^{13}C chemical shift values were reported and compared with experimental data, showing good agreement for both ^{1}H and ^{13}C. HOMO and LUMO energy gaps explain the reactivity of studied compounds, where A23 and A25 were predicted to be the most reactive with least ΔE energy gaps. The MESP at the 6-311G++(d,p) optimized geometry was calculated to predict the reactive sites for electrophilic and nucleophilic attack for the studied molecules. Finally, the ADME parameters, pharmacokinetic properties and drug-like nature were predicted and showed good bioavailability and absorption.

References

1. Kealy, T. and P. Pauson, A new type of organo-iron compound. Nature, 1951. **168**(4285): p. 1039-1040.
2. Wilkinson, G., et al., The structure of iron bis-cyclopentadienyl. Journal of the American Chemical Society, 1952. **74**(8): p. 2125-2126.
3. Skrehota, O., Quantitative structure-property relationship modeling algorithms, challenges and IT solutions. 2010, Thesis. Masaryk University Faculty of Informatics. Czechoslovakia.
4. Radhia, M., Nouvelle approche des relations structures-activités dans des molécules antibiotiques. 2015, University of Mohamed Khider Biskra.
5. Prasad, O., L. Sinha, and N. Kumar, Theoretical Raman and IR spectra of tegafur and comparison of molecular electrostatic potential surfaces, polarizability and hyerpolarizability of tegafur with 5-fluoro-uracil by density functional theory. J At Mol Sci, 2010. **1**: p. 201-214.
6. SALAH, T., Etude par la chimie computationnelle des corrélations 2D-QSAR et 3D-QSAR de quelques composés bioactifs. 2017, University Mohammed khidher of Biskra. p. 130.
7. Broughton, H.B. and I.A. Watson, Selection of heterocycles for drug design. Journal of Molecular Graphics and Modelling, 2004. **23**(1): p. 51-58.
8. Lessigiarska, I., Development of structure-activity relationships for pharmacotoxicological endpoints relevant to European Union legislation. 2006, Liverpool John Moores University.
9. Lewars, E.G., Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics2011, London New York: Springer Dordrecht Heidelberg.
10. Belgacem, T., synthèse, cyclisation, étude electrochimique et structurale de quelques N-acyl-n'alkylférrocenyl méthyle hydrazide. 2007, university Elhadj Lakhder of batna.
11. Oumelkhair, R., Synthèse et étude électrochimique et structurale de quelques N-ferrocényl méthyle anilines substitués Applications à l'inhibition de la corrosion aqueuse. 2014, University Kasdi Merbah of Ouergla.
12. Hadia, H., Synthèse, caractérisation et évaluation in vitro de l'activité biologique et du pouvoir mutagène de quelques N-(ferrocenylmethyl)nitroaniline. 2017, University of El Oued.

13. Meriem, H., Synthèse, caractérisation et évaluation in vitro de l'activité biologique et du pouvoir mutagène de quelques N-(ferrocenylmethyl) aminobenzonitrile, in chemistry departement 2017, university of Eloued.
14. Elhafnaoui, L., Synthesis, characterization and in vitro and in silico study of biological activities of some ferrocenylmethyl-nucleic bases derivatives. 2020, University Kasdi Merbah of Ouargla.
15. Zhong, H.A., ADMET properties: Overview and current topics. Drug design: Principles and applications, 2017: p. 113-133.
16. Chandrasekaran, B., et al., Computer-aided prediction of pharmacokinetic (ADMET) properties, in Dosage Form Design Parameters. 2018, Elsevier. p. 731-755.
17. Rahnamaye Aliabad, H. and M. Chahkandi, Comprehensive SPHYB and B3LYP-DFT Studies of Two Types of Ferrocene. Zeitschrift für anorganische und allgemeine Chemie, 2017. **643**(6): p. 420-431.
18. Yáñez-S, M., et al., Theoretical assessment of TD-DFT applied to a ferrocene-based complex. Computational and Theoretical Chemistry, 2017. **1118**: p. 65-74.
19. Karabacak, M., et al., DFT calculations and experimental FT-IR, FT-Raman, NMR, UV–Vis spectral studies of 3-fluorophenylboronic acid. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2015. **136**: p. 306-320.
20. Sivaranjani, T., S. Xavier, and S. Periandy, NMR, FT-IR, FT-Raman, UV spectroscopic, HOMO–LUMO and NBO analysis of cumene by quantum computational methods. Journal of Molecular Structure, 2015. **1083**: p. 39-47.
21. Ghosh, K., et al., Three mononuclear octahedral cobalt (III) complexes with salicylaldimine Schiff bases: Synthesis, characterization, phenoxazinone synthase mimicking activity and DFT study on supramolecular interactions. Polyhedron, 2016. **112**: p. 6-17.
22. Tomberg, A., GAUSSIAN 09W TUTORIAL AN INTRODUCTION TO COMPUTATIONAL CHEMISTRY USING G09W AND AVOGADRO SOFTWARE.
23. Varsányi, G., M.A.e. Kovner, and L. Láng, Assignments for vibrational spectra of 700 benzene derivatives. 1973: Akademiai Kiado.
24. Demir, S., et al., Vibrational spectroscopic studies, NMR, HOMO–LUMO, NLO and NBO analysis of 1-(2-nitrobenzoyl)-3,5-diphenyl-4,5-dihydro-1 H -pyrazole with use X-ray diffractions and DFT calculations. Journal of Molecular Structure, 2016. **1108**: p. 637-648.
25. Lippincott, E.R. and R.D. Nelson, The vibrational spectra and structure of ferrocene and ruthenocene. The Journal of Chemical Physics, 1953. **21**(7): p. 1307-1308.

26. Noh, D.-Y., et al., Syntheses and characterization of heterobimetallic complexes (dppf) Pt (dithiolate)(dppf: bis (diphenylphosphino) ferrocene); X-ray crystal structures of (dppf) PtL where L= dmit, phdt and i-mnt. Polyhedron, 2001. **20**(15-16): p. 1939-1945.
27. Lipinski, C.A., Lead-and drug-like compounds: the rule-of-five revolution. Drug Discovery Today: Technologies, 2004. **1**(4): p. 337-341.
28. Ghose, A.K., V.N. Viswanadhan, and J.J. Wendoloski, A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. Journal of combinatorial chemistry, 1999. **1**(1): p. 55-68.
29. Veber, D.F., et al., Molecular properties that influence the oral bioavailability of drug candidates. Journal of medicinal chemistry, 2002. **45**(12): p. 2615-2623.
30. Egan, W.J., K.M. Merz, and J.J. Baldwin, Prediction of drug absorption using multivariate statistics. Journal of medicinal chemistry, 2000. **43**(21): p. 3867-3877.
31. Muegge, I., S.L. Heald, and D. Brittelli, Simple selection criteria for drug-like chemical matter. Journal of medicinal chemistry, 2001. **44**(12): p. 1841-1846.
32. Daina, A. and V. Zoete, A boiled-egg to predict gastrointestinal absorption and brain penetration of small molecules. ChemMedChem, 2016. **11**(11): p. 1117-1121.
33. Daina, A., O. Michielin, and V. Zoete, SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Scientific reports, 2017. **7**: p. 42717.



Chapter III

In vitro and in silico evaluation of anti-proliferative activity against breast cancer

III.1. Introduction

Depending on the type of breast cancer, different types of drug treatment might be used, including chemotherapy, hormone therapy, targeted therapy and immunotherapy. Hormone therapy treats the types of breast cancer affected by hormones, like estrogen and progesterone. In this case, breast cancer cells have receptors (proteins) that attach to estrogen and progesterone, which helps them to grow. Therefore, these treatments can either prevent these hormones from attaching to their receptors, or reduce their level. It can reach cancer cells almost anywhere in the body and not only in the breast. It is recommended for women with tumors that are hormone receptor-positive (ER+), which account for about 70% of breast cancers.

In this chapter, we aimed to find new potentially active ferrocenic compounds against hormone dependent breast cancer. First, the anti-proliferative activity of a series of 29 ferrocenylmethylanilines against both hormone-dependent (MCF7) and -independent (MDA-MB-231) breast cancer cell lines was performed using MTT assay. Moreover, docking study was carried out against Estrogen and Progesterone receptor to understand the efficacy of the investigated compounds against the hormone dependent cells. Finally, another docking study was performed with aromatase and 17β-HSD1, the enzymes that catalyse the estrogen synthesis.

III.2. Chemistry

All compounds (Table II. 1, chapter II) were synthetized in the Laboratory of "Valorisation et Technologie des Ressources Sahariennes (VTRS)" of the university of El-Oued, following our previously reported procedure [1-5].

III.3. Anti-proliferative activity against MDA-MB-231 and MCF-7

III.3.1. General terms

- **Cell line:** A culture which is subcultured beyond the initial primary culture phase, it has an unlimited number of passages, giving a continuous line.
- **Primary culture**: The initial culture derived from in vivo material
- **Cell culture**: Lab technique for bringing cells to life in vitro.
- **Substrate:** The container on which a culture is grown (the flask).
- **Passage:** Subculture of cells from one container to another.
- **Confluent:** Situation wherein cells fully cover the substrate.
- **Medium**: It is a liquid which support the survival of cells in their incubation environment, it generally contains an appropriate source of energy and compounds which regulate the cell cycle.

III.3.2. Cell culture

The initial culture taken directly from an individual is referred to as the "primary culture" and when diluted and transferred into further containers (a procedure called "subculture" or "passage"), it becomes a "cell line." Cell lines may be categorized as either "continuous" lines, with the potential for unlimited population expansion, or "finite," lines, which undergo a limited number of population doublings before "senescence." Becoming a continuous cell line necessitates "transformation" and this requires either the presence of cells already transformed at the beginning of the culture or undergoing transformation in the early generations.

Cell lines can either exist as adherent cultures or they may grow in "suspension." Most cell types adhere to a "substrate" such as plastic or glass and proliferate as a monolayer, whereas suspension cultures do not adhere to a substrate surface and will float in the medium[6].

III.3.2.1. Basic Requirements of Cell Culture

- **Cell Culture Hood:** working under aseptic conditions is a basic necessity of all cell culture laboratories worldwide. Successful research in the cell culture, as well as obtaining accurate and reproducible results, relies on full control of undesired contaminations. In order to avoid contamination from airborne particles, all manipulations with cells are carried out in a cell culture hood (biosafety cabinet), a

closed workspace where the air is continuously filtered. Before and after work, the hood and all the items that you bring into the hood must be cleaned with disinfection solution (e.g., 70% ethanol or isopropanol) [7].

- **Media:** Cell culture media are the key constituent providing an environment ensures the survival, further continuous proliferation, and/or differentiation for cells that have been either freshly explanted from an organism or have been transformed from other primary cultures. Cell culture media shall in the first line provide the energy sources, confer oxygen to cells, contain a salt composition and pH that are beneficial to the cells being cultivated, and take up metabolites and debris of cultured cells. Additionally, culture media are often used to provide signals for the survival, growth, and/or differentiation of the cultured cells. Thus, cell culture media must replace the natural environment of the cultured cells in fluid form.

  Dulbecco's Modified Eagle's Medium (DMEM) is rich medium, containing high concentration of amino acids and vitamins comparing to another medium named EMEM. Additionally, ferric nitrate, sodium pyruvate, and other amino acids are included. Originally, it contained glucose at 1000 mg/L, but a variation with 4500 mg/L glucose has been proved to be better to culture a number of cells. DMEM is a basal medium and does not contain proteins or growth-promoting agents and requires fetal calf serum (FBS) supplementation. DMEM contains sodium bicarbonate buffer system (3.7 g/L) and required the presence of CO_2 to preserve the required pH. DMEM is applied widely for culturing primary mouse, chicken, and human cells including fibroblasts. It is used as a basal medium also for embryonic stem cells[7].

- **Serum:** Although media contains many of the essential nutrients needed for growth, serum provides additional key elements. Serum supports the survival and growth of cells in culture to the extent that it can replace many of the in vivo hormonal, nutritional, and stromal elements present in the in vivo cell environment. Serum proteins include hormones, growth factors, lipids, transport (binding) proteins, enzyme cofactors, and attachment factors. The concentrations of individual components of serum will differ with the age and health status of the animals of origin, therefore, fetal and new-born calf sera are commonly used in most cancer culture studies although both human and equine sera are also used. Typically, concentrations of 5–20% serum in media are considered optimal depending on cell type. Higher percentages generally add little benefit for increased cost[6].

- **The Substrate:** Almost all types of cancer cells require interaction with a substrate for their growth and differentiation in culture. Most cancer cell types will grow as monolayers on plastic or glass. Modern cell culture flasks and microplates are disposable, single-use vessels which are mainly made of polystyrene due to its good optical properties, easy casting during manufacturing, reasonable costs, and surface modification possibilities[7].
- **Physical Environment:** Maintaining the physical environment conditions within certain limits is important for optimal cell growth. The preferred temperature for most mammalian cultures is 36.5° ± 1°C, although cells can still grow at lower temperatures, temperatures above 40°C leads quickly to the cell death. Culture media contain buffering systems which generally require a CO_2 atmosphere and frequently 5% CO_2 is preferred. Most cancer cells need a pH value of about 7.2–7.4 and media should be periodically changed as cell growth creates respiratory by-products that acidify the media. The pH indicator commonly added to medium is phenol red[6].

In our case, the medium includes 500ml of DMEM, 5ml of L-Glutamin, 5ml of Pen Strep Solution and 50ml of Fetal Bovine Serum (FBS). Ethanol 70% was used for sterilization of the hood and all materials. The 75ml flasks were used, we added 13 ml of the medium and carefully added 1000µl of the cells which had been frozen at -80°C. The flasks were then incubated at 37°C in a humidified atmosphere containing 5% CO_2. After 48 hours the cells were multiplied, we poured the medium and added Phosphate Buffered Saline solution (PBS) which is used to wash the cells before dissociation, transportation or diluting cells for counting, we poured it and added 500µl of Trypsin EDTA which helps to remove the cells from the substrate. Finally, we added 3.5ml of the medium and the cells are ready for the cytotoxicity test.

Afterward, we counted the cells in the obtained mixture using a drop of Trypan Blue and hymocytometer and then made dilution to obtain 4000 cells per well.

III.3.3. Breast cancer cell lines (MCF-7 and MDA-MB-231)

III.3.3.1. MCF-7

MCF7 is the most studied human breast cancer cell line in the world, it was named after the Michigan Cancer Foundation (MCF) and it was derived from the pleural effusion from a 69 years old female with breast adenocarcinoma [8, 9].

MCF7 cells are of interest as they retain a variety of characteristics similar to mammary epithelium. It is one of few breast cancers which express the oestrogen receptor alpha (ERα) [10-12]. Moreover, these cells express androgen, progesterone and glucocorticoid receptors

which make them valuable tools in medical research. Treatment with oestrogens for MCF7 has been shown to have an anti-apoptotic effect [13, 14], whereas treatment with anti-oestrogen chemotherapy drugs (e.g. Tamoxifen) can reduce growth of cultures by inhibiting proliferation and inducing apoptosis [15, 16]. MCF7 are mainly used as an in vitro model for breast cancer biology researches, because of the number of variants available. It is used in development of chemotherapeutic drugs and understanding drug resistance.

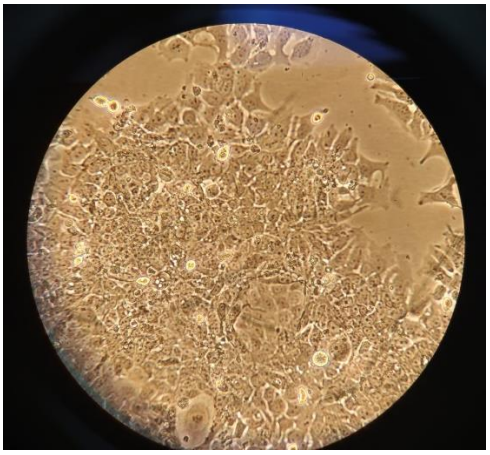


Figure III. 1. MCF-7 cell lines.
(personal photo.)

III.3.3.2. MDA-MB-231

The MDA-MB-231 cell line is an epithelial, human breast cancer cell line isolated at MD Anderson cancer center from a pleural effusion of a 51 years old Caucasian female with a metastatic mammalian adenocarcinoma. It is one of the most widely used breast cancer cell lines in medical laboratories [17, 18].

The MDA-MB-231 cell line is a triple-negative breast cancer (TNBC) cell line, highly aggressive, invasive and poorly differentiated. TNBCs lacks estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2). The invasiveness of the MDA-MB-231 cells is mediated by proteolytic degradation of the extracellular matrix [18, 19]. MDA-MB-231 cells are different from other breast cancer cells due to the unique presence of several proteins including MKP3 and Nedd4 [20]. Also, MDA-MB-231 cells have more highly phosphorylated ERK1/2 compared with other breast cancer cell lines and exhibit very low basal Akt activity [20, 21].

Presently, MDA-MB-231 have no clinically established targeted therapies [22]. Therefore, understanding their molecular basis is important for successful development of new drugs and as a result several studies on potentially active agents for this specific form of breast cancer were performed using the MDA-MB-231 cell line.

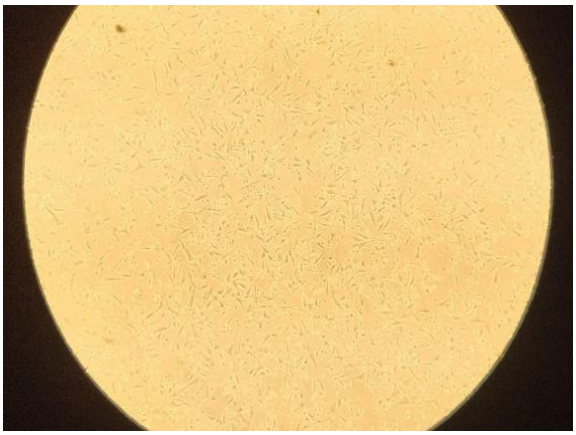


Figure III. 2. MDA-MB-231 cell line
(personal photo.)

III.3.4. Cytotoxicity essay (MTT assay)

The use of a tetrazolium dye (MTT) as an indirect measure of cell number was first reported in the early 1980s. The American National Cancer Institute (NCI) evaluated MTT dye reduction as a possible endpoint in a rapid screening assay, and this stimulated interest in the wider scientific community.

MTT assay is fast, simple, and allows to perform a large number of assays in one batch. This is an important factor when comparing cell lines, cytotoxic agents, or when evaluating combinations of drugs.

III.3.4.1. MTT assay basics

It is based on the ability of mitochondrial succinate dehydrogenase to precipitating the tetrazolium salt MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl bromide) tetrazolium) crystals. Intracellular violet formazan crystals are dissolved in DMSO and revealed by OD reading. The intensity of this coloration is proportional to the number of living cells present during the test but also to their metabolic activity. A standard range of number of known cells is carried out in parallel with each experiment [6].

III.3.4.2. The MTT Reagent

MTT is a yellow water-soluble tetrazolium dye that is reduced by live cells to a purple formazan product that is insoluble in aqueous solutions.

The MTT reagent can be prepared by adding 5mg of MTT in 1ml of phosphate buffered saline, it is stable at -20°C in the dark for two months, provided there is no contamination. However, it is advised to prepare only the amount needed for the current experiment. Solution must be filter sterilized after adding MTT [6].

III.3.4.3. Equipment

The FLUOstar Omega is a multi-mode microplate reader with six detection modes. It utilizes an ultra-fast UV/vis spectrometer or filters for absorbance as well as highly sensitive filters for all other detection modes.

III.3.4.3. MTT protocol

The experiment was carried out at the department of biochemistry at the faculty of pharmacy, Hacettepe University, Ankara, Turkey.

The cancer cell lines MDA-MB-231 (ATCC® HTB-26) and MCF-7 (ATCC® HTB-22) were cultured in Dulbecco's Modified Eagle's Medium (DMEM) including 1% L-glutamine, 1% pen-strep solution and 10% of fetal bovine serum (FBS), at 37°C in a humidified atmosphere containing 5% CO_2 for 48 hours. Then, the cells were seeded in 96-well plate at 4000 cells per well in 100 µl of DMEM, after 6 hours, the medium was changed and DMEM with 1% L-glutamine, 1% per-strep solution and 2% of FBS was added. After 8 hours, the tested compounds were added at different concentrations and incubated for 48 hours at 37°C in the presence of 5% CO_2. After incubation, 10µl MTT reagent (5mg/ml) was added to each well and well plates were incubated for 2 hours at 37°c in a stirrer. The MTT formazan crystals were dissolved by adding 100µL DMSO. Crizotinib (Sigma, PZ0191, USA), an approved anti-cancer drug, was also added as a positive control[23].

Relative cell viability was estimated by comparing absorbance value at 570 nm (for absorbance of MTT formazan) and 690 nm (for the reference wavelength) of treated group and untreated control group using FLUOstar Omega microplate reader. IC_{50} values were calculated using GraphPad Prism software.

III.3.4.4. Applications of MTT assay

- Analysis of cytotoxic and cytostatic compounds such as anticancer drugs and other pharmaceuticals.
- Measurement of cell proliferation in response to growth factors, cytokines, mitogens and nutrients.
- Assessment of growth inhibitory antibodies and physiological mediators.

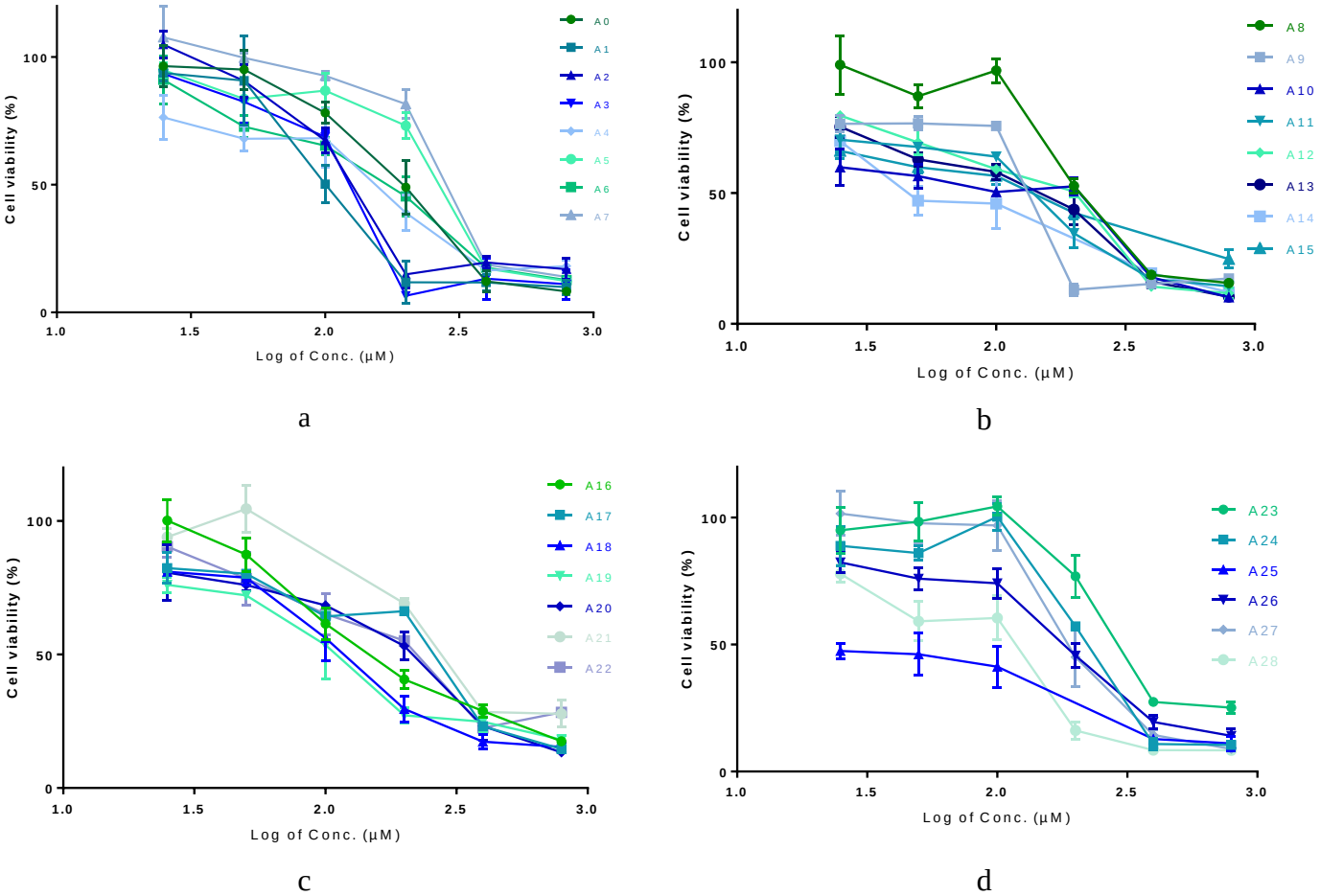
III.3.5. Results and discussion

The results were obtained after 2 days of culture, and were expressed as percentages of cell vs. control. Control experiments were conducted in the same conditions, but without any additional chemical compound.

For first general screening, the cell viability was evaluated at 4 concentrations, 200, 100, 50 and 25 µM for both cell lines.

The cell viability data for MDA-MB-231 cell line revealed that the cytotoxicity of all the compounds increases with higher concentration where most of them were inactive at 25µM. In particular compounds A1, A2, A3, A9, A25 and A28 are more potent at 200µM than the rest of the compounds. For this reason, a broader range of concentrations (800, 400, 200, 100, 50, 25 µM) was chosen to assess the activity of these compounds in more detail. The results corresponding to the percentage cell viability are presented in graph III.1.

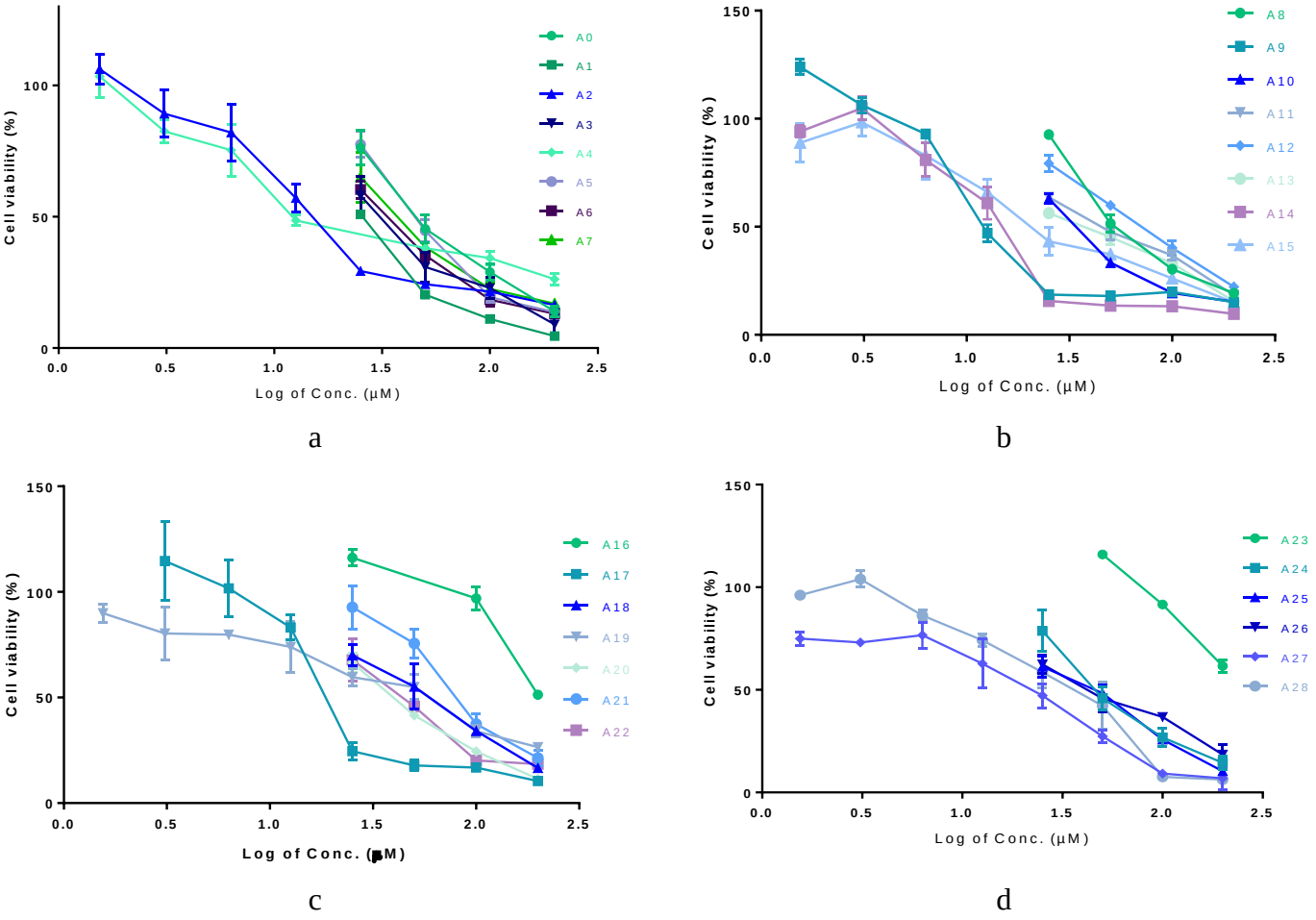
It can be seen from the graphs that the cytotoxicity of all the compounds against MDA-MB-231 is directly related to the concentration, and all compounds were potent at 800µM. In particular, A25 is the most potent one with cell viability ratio less than 50% with the lowest concentration 25µM and around 11% in the concentration 800µM (Graph III.1.d).



Graph III. 1. MDA-MB-231 cell viability ratio versus logarithm of the concentration obtained for compounds a: A0—A7; b: A8—A15; c: A16—A22; d: A22—A28.

Unlike MDA-MB-231 cell line, the cell viability results for MCF-7 showed that the cytotoxicity of the studied compounds is high at the lowest concentration 25µM, where the most potent ones are A2, A4, A9, A14, A15, A17, A19, A27 and A28 with cell viability ration less than 50%. Thus, a wider range of concentrations (200, 100, 50, 25, 12.5, 6.25, 3.125, 1.562

µM) was selected to further analyse the activity of these compounds in more detail. The results of the cell viability ratio are presented in graph. III.2. It can be noticed that A27 is the most potent one which inhibit almost 70% of the cells in the lowest concentration equal to 1.562µM.



Graph III. 2. MCF-7 cell viability ratio versus logarithm of the concentration obtained for compounds a: A0—A7; b: A8—A15; c: A16—A22; d: A22—A28.

The following figures show examples of 96 well plate after formazan crystals were dissolved. Control experiment were conducted without any additional chemical compound which is clear in figure III.3.a with most intensive color. A14 appears with a clear purple with all concentrations with MCF7 cells (figure III.3.b).

Figure III.4.a shows an intensive purple with all concentrations with MDA-MB-231 cells, however, they were clear with the high concentrations (800µM and 400µM) figure III.4.b.

The IC50 values along with SEM (standard error of the mean) and R^2 (squared correlation coefficient) are presented in table III.1.

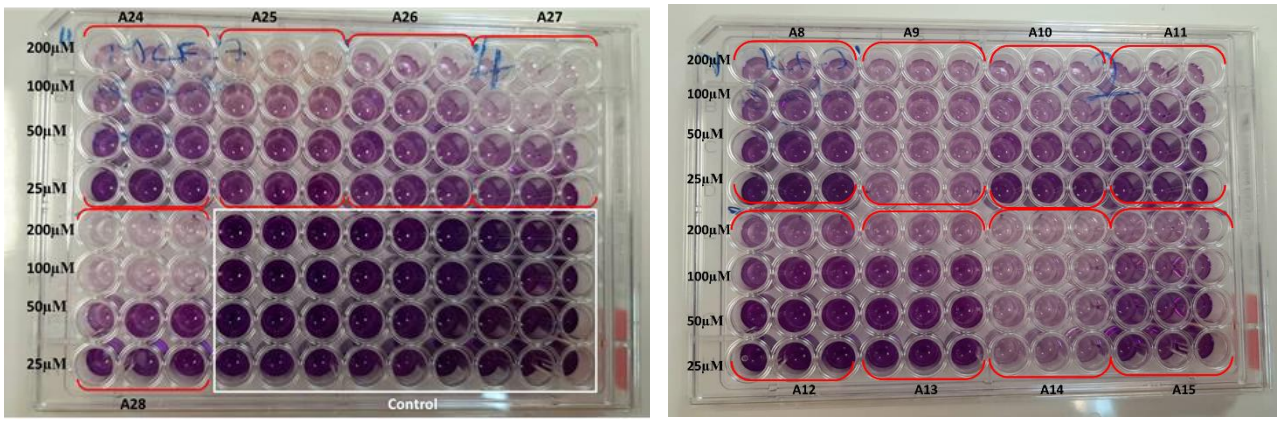


Figure III. 3. The 96-well micro plate used in MTT assay for MCF-7 cell line with different concentration of some of the compounds.

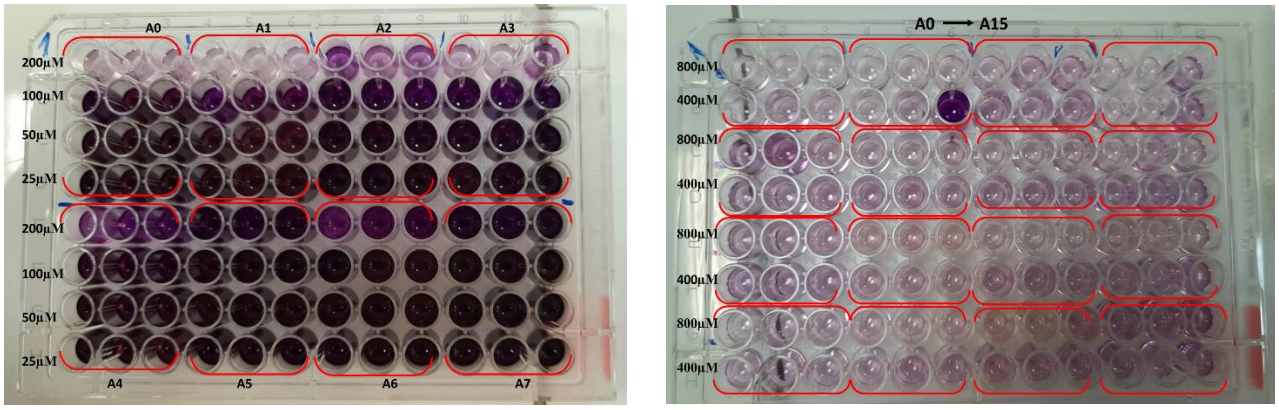


Figure III. 4. The 96-well micro plate used in MTT assay for MDA-MB-231 cell line with different concentration of some of the compounds.

As the data shows in Table III.1, all compounds have shown a promising cytotoxicity against MCF-7, the most potent ones are A2, A9, A14, A17 and A27 with IC_{50} equal to 17.9, 14.14, 14.06, 19.34 and 15.9 µM, respectively. where, the only inactive compounds are A16 and A23 with IC_{50} equal to 201.6 and 227.2 µM, respectively.

On the other hand, it can be seen from table III.1 that most of studied compounds were inactive against MDA-MB-231 with IC50 in the range of 101.7 - 318.4 µM except for A10, A14, A25 and A28 that have moderate cytotoxicity with IC_{50} equal to 83.24, 60.93, 31.4 and 79.56 µM, respectively, where A25 was the best one. Similar result has been found previously, where ferrocenic derivative (1,2-bis(4′-hydroxyphenyl)-2-ferrocenyl-but-1-ene) has shown a very weak cytotoxicity against MDA-MB-231. However, 1,1-bis(4′-hydroxyphenyl)-2-ferrocenyl-but-1-ene has indicated strong effect against those cells [24]. In addition, A25 showed cytotoxicity against both MDA-MB-231 and MCF-7, which may be explained by the existence of ERβ in both MDA-MB-231 and MCF-7 that could mediate the cytotoxicity[24, 25].

It can be seen from the results that none of the compounds was more potent than A14 at both cancer cell lines which also shows a good affinity to DNA with binding energy ΔG equal to -25,147 KJ/mol [4]. Regarding the substituted CN group at phenyl ring, we can say that the cyano group in meta position is the best comparing with ortho and para positions where it shows moderate cytotoxicity. Furthermore, it can be seen that A27 has shown better result than A26 and A28 with MCF-7 cell line. We can say that the best position for the cyano group in the acylated and the non-acylated compounds is the meta position. The compounds bearing NO_2 substituent and their acylated forms have also shown similar results, in contrast the ortho NO_2 substituent were totally inactive (A16 and A23). Moreover, the compounds with NO_2 in the para position show moderate cytotoxicity.

Based on the obtained results, it can be assumed that all potent compound with MCF-7 cell line (A9, A14, A17 and A27) are substituted in meta position with one of the following Cl, NO_2 or CN group.

Furthermore, A9, A17 and A27 are expected to be selective against MCF-7 since they were totally inactive against MDA-MB-231 with selectivity equal to 8.4, 10.4 and 12.5, respectively.

Although most of the studied compounds showed remarkable cytotoxicity, but Crizotinib which is used as a standard anti-cancer drug in this study showed stronger cytotoxicity against both cell lines with IC50±SEM values equal to 3.5±1.5 µM for MDA-MB-231 and 3.5±1.5 µM for MCF7.

Table III. 1. IC50 values of compounds A0-A28 of cell proliferation assay against MCF-7 and MDA-MB-231.

Compounds	MCF7			MDA-MB-231		
	IC50(µM)	R^2	SEM	IC50(µM)	R^2	SEM
A0	50.34	0.9553	0.0248	187.3	0.9716	0.0227
A1	24.97	0.9832	0.0143	101.7	0.9387	0.0301
A2	17.9	0.9241	0.0458	126.5	0.917	0.037
A3	30.6	0.9123	0.0386	116.3	0.913	0.0333
A4	27.88	0.872	0.0792	130.5	0.8667	0.0559
A5	46.97	0.9664	0.0206	257.5	0.9292	0.0311
A6	33.24	0.9718	0.0200	144.3	0.9617	0.0334
A7	38.23	0.926	0.0328	284.6	0.957	0.0232
A8	62.96	0.928	0.0327	225.3	0.939	0.0309
A9	14.14	0.8942	0.0533	118.2	0.807	0.0625
A10	34.26	0.9458	0.0276	83.24	0.752	0.0950
A11	47.06	0.9674	0.0216	107.4	0.924	0.0417
A12	71.76	0.9875	0.0128	131.5	0.922	0.0399
A13	37.08	0.9486	0.0299	109.9	0.927	0.0386
A14	14.06	0.9491	0.0351	60.93	0.901	0.0556
A15	26.96	0.9376	0.0410	114.4	0.944	0.0357

A16	201.6	0.8981	0.0223	172.9	0.950	0.0326
A17	19.34	0.9014	0.0452	201.5	0.885	0.0529
A18	55.83	0.9397	0.0290	116.9	0.955	0.0293
A19	47.07	0.9231	0.0506	106.9	0.919	0.0401
A20	40.83	0.9878	0.0128	168	0.927	0.0411
A21	85.32	0.946	0.0277	302.8	0.918	0.0421
A22	43.14	0.92	0.0371	194.1	0.918	0.0475
A23	227.2	0.8054	0.0493	318.4	0.909	0.0350
A24	51	0.9354	0.0299	220.5	0.938	0.0258
A25	41.11	0.959	0.0245	31.4	0.833	0.0883
A26	44.61	0.919	0.0355	166.7	0.935	0.0362
A27	15.9	0.9084	0.0566	198.7	0.958	0.0243
A28	31.32	0.9616	0.0316	79.56	0.916	0.0476
Crizotinib	3.6	-	0.6	3.5	-	1.5

SEM: standard error of the mean; R^2: squared correlation coefficient.

III.4. In silico evaluation of the anti-breast cancer activity

To predict and explain the mechanism of anticancer activity and further visualized the interactions of the studied compounds with MCF-7, molecular docking study was performed, since MDA-MB-231 lacks the estrogen (ERα) and progesterone receptors (PR), these receptors were chosen for this part.

Moreover, the compounds were docked into the human placental aromatase cytochrome P450 and 17-beta-hydroxysteroid dehydrogenase type 1 (17β-HSD1) enzymes.

The docking study was performed for the four targets with the 29 studied compounds using AutoDock 4.2 software with AutoDock Tools (ADT) interface.

III.4.1. The Targets

III.4.1.1. Estrogen Receptor α (ERα)

It is one of two major estrogen receptor types (ERα and β), a nuclear receptor activated by the sex hormone estrogen. ERα regulates essential physiological functions of different. it mainly occurs in mammary gland, uterus, ovary (thecal cells), bone, male reproductive organs (testes and epididymis), prostate (stroma), liver, and adipose tissue.

In several breast cancers, ERα activation by estrogens is commonly considered responsible for enhanced proliferation. Therefore, theoretically, breast cancer patients with estrogen responsive disease should respond positively to treatment with ERα-antagonists [26].

III.4.1.2. Progesterone Receptors (PR)

Progesterone effects are mediated through the binding to the nuclear progesterone receptor (PR). PR is a member of a large family of ligand-activated nuclear transcription regulators, that are characterised by organisation into specific functional domains and are maintained, to differing degrees, between species and family members. The general trend for PR coactivators is to be overexpressed and amplified in cancers, suggesting that increased availability of these proteins may enhance PR activity in tissues such as the breast, where it is known to play a role in cancer development [27].

III.4.1.3. The human placental aromatase cytochrome P450

Aromatase, a cytochrome P450 enzyme complex present in breast tissues, plays a significant role in the biosynthesis of estrogen from androgen. The aromatase converts the ring situated on the left (the ring labelled A) of the steroid into an aromatic ring (hence the name aromatase) through oxidation and the loss of a methyl group (figure III.5) [28].

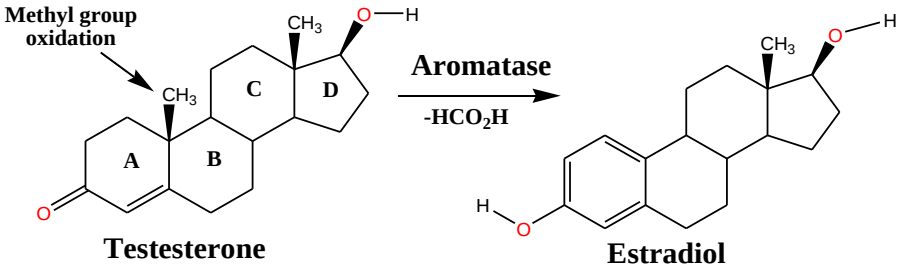


Figure III. 5 Aromatase converting testosterone to estradiol.

Consequently, the inhibition of the aromatase is a crucial strategy to prevent the growth stimulation effect of estrogens in breast cancer tissues in postmenopausal women. Aromatase inhibitors (AIs) have been developed and used as anti-breast cancer agents [29, 30].

III.4.1.4. 17β-hydroxysteroid dehydrogenase type 1 (17β-HSD1)

The enzyme 17β-HSD1 is considered to catalyse the final step in the hormone Oestradiol synthesis. It catalyses the NAD(P)H dependent transformation of estrone (E1) to 17β-estradiol (E2) which represents the most potent estrogen in humans. This hormone stimulates breast cancer and other oestrogen-related diseases [31, 32].

Therefore, the selective inhibition of this enzyme is considered as a valuable treatment option. This intracrine concept has been supported by the observation that an 17β-HSD1 inhibitor tool compound can effectively block the increased E2 synthesis in endometriotic tissue specimens[33].

Because of the tissue-selective expression and the intracrine mode of action of 17b-HSD1 [34], its inhibition should be associated with less side effects compared to established therapies

using GnRH- analogues, aromatase inhibitors, anti-estrogens or selective estrogen receptor modulators (SERMs) [35-37].

III.4.2. Docking using AUTODOCK

III.4.2.1. The theory of AUTODOCK

A semi-empirical free energy force field is used while the docking simulation procedure. the force field evaluates conformation in 2 steps, in the first step is the estimation of the intramolecular energy of transformation from the unbounded to the bounded stats of ligand-target conformation, the second is the evaluation of the intramolecular energy of combining the ligand and the receptor in the bound state[38, 39].

$$\Delta G = \Delta G_{vdw} + \Delta G_{Hbond} + \Delta G_{elec} + \Delta G_{tor} + \Delta G_{sol} \qquad (Eq \text{ III. } 1)$$

Where:

ΔG_{vdw} : is the van der waal's energy

ΔG_{Hbond} : is the hydrogen bonding energy

ΔG_{elec} : is the electrostatics interaction energy

ΔG_{tor} : is a term which translates the increase of energy of the system due to the restriction of the free rotors of the ligand and the restriction of the rotations and translations of the ligand during complexation with the receptor.

ΔG_{sol} : is another term related to entropy that describes the energy changes of the system during the desolvation of the ligand at the time of complexation at the receptor.

IV.4.2.2. AUTODOCK steps

The ligand / protein interaction is evaluated by the following steps [40, 41]:

- **Preparing Coordinates:** The first step is to prepare the ligand and receptor coordinate files to include the information needed by AutoGrid and AutoDock. after downloading the PDB file of the target as mentioned above, all the hydrogen atoms of the ligand are present, whereas only the essential (polar) hydrogens of the target are kept. Partial charges are assigned by the appropriate method (Gasteiger for the ligand and KOLL UNI for the target). The target.pdb file is converted to ".pdbqt" format in which the partial charges, polar hydrogen atoms and parameters of the target atoms are found. For the ligand.pdb is also converted to ".pdbqt" file.
- **Generation of a box containing the grid (Running AutoGrid):** AutoDock needs pre-calculated *grid maps*, one for each atom type present in the ligand being docked. This makes the docking calculations quicker. These maps are calculated by **AutoGrid**. A

grid map consists of a 3D lattice of regularly spaced points, surrounding (either entirely or partly) and centered on some region of interest of the macromolecule under study (catalytic site). Typical grid point spacing varies from 0.2Å to 1.0Å, and the default is 0.375Å (almost a quarter of the length of a carbon-carbon single bond). Each point within the grid map stores the potential energy of a '*probe*' atom or functional group that is due to all the atoms in the macromolecule. AutoGrid requires a grid parameter file which usually has the extension ".GPF".

- **Running AUTODOCK:** AutoDock uses one of many conformational search algorithms to explore the conformational states of a flexible ligand, using the maps produced via AutoGrid to estimate the ligand-protein interaction at each point in the docking process. In a typical docking, the user will dock a ligand many times, to get numerous docked conformations. AutoDock requires the grid maps that is calculated by AutoGrid, a ".pdbqt" file for the ligand and a docking parameter file that specifies the files and parameters for the docking calculation which has the extension ".dpf".
- **Result analysis:** At the end of the docking process, a dock.dlg file is generated. It contains the final docked coordinates regrouping the solutions proposed (confirmations) by similarity and the classes by increasing energies. The similarity among two solutions is calculated using the Root Mean Square Deviation (RMSD). these docked conformations might be visualized using AutoDockTools, they might be written as ".pdbqt" files using AutoDockTools, or they might be read directly from the dock.dlg file using a text editor.

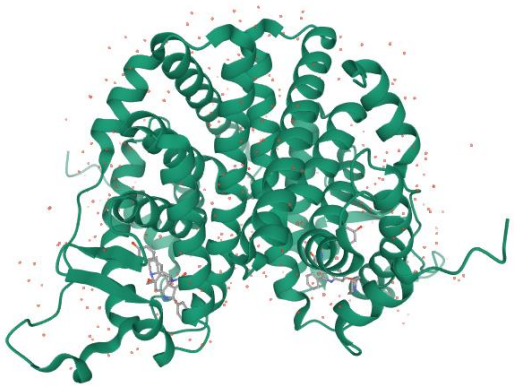
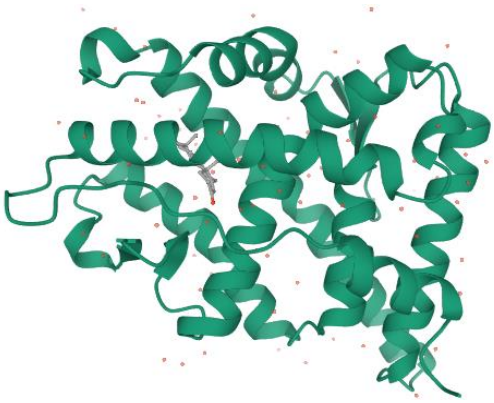
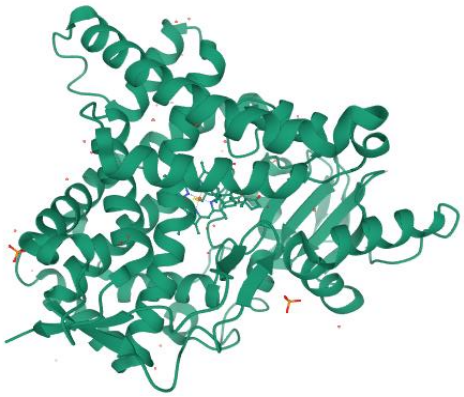
  Theoretically, the best solution is the one with the lowest free binding energy which should be negative ($\Delta G < 0$).

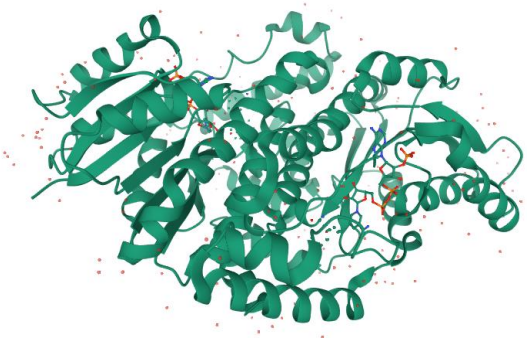
III.4.3. Molecular docking

The docking study was performed using AutoDock 4.2 software with AutoDock Tools (ADT) interface [42]. Table III.2 shows the targets' information chosen for docking studies along with their PDB codes. After downloading the PDB files from protein data bank (www.rcsb.org), all receptors were first prepared using Discovery Studio Software [43], all water molecules, ligands and cofactors were removed to obtain the raw protein's structure and the active site was defined. The polar hydrogen atoms were then added along with the charges using ADT. The ligands were fully optimized using DFT/B3LYP method with 6-311G++(d, p) basis set using Gaussian 09W software (chapter II) [44]. Then, the structures were saved as ".pdbqt" file using ADT.

The docking procedure consists of two main steps; the first one is the generation of interactions maps of the receptor depends on the active site using the AutoGrid tool, the grid boxes size are tabulated in table III.3. The second step is docking the ligand into the receptor with AutoDock4. Lamarckian genetic algorithm was utilized for docking search with number of evaluations: 2,500,000, population size: 150, GA run: 15. The graphical interface of ADT was used to dock and view the results. Finally, different values of binding energy were obtained, the best poses were selected with the lower docking energy [45-47] and were used in the docking analysis using Protein-Ligand Interaction Profiler (PLIP) [48] .

Table III. 2. Target receptors' information chosen for docking studies.

	**Estrogen receptors (ER) and**	
	PDB ID	**2IOG**
	Resolution (Å)	1.60
	R-Value Free	0.328
	R-Value	0.212
	Space Group	P 6_5 2 2
	Chains	A
	Organism	Homo sapiens
	**progesterone receptors (PR)**	
	PDB ID	**1A28**
	Resolution (Å)	1.80
	R-Value Free	0.227
	R-Value	0.191
	Space Group	P 1 2_1 1
	Chains	A, B
	Organism	Homo sapiens
	**Aromatase**	
	PDB ID	3S79
	Resolution (Å)	2.75
	R-Value Free	0.236
	R-Value	0.220
	Space Group	P 3_2 2 1
	Chains	A
	Organism	Homo sapiens



17β-HSD1	
PDB ID	3HB5
Resolution (Å)	2.00
R-Value Free	0.248
R-Value	0.202
Space Group	C 1 2 1
Chains	X
Organism	Homo sapiens

Table III. 3. Grid parameters used to generate the AutoGrid

	2IOG	1A28	3S79	3HB5
Size	50 x 50 x 50	50 x 50 x 50	70 x 70 x7 0	50 x 50 x 50
Spacing	0.375	0.375	0.375	0.375
Coordinates	X = 24.685 Y = 5.51 Z = 23.906	X= 26.573 Y= 8.374 Z= 64.687	X=83.347 Y= 50.057 Z= 46.452	X= 12.472 Y= -1.756 Z= -3.474

III.4.4. Results and discussion

III.4.4.1. Estrogen and Progesterone receptors

The molecular docking study was carried out using estrogen receptor (ERα) and progesterone receptor (PR) as targets to clarify the efficacy of studied ferrocenylmethylanilines derivatives against hormone-dependent (MCF7) and to predict binding mode of these interactions. Table III.4. represents the binding free energies (ΔG) along with the binding constants (K) of the studied compounds interacted with ERα and PR.

Table III. 4. binding free energies (ΔG) and binding constant (K), for the studied molecules with ERα and PR.

Compounds	ERα		PR	
	ΔG (Kcal.mol^{-1})	K (M^{-1})	ΔG (Kcal.mol^{-1})	K (M^{-1})
A0	-7.8	5.14 ×10^5	-7.66	4.06 ×10^5
A1	-7.99	7.08 ×10^5	-8.95	35.732 ×10^5
A2	-8.12	8.82 ×10^5	-8.93	34.547 ×10^5
A3	-9.85	162.939 ×10^5	-10.62	596.771 ×10^5
A4	-9.02	40.207 ×10^5	-10.13	261.238 ×10^5
A5	-8.09	8.38 ×10^5	-8.63	20.833 ×10^5
A6	-7.88	5.88 ×10^5	-8.83	29.187 ×10^5
A7	-7.95	6.62 ×10^5	-8.51	17.017 ×10^5
A8	-8.24	10.794 ×10^5	-8.86	30.701 ×10^5
A9	-8.01	7.32 ×10^5	-8.97	36.957 ×10^5
A10	-7.65	3.99 ×10^5	-8.30	11.943 ×10^5
A11	-7.90	6.09 ×10^5	-9.09	45.244 ×10^5
A12	-7.76	4.81 ×10^5	-8.59	19.474 ×10^5

A13	-7.53	3.26×10^5	-8.14	9.12×10^5
A14	-8.23	10.614×10^5	-8.75	25.504×10^5
A15	-7.88	5.88×10^5	-8.90	32.843×10^5
A16	-6.77	0.9×10^5	-8.68	22.665×10^5
A17	-7.48	3×10^5	-9.82	154.902×10^5
A18	-7.45	2.85×10^5	-9.86	165.709×10^5
A19	-7.92	6.29×10^5	-8.35	12.994×10^5
A20	-7.88	5.88×10^5	-9.51	91.85×10^5
A21	-7.53	3.26×10^5	-8.57	18.829×10^5
A22	-9.03	40.891×10^5	-10.16	274.791×10^5
A23	-7.14	1.69×10^5	-9.20	54.463×10^5
A24	-7.82	5.32×10^5	-10.39	404.953×10^5
A25	-7.97	6.85×10^5	-10.15	270.197×10^5
A26	-7.92	6.29×10^5	-9.30	64.464×10^5
A27	-9.05	42.293×10^5	-9.83	157.536×10^5
A28	-8.74	25.078×10^5	-9.37	72.539×10^5

In summary, the obtained results have shown high affinity between all molecules and both targets ERα and PR with binding free energy in the range from -6.77 to -10.62 Kcal.mol^{-1} (table. III. 4).

In the binding pattern of the compounds and ERα, it can be seen that A3 has shown the lowest binding free energy equal to -9.85 Kcal.mol^{-1}. In addition, A4, A22 and A27 have also shown low binding free energies, equal to -9.02, -9.03 and -9.05 Kcal.mol^{-1} , respectively. A16 has shown the highest binding free energy equal to -6.77 Kcal.mol^{-1}, which was inactive against MCF-7 cell lines with IC_{50} value equal to 201.6 µM (table III.1).

Moreover, all compounds have shown at least 5 hydrophobic bonds with different amino acids, some of them made one H-bond or more with the amino acids Glu353, Arg394, Leu346, Ala350, Asn719, Pro325, Leu391, Arg766, Gln725, Leu763 and Leu718. Furthermore, another bond of the type π- π stacking was observed with the amino acids Phe404, Phe778 and Trp393 (table III.5 and figure III.6).

It can be seen that all compounds are located at the estrogen binding cavity(table.III. 5), which is formed by the following amino acids: Met 342 to Leu 354, Trp 383 to Arg394, Val418 to Leu 428, Met 517 to Met528, Leu 539 to His 547 and Leu 402 to Leu 410 [49].

It has been reported and proved in the literature [49] that estrogen binds to his receptor by several amino acids which are Ile 424, Gly 521, Leu 525, Phe404, Glu353, His524, Arg394, Ala 350 and Leu 387. It can be seen from table III. 6, that each of the studied compounds has shown similar bonding pattern to Estrogen's, for instance A9, A14, A17 and A27 have presented key hydrogen bonds with Glu353 and Arg394. Moreover, A2, A14 and A27 have made key hydrophobic interactions with Leu525 and Ile424. Therefore, these docking analyses

indicates the good inhibitory effect of the studied compounds against ERα. The compound A27 was chosen for secondary structure view to show its positioning into the receptor (figure III.7).

Table III. 5. Distances of formed bonds between the ligands and ERα residues.

	ERα													
Bonds type	**A0**		**A1**		**A2**		**A3**		**A4**		**A5**		**A6**	
	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist
Hydrophobic	PRO324	2.96	LEU349	3.87	LEU346	3.36	LEU346	3.20	LEU349	3.35	PRO324	3.04	LEU346	3.25
	ILE326	3.16	ALA350	3.43	LEU384	3.27	THR347	3.32	ALA350	3.34		3.20	LEU349	3.36
	LEU327	3.18	LEU384	3.24	LEU387	3.75	ALA350	2.95	LEU384	3.21	ILE326	3.17	ALA350	3.82
	LEU349	3.48		3.47	LEU391	3.59		3.20	LEU387	3.64	GLU353	3.76	LEU387	3.89
	GLU353	3.21	LEU387	3.14	ILE424	3.41	TRP383	3.33		3.18		3.54	LEU391	3.96
		3.77	MET388	3.95		3.73	LEU384	3.44	LEU391	3.90	HIS356	3.16		3.59
	MET357	3.72	LEU391	3.28	PHE425	3.35	LEU387	3.24	PHE404	3.26	MET357	3.92	LEU525	3.57
	ILE386	3.29	PHE404	3.24	HIS524	3.54		3.46		3.23	ILE386	3.61		3.21
	LEU387	3.39		3.52	LEU525	3.85	LEU391	3.13	ILE424	3.95	LEU387	3.84		
	LYS449	3.97	ILE424	3.09			PHE404	3.51		3.08	ARG394	3.18		
			PHE425	3.92				3.78	LEU525	3.26	PHE445	3.20		
							LEU525	3.22	MET528	3.74		3.72		
								3.31						
							LYS531	3.72						
Hydrogen Bonds											GLU353	1.92	LEU346	2.18
											ARG394	2.80	ALA350	3.68
π-Stacking					PHE404	5.05					TRP393	4.30	PHE404	5.04

Bonds type	A7		A8		A9		A10		A11		A12		A13	
	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist
Hydrophobic	PRO324	3.61	PRO324	3.23	PRO324	3.87	PRO324	3.21	LEU349	3.28	LEU346	3.46	LEU346	3.61
		3.01	ILE326	3.14	ILE326	3.49	ILE326	3.92	ALA350	3.74		3.56		3.47
	ILE326	3.88	MET357	3.42	GLU353	3.35		3.22	TRP383	3.31	LEU349	3.54		3.45
		3.46	ILE386	3.20	MET357	3.37	LEU327	3.68	LEU384	3.39	ALA350	3.02	LEU349	3.41
	LEU327	3.81	ARG394	3.21	LEU387	3.89		3.46	LEU391	3.88	LEU384	3.51	ALA350	3.38
		3.52	PHE445	3.18	ARG394	3.31	GLU353	3.14	LEU525	3.63	LEU387	3.68	LEU384	3.90
	GLU353	3.11		3.75	PHE445	3.14	ILE386	3.89		3.80	LEU391	3.71	PHE404	3.25
		3.70	LYS449	3.91		3.59	PHE445	3.28				3.91		3.64
	HIS356	3.41					LYS449	3.30				3.84	ILE424	3.10
	MET357	3.61									PHE404	3.69		3.98
	ILE386	3.12									LEU428	3.35	PHE425	3.82
	LEU387	2.97											LEU428	3.32
	LYS449	3.67												
Hydrogen Bonds			GLU353	2.11	GLU353	2.01	PRO325	2.15	LEU346	2.07			LEU391	3.59
			ARG394	3.11	ARG394	2.89			ALA350	3.46			ARG394	2.13
														3.45
π-Stacking			TRP393	4.35	TRP393	4.46			PHE404	4.97			PHE404	5.24

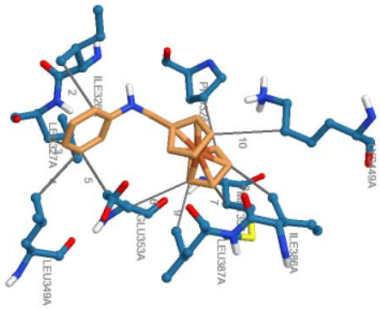
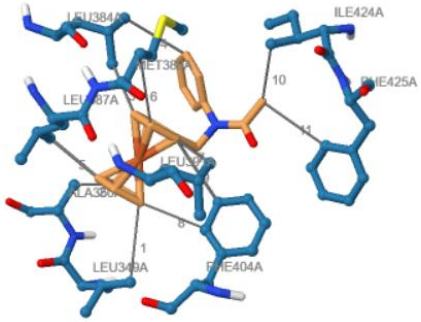
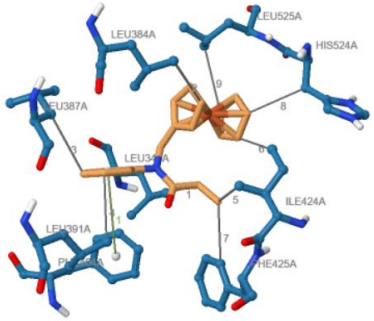
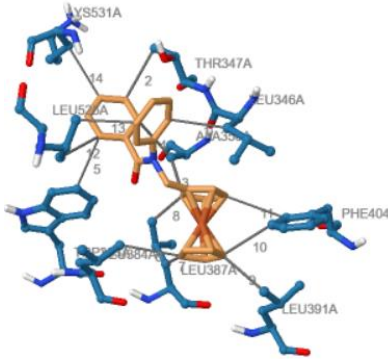
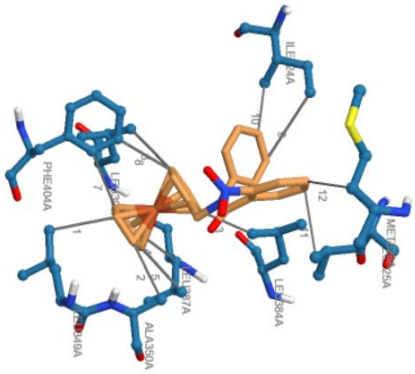
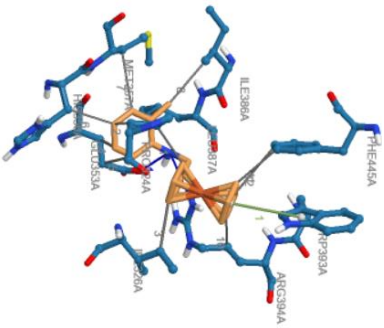
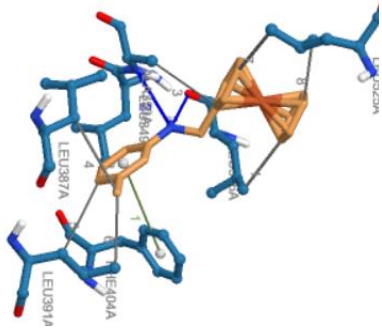
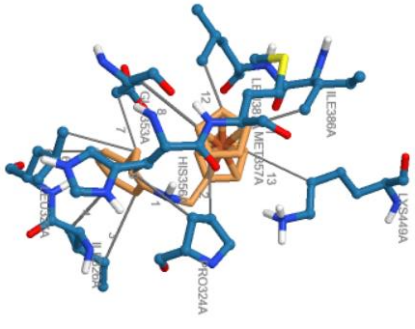
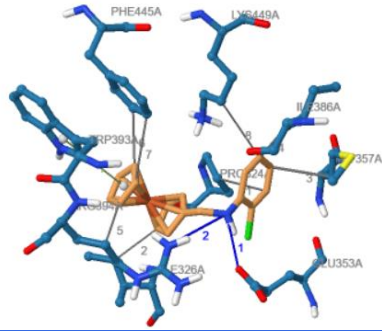
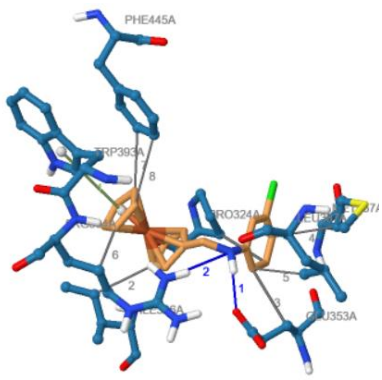
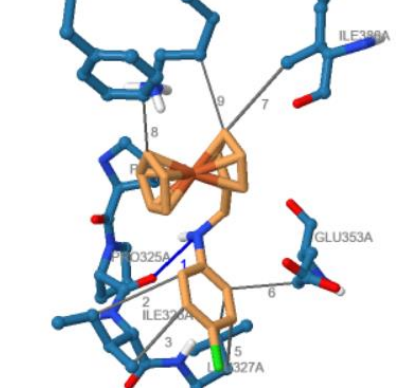
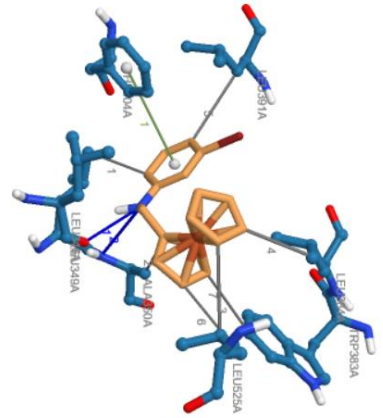
	A14		A15		A16		A17		A18		A19		A20	

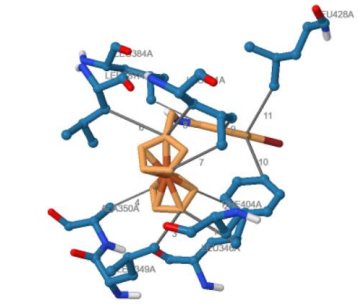
	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist
Hydrophobic	LEU346	3.94	LEU346	3.23	LEU346	3.90	LEU346	3.35	THR347	3.35	PRO324	2.96	LEU349	3.22
	THR347	3.51	LEU349	3.53		3.60	ALA350	3.58	ALA350	3.03		3.31	LEU384	3.25
	LEU349	3.34	LEU387	3.22	LEU349	3.54	LEU384	3.62	TRP383	3.63	ILE326	3.27	LEU391	3.25
	ALA350	3.18		3.92	ALA350	3.80	LEU387	3.78	LEU384	3.66	GLU353	3.72	PHE404	3.47
		3.20	PHE404	3.56	LEU384	3.66	LEU391	3.62	LEU387	3.41		3.95		3.38
	TRP383	3.17		3.38	ILE424	3.71				3.69	HIS356	3.34	ILE424	3.44
	LEU387	3.40	ILE424	3.31	LEU525	3.83			LEU525	3.50	ILE386	3.84	HIS524	3.51
	PHE404	3.47	LEU428	3.35						3.33	LEU387	3.60	LEU525	3.52
	LEU525	3.78									TRP393	3.92		3.64
		3.32									ARG394	3.41		
											PHE445	3.75		
Hydrogen Bonds	ARG394	1.90	LEU391	3.75	LEU391	3.69	ARG394	1.67	LEU391	3.38	GLU353	2.02	LEU391	3.60
		2.98	ARG394	1.77	ARG394	2.04	ARG394	2.97	ARG394	1.85	ARG394	2.09	ARG394	2.15
				2.74		3.45				2.94				3.45
π-Stacking					PHE404	5.02								

	A21		A22		A23		A24		A25		A26		A27	
	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist
Hydrophobic	LEU346	3.42	LEU346	3.82	LEU384	3.20	LEU346	3.88	LEU346	3.79	LEU346	3.39	THR347	3.52
	LEU349	3.69		3.80	LEU387	3.48		3.56	THR347	3.35	LEU349	3.78	LEU349	3.10
	ALA350	3.61		3.61	MET388	3.98	ALA350	3.67	ALA350	2.97	ALA350	3.60	ALA350	3.13
	LEU384	3.37	LEU384	3.38	LEU391	3.91	LEU384	3.91	TRP383	3.58	LEU384	3.37	TRP383	3.12
	LEU384	3.30	LEU387	3.11	PHE404	3.94		3.51		3.30	LEU387	3.75	LEU387	3.62
	LEU387	3.20	LEU391	2.96		3.19	ILE424	3.57	LEU387	3.23		3.08	PHE404	3.79
		3.70	PHE404	3.26	ILE424	3.50	HIS524	3.64	PHE404	3.79	LEU391	3.30	LEU525	3.20
	MET388	3.98	ILE424	3.95	LEU428	3.22	LEU525	3.84	LEU525	3.30		3.91		
	LEU391	3.67		3.37	HIS524	3.57					PHE404	3.45		
	PHE404	3.14	PHE425	3.29	LEU525	3.46					ILE424	3.06		
		3.75										3.23		
	PHE425	3.75									PHE425	3.64		
	LEU525	3.86										3.85		
											LEU428	3.28		
Hydrogen Bonds							LEU391	3.67	LEU391	3.74	ALA350	3.09	ARG394	1.98
							ARG394	2.24	ARG394	1.78				3.11
										3.06				
π-Stacking							PHE404	5.00			PHE404	5.05		

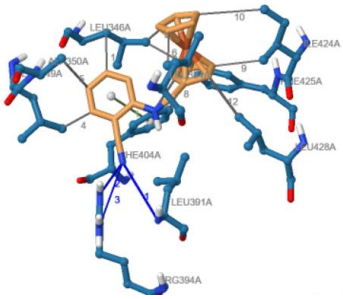
	A28	
	AA	Dist
Hydrophobic	LEU346	3.30
	LEU349	3.97
	ALA350	3.45
	LEU387	3.56

	ILE424	3.16	
	LEU428	3.61	
	LEU525	3.12	
Hydrogen Bonds	LEU391	3.75	
	ARG394	1.78	
		2.73	
π-Stacking	PHE404	5.13	

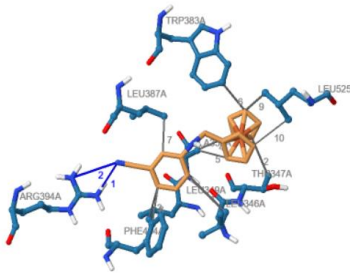
A0-ERα	A1-ERα	A2-ERα
		
A3-ERα	A4-ERα	A5-ERα
		
A6-ERα	A7-ERα	A8-ERα
		
A9-ERα	A10-ERα	A11-ERα
		
A12-ERα	A13-ERα	A14-ERα



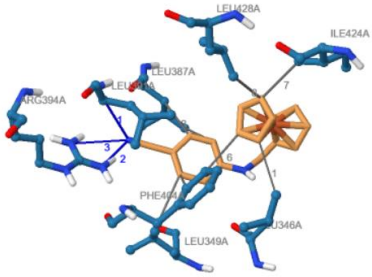
A15-ERα



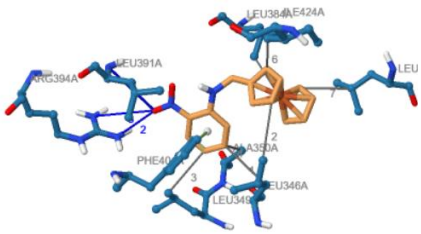
A16-ERα



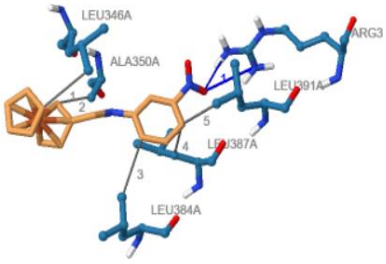
A17-ERα



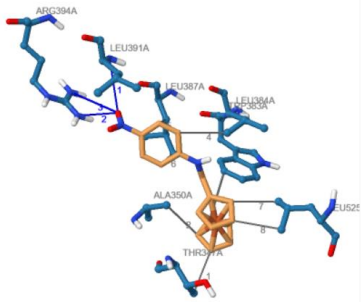
A18-ERα



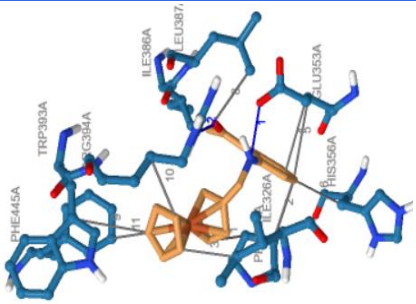
A19-ERα



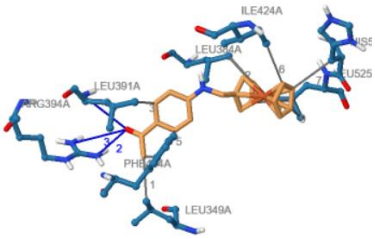
A20-ERα



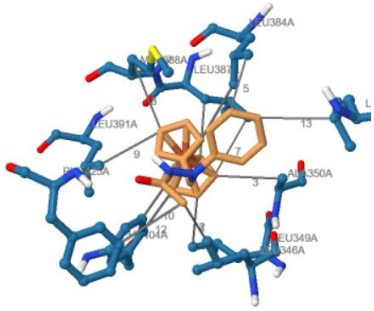
A21-ERα



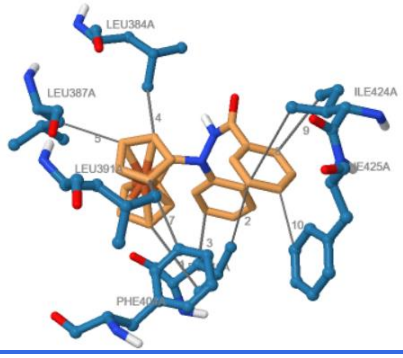
A22-ERα



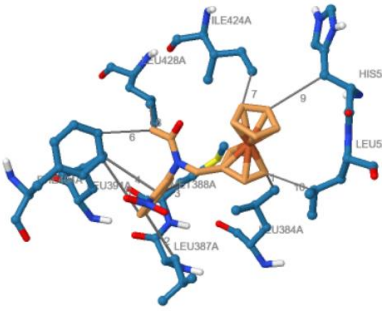
A23-ERα



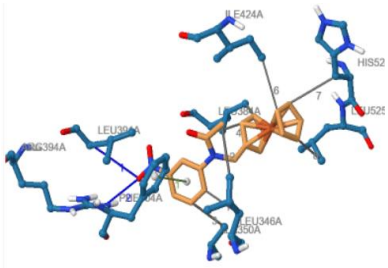
A24-ERα

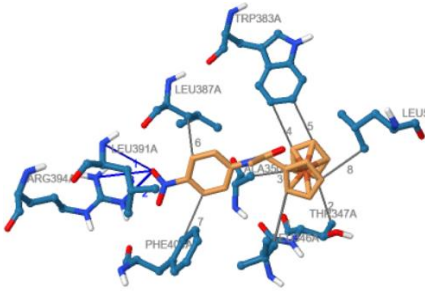


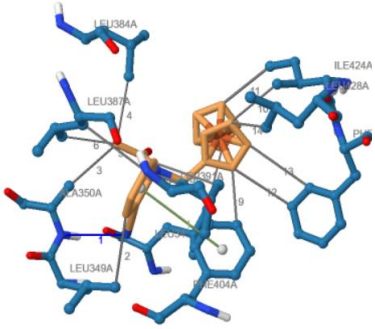
A25-ERα



A26-ERα







A27-ERα	A28-ERα

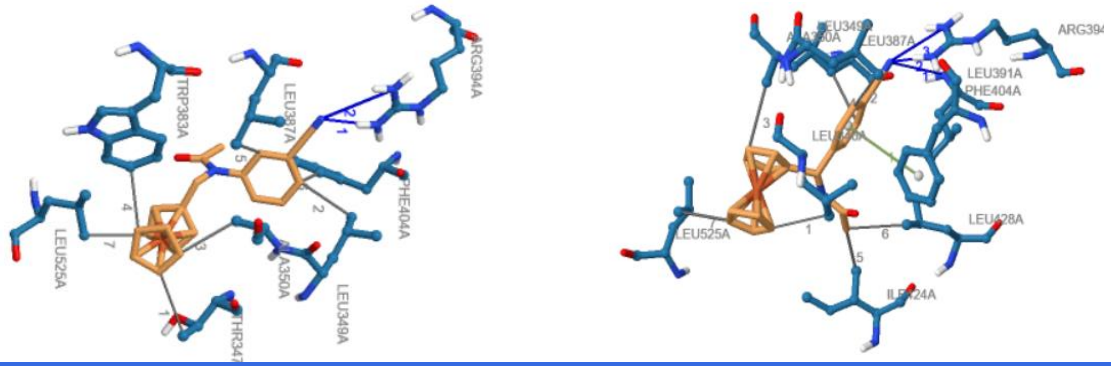


Figure III. 6. 3D interaction between studied ligands and the target ERα (PDB ID: 2IOG) where the grey lines show hydrophobic bonds and the bleu lines show the H-bonds.

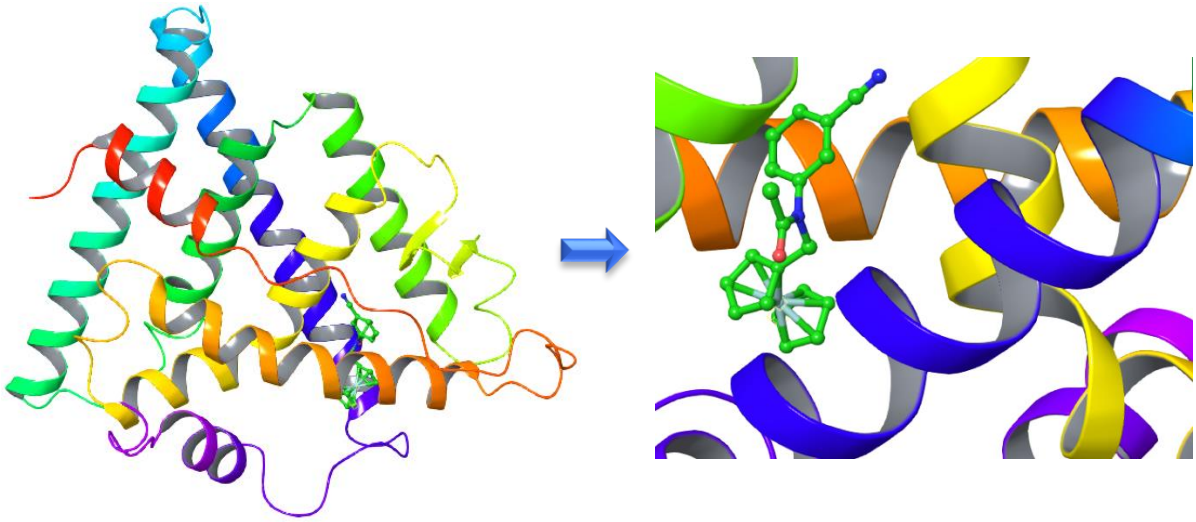


Figure III. 7. Docking pose of A27 with ERα illustrating (a) the secondary structure view (b) zoomed view of the interactions.

Previous studies found that hydroxytamoxifen made interaction between His524 and the ferrocenyl group, and between Asp351 and the nitrogen of its basic chain, moreover, it was found that these interactions provide correct positioning of the organometallic hormone and particularly of the dimethyl amino side chain which is known to be the source of the antihormonal effect. Looking at figure III.6 and table III.5 we can see that the compounds A2, A20 and A24 made similar hydrophobic bond between His524 and the ferrecenyl group of these compounds, which may be important for the antiproliferative activity of these compounds.

In addition, all tested compounds have indicated a good relationship to PR with binding free energies lower than the ER's in the range from -7.66 to -10.62 Kcal.mol^{-1}. Similar to ERα results, the best compound interacted with PR is A3 with the lowest binding free energy equal to -10.62 Kcal.mol^{-1}. Furthermore, A4, A22, A24 and A25 have also shown low binding free energies equal to -10.13, -10.16, -10.39 and -10.15 Kcal.mol^{-1}, respectively.

Previous studies have found that progesterone binds to its receptor via hydrogen bonds to Gln 725, Tyr 777, Asp 719, Thr 894 and Arg 766, and Van Der Waals bond with Phe778 [50, 51]. As it is shown in table III.6, A14, A15, A16, A17, A18, A20, A23, A24, A25, A27 and A28 have made similar hydrogen bonds with Gln725 and Arg766. Other bonds, π-stacking and π-alkyl, were made with Thr894 and Phe778, the amino acids which bind to progesterone, these interactions as well could disturb its binding pattern. These results indicate that these compounds have high binding affinity with PR (table III.6 and figure III.8).

Table III. 6. Distances of formed bonds between the ligands and PR residues

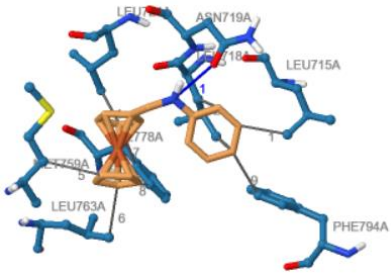
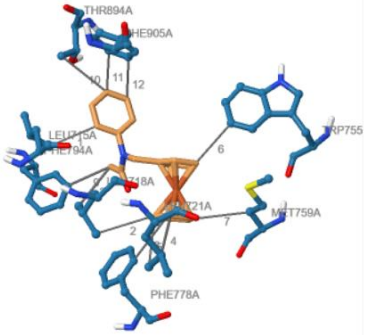
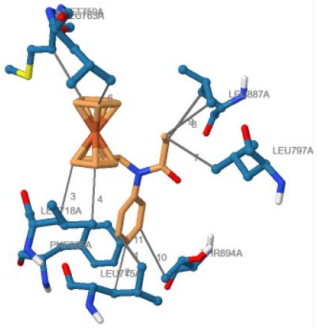
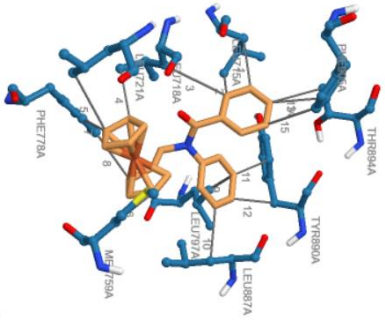
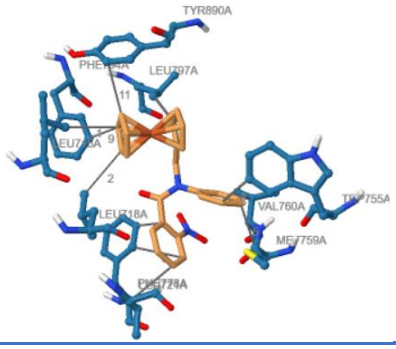
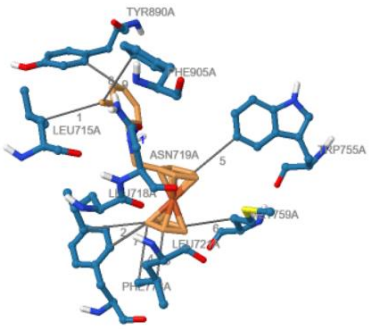
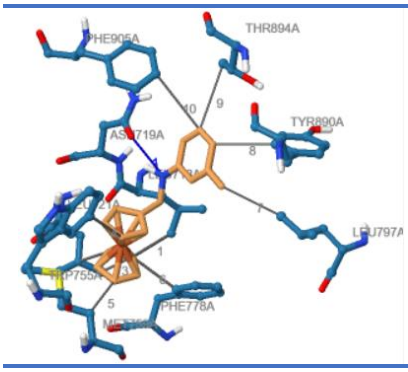
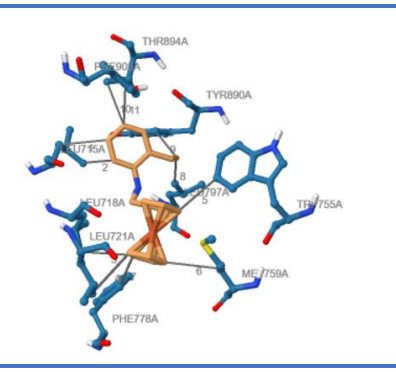
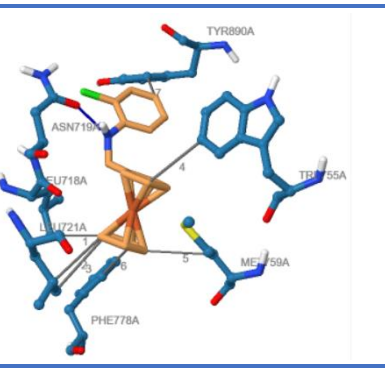
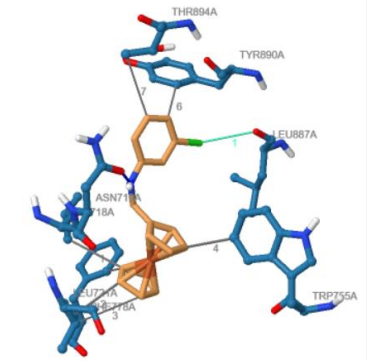
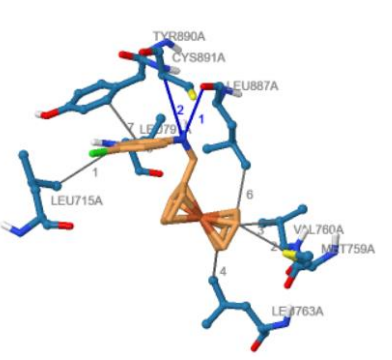
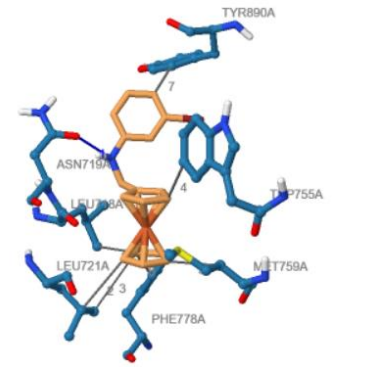
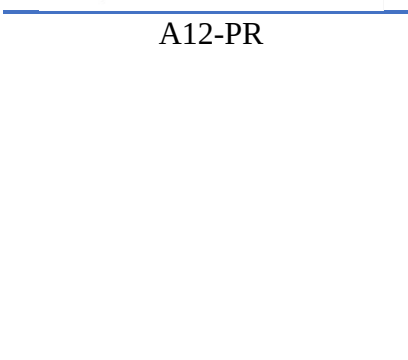
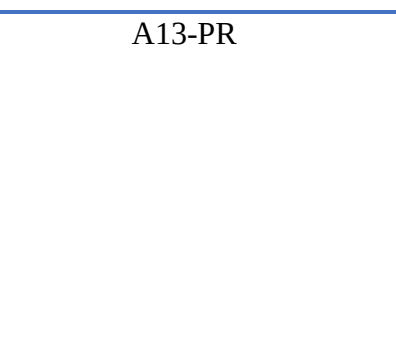
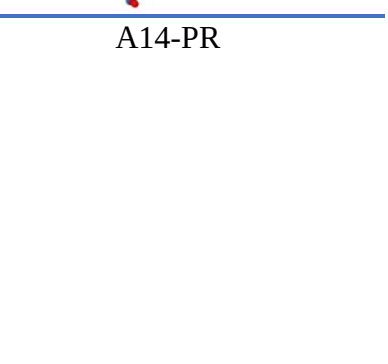
PR														
	A0		A1		A2		A3		A4		A5		A6	
	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist
Hydrophobic	LEU715	3.18	LEU715	3.56	LEU715	3.40	LEU715	3.42	LEU715	3.23	LEU715	3.92	LEU718	3.93
	LEU718	3.42	LEU718	3.81		3.93		3.56	LEU718	3.59	LEU718	3.93	LEU721	3.45
		3.89		3.76	LEU718	3.36	LEU718	3.90	LEU721	3.39	LEU721	3.47		3.86
	LEU721	3.63	LEU721	3.53		3.51		3.91	TRP755	3.27		3.94	TRP755	3.62
	MET759	3.30		3.95	MET759	3.47	LEU721	3.68	MET759	2.83	TRP755	3.54	MET759	3.66
	LEU763	3.80	TRP755	3.50	LEU763	3.80	MET759	3.59	VAL760	3.96	MET759	3.70	PHE778	3.54
	PHE778	3.61	MET759	3.95	LEU797	3.46	PHE778	3.23	PHE778	3.40	PHE778	3.59	LEU797	3.73
		3.31	PHE778	3.49	LEU887	3.27		3.81		3.95	TYR890	3.22	TYR890	3.41
	PHE794	3.88	PHE794	3.57		3.54	LEU797	2.97	PHE794	3.46	PHE905	3.98	THR894	3.88
			THR894	3.65	THR894	3.60	LEU887	3.22	LEU797	3.25			PHE905	3.92
			PHE905	3.41	PHE905	3.14	TYR890	3.24	TYR890	3.77				
				3.95				3.79						
							THR894	3.93						
							PHE905	3.32						
								3.42						
Hydrogen Bonds	ASN719	2.16									ASN719	1.99	ASN719	1.93

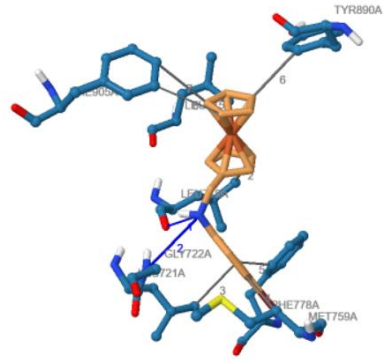
	A7		A8		A9		A10		A11		A12		A13	
	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist
Hydrophobic	LEU715	3.61	LEU718	3.89	LEU718	3.76	LEU715	3.92	LEU718	3.96	LEU715	3.61	LEU718	3.39
		3.60	LEU721	3.49	LEU721	3.42	MET759	3.80	LEU721	3.44	LEU718	3.77		3.45
	LEU718	3.90		3.93		3.97	VAL760	3.47		3.85	LEU721	3.90	MET759	3.44
	LEU721	3.51	TRP755	3.55	TRP755	3.78	LEU763	3.63	TRP755	3.64	MET759	3.41	PHE778	3.53
	TRP755	3.68	MET759	3.77	PHE778	3.35	LEU797	3.81	MET759	3.63	PHE778	3.38		3.77
	MET759	3.66	PHE778	3.54	TYR890	3.28	LEU887	3.65	PHE778	3.54	TYR890	3.42	LEU797	3.94
	PHE778	3.33	TYR890	3.18	THR894	3.81	TYR890	3.23	TYR890	3.32	PHE905	3.48	TYR890	3.39
	LEU797	3.60										3.74		
	TYR890	3.46												
	THR894	3.57												
	PHE905	3.31												
Hydrogen Bonds			ASN719	1.89	ASN719	1.92	LEU887	2.28	ASN719	1.85	LEU718	1.95	ASN719	2.38
							CYS891	3.72			GLY722	3.40		2.08

	A14		A15		A16		A17		A18		A19		A20	
	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist
Hydrophobic	LEU715	3.38	LEU715	3.57	LEU718	3.29	LEU715	3.37	LEU715	3.30	LEU715	3.25	LEU715	3.50
	VAL760	3.85	LEU718	3.43		3.36	LEU718	3.18	LEU718	3.21	LEU718	3.22	LEU718	3.10
	LEU763	3.88	PHE778	3.81	LEU721	3.41		3.48		3.43		3.50		3.31
	PHE778	3.67	LEU797	3.07		3.33	PHE794	3.90	LEU721	3.32	LEU721	3.43	LEU721	3.88
	LEU797	3.35	PHE905	3.88	PHE778	3.05	TYR890	3.79	TYR890	3.43	MET759	3.33	MET759	3.30
	TYR890	3.60			PHE794	3.33			PHE905	3.99	PHE778	3.62	LEU763	3.25
	PHE905	3.87			LEU887	3.88			PHE905	3.36	TYR890	3.83	TYR890	3.36
		3.83									PHE905	3.53	PHE905	3.73
Hydrogen Bonds	GLN725	2.18	ASN719	2.12	GLN725	1.78	LEU718	2.18	LEU718	1.82	LEU718	2.13	LEU718	2.26
	ARG766	2.08	GLN725	2.21	MET759	3.47	GLN725	1.88	GLN725	1.85			GLY722	3.42
			ARG766	3.13	LEU763	3.64	LEU763	3.70	LEU763	3.58			GLN725	1.72
					ARG766	1.97	ARG766	1.75	ARG766	1.89			ARG766	2.00
π-Stacking							PHE778	5.19	PHE778	5.34				

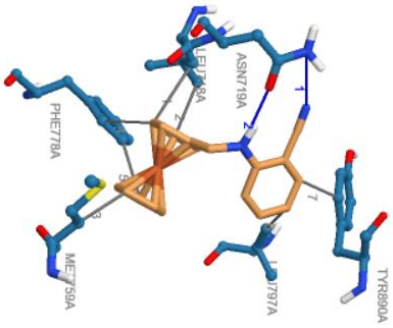
	A21		A22		A23		A24		A25		A26		A27	
	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist
Hydrophobic	LEU718	3.78	LEU715	3.44	LEU718	3.19	LEU715	3.61	LEU718	3.79	LEU715	3.63	LEU715	3.37
		3.21	LEU718	3.75	LEU721	3.59	LEU718	3.20	LEU721	3.70		3.75	LEU718	3.18
	LEU721	3.76		3.41		3.85	LEU721	3.59	TRP755	3.88	LEU718	3.97		3.48
	TRP755	3.49		3.38	MET759	3.01	PHE778	3.98	PHE794	3.34	LEU721	3.59	PHE794	3.90
	MET759	3.99	LEU721	3.84	PHE778	3.74		3.36	LEU797	3.11		3.88	TYR890	3.79
	PHE778	3.58	MET759	3.17	PHE794	3.09	PHE794	3.96	TYR890	3.14	TRP755	3.45		
		3.65	LEU763	3.71	LEU797	3.21	LEU797	3.67			MET759	3.71		
	PHE794	3.25	PHE778	3.54	LEU887	3.52	LEU887	3.77			PHE778	3.52		
	LEU797	3.37	PHE794	3.35			TYR890	3.39			PHE794	3.38		
	PHE905	3.89	LEU797	3.30							THR894	3.53		
			LEU887	3.48							PHE905	3.31		
				3.52										
			TYR890	3.66										
Hydrogen Bonds	LEU715	2.47			GLN725	2.23	GLN725	1.92	GLN725	1.95			LEU718	2.18
	ASN719	2.45			LEU763	3.59	LEU763	3.52	LEU763	3.51			GLN725	1.88
					ARG766	2.09	ARG766	1.85	ARG766	1.80			LEU763	3.70
													ARG766	1.75
π-Stacking									PHE778	4.86			PHE778	5.19

	A28	
Hydrophobic	aa	Dist
	LEU715	3.21
	LEU718	3.07
	LEU797	3.83
	LEU887	2.91
	TYR890	3.42
	THR894	3.89
	PHE905	3.44
Hydrogen Bonds	GLN725	2.06
	ARG766	2.31
π-Stacking	PHE778	4.80

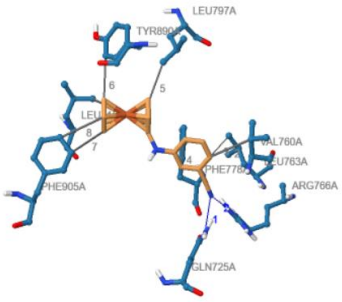
A0-PR	A1-PR	A2-PR
		
A3-PR	A4-PR	A5-PR
		
A6-PR	A7-PR	A8-PR
		
A9-PR	A10-PR	A11-PR
		
A12-PR	A13-PR	A14-PR
		



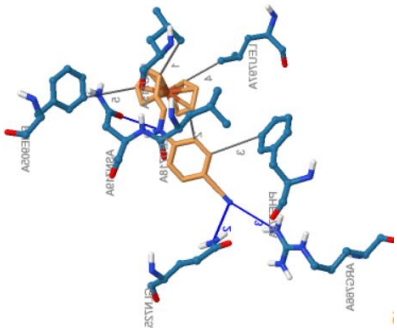
A15-PR



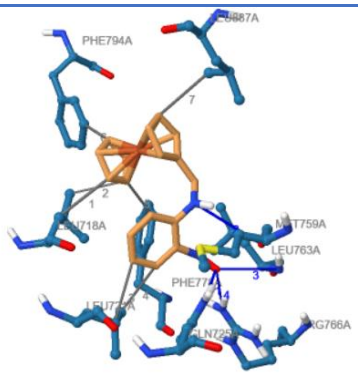
A16-PR



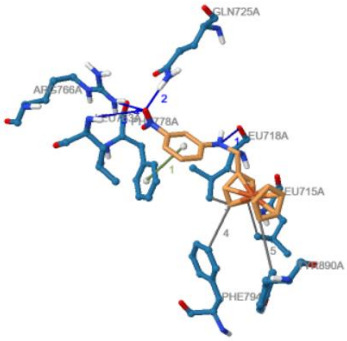
A17-PR



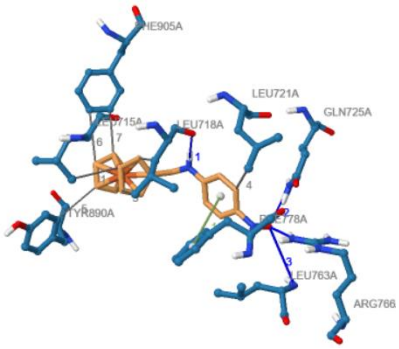
A18-PR



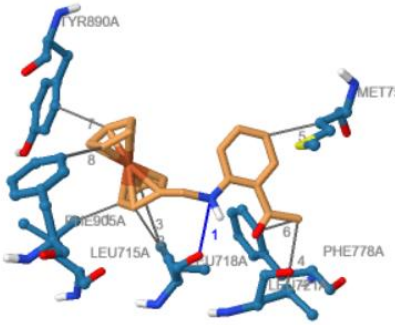
A19-PR



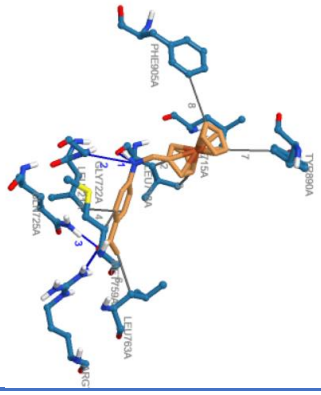
A20-PR



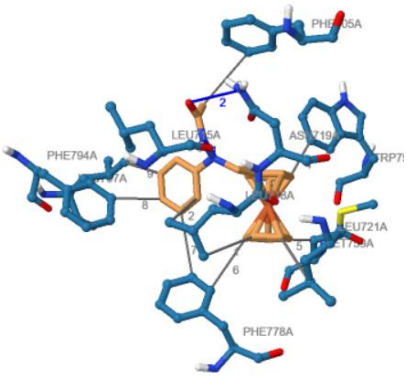
A21-PR

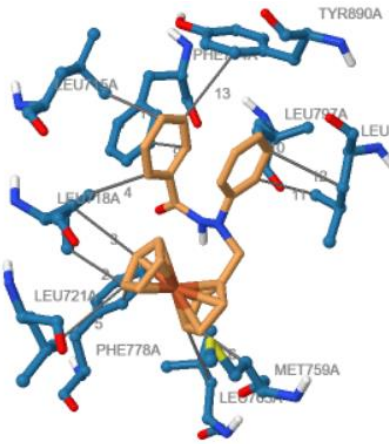


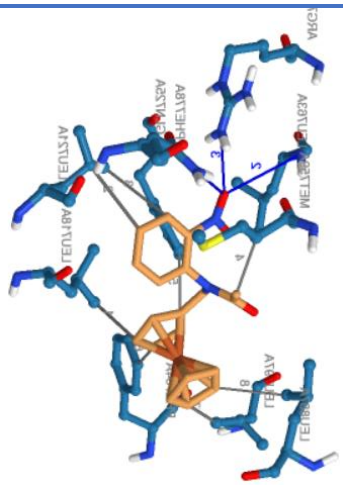
A22-PR



A23-PR







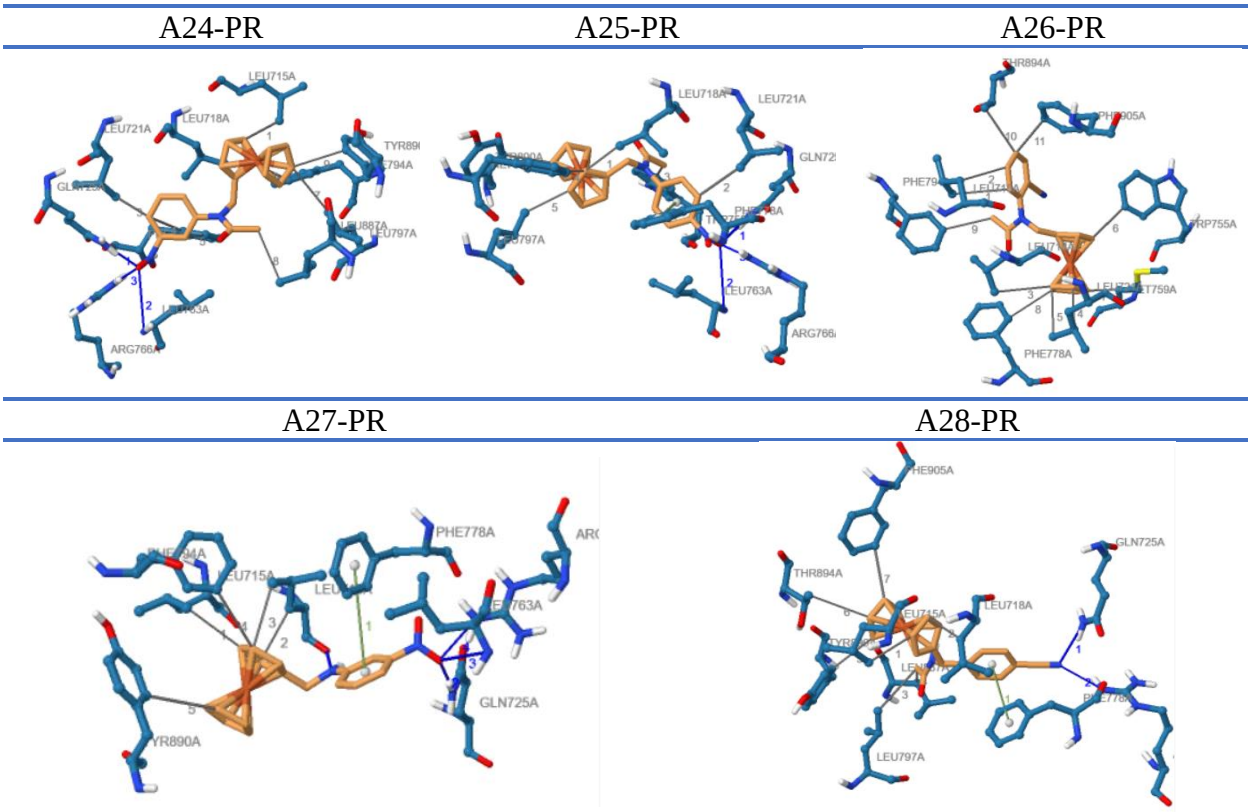


Figure III. 8. 3D interaction between studied ligands and the target PR (PDB ID: 1A28) where the grey lines show hydrophobic bonds and the bleu lines show the H-bonds.

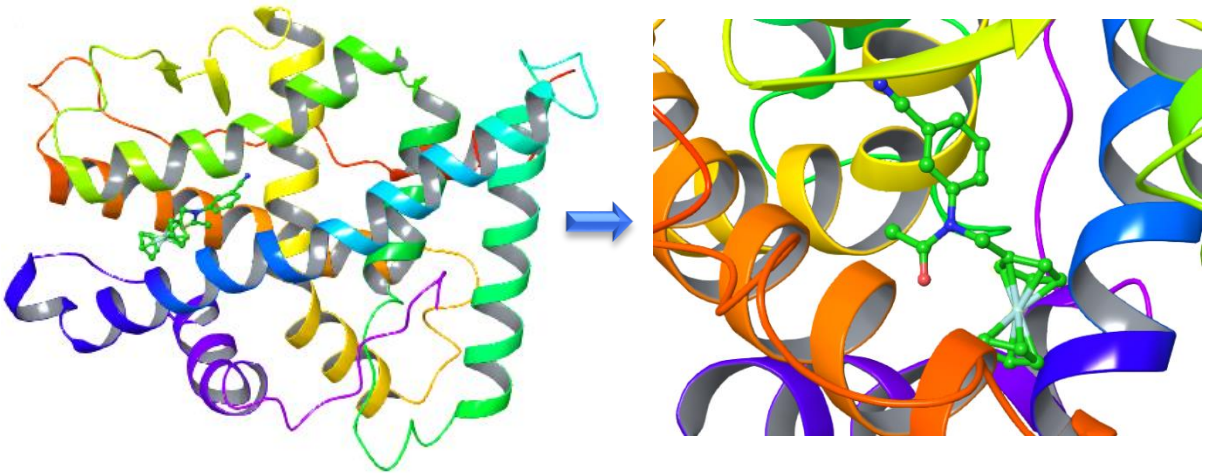


Figure III. 9. Docking pose of A27 with PR illustrating (a) the secondary structure view (b) zoomed view of the interactions.

III.4.4.2. Aromatase and 17β-HSD1

The inhibition of both aromatase and 17β-HSD1 enzymes is considered as crucial strategy to reduce the estrogen level. Table III.7. represents the binding free energies (ΔG) and the binding constants (K) of studied compounds interacted with Aromatase and 17β-HSD1.

Table III. 7. binding free energies (ΔG) and binding constant (K), for the studied molecules with Aromatase and 17β-HSD1.

Compounds	Aromatase		17 beta-HSD type 1	
	ΔG (Kcal.mol^{-1})	**K (M^{-1})**	**ΔG (Kcal.mol^{-1})**	**K (M^{-1})**
A0	-7.21	1.9 ×10^5	-6.76	0.89 ×10^5
A1	-7.77	4.89 ×10^5	-7.89	5.98 ×10^5
A2	-7.53	3.26 ×10^5	-8.81	28.22 ×10^5
A3	-9.28	62.32 ×10^5	-8.78	26.82 ×10^5
A4	-10.26	325.25 ×10^5	-9.12	47.59 ×10^5
A5	-7.85	5.59 ×10^5	-7.55	3.37 ×10^5
A6	-8.18	9.76 ×10^5	-7.64	3.93 ×10^5
A7	-7.80	5.14 ×10^5	-7.79	5.06 ×10^5
A8	-7.96	6.73 ×10^5	-7.73	4.57 ×10^5
A9	-8.28	11.54 ×10^5	-7.72	4.49 ×10^5
A10	-7.15	1.72 ×10^5	-7.25	2.03 ×10^5
A11	-8.56	18.51 ×10^5	-7.99	7.08 ×10^5
A12	-8.42	14.62 ×10^5	-7.51	3.15 ×10^5
A13	-7.69	4.27 ×10^5	-7.57	3.49 ×10^5
A14	-8.19	9.92 ×10^5	-7.71	4.42 ×10^5
A15	-8.86	30.70 ×10^5	-7.97	6.85 ×10^5
A16	-9.19	53.55 ×10^5	-8.30	11.94 ×10^5
A17	-9.79	147.26 ×10^5	-8.83	29.19 ×10^5
A18	-9.62	110.56×10^5	-8.23	10.61 ×10^5
A19	-7.8	5.14 ×10^5	-7.56	3.43 ×10^5
A20	-7.84	5.50 ×10^5	-7.77	4.89 ×10^5
A21	-7.33	2.33 ×10^5	-8.02	7.45 ×10^5
A22	-8.73	24.65 ×10^5	-8.57	18.83 ×10^5
A23	-8.6	19.80 ×10^5	-8.53	17.60 ×10^5
A24	-9.97	199.47 ×10^5	-9.83	157.54 ×10^5
A25	-10.07	236.10 ×10^5	-9.34	68.96 ×10^5
A26	-7.8	5.14 ×10^5	-8.29	11.74 ×10^5
A27	-8.45	15.38 ×10^5	-8.71	23.84 ×10^5
A28	-8.42	14.62 ×10^5	-8.80	27.75 ×10^5
ASD	-9.25	59.25 ×10^5	-	-
E2B	-	-	-11.61	3167.16 ×10^5

In summary, the obtained results showed high affinity between all molecules and both targets.

The studied compounds interacted to aromatase enzyme with binding free energy in the range from -7.15 to -10.26 Kcal.mol^{-1} (table. III. 7), where A4 and A25 showed the lowest binding free energy equal to -10.26 and -10.07 Kcal.mol^{-1}, respectively. However, A10 showed the highest binding free energy equal to -7.10 Kcal.mol^{-1}. Androstenedione (ASD), which is an androgen steroid hormone and a precursor of estrogen, was also docked into the aromatase as a reference ligand. It can be seen that the compounds A3, A4, A17, A18, A24 and A24 showed a higher docking score than Androstenedione's which is equal to -9.25 Kcal.mol^{-1} .

Molecular docking results revealed that most of the compounds could occupy at the same binding site of the natural substrate Androstenedione on the aromatase pocket, and share the same binding residues with those of the Androstenedione which are Arg 115, Ile 305, Ala 306, Asp 309, Thr 310, Phe 221, Trp 224, Ile133, Phe 134, Val 370, Leu372, Val 373, Met 374, Leu 477 and Ser 478 [52]. These suggested that the investigated compounds may elicit their aromatase inhibitory activities via acting as competitive inhibitors (table III.8 and figure III.10).

Table III. 8. Distances of formed bonds between the ligands and aromatase residues

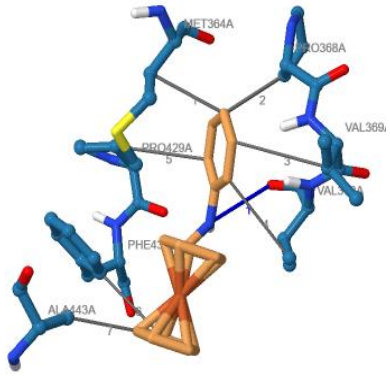
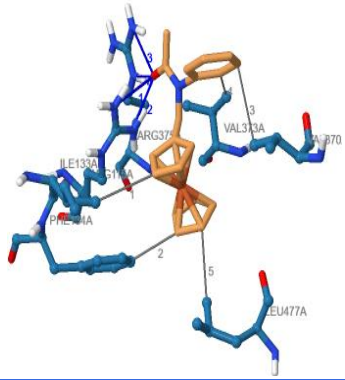
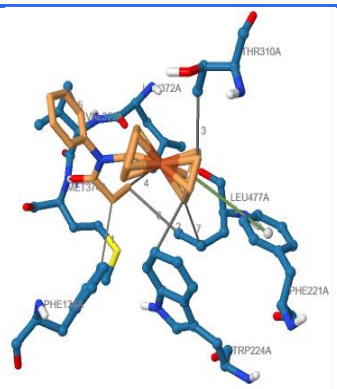
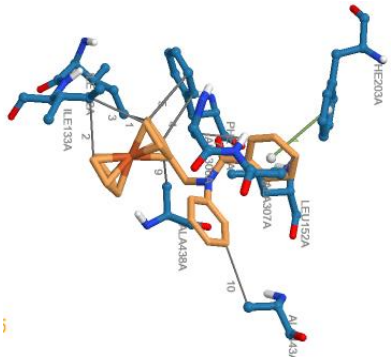
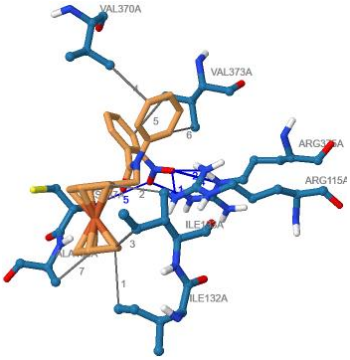
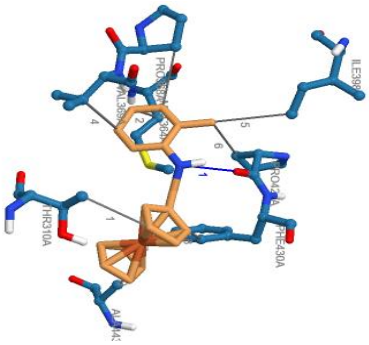
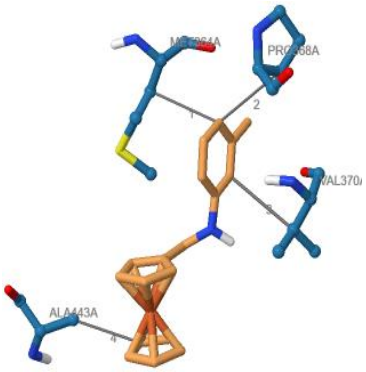
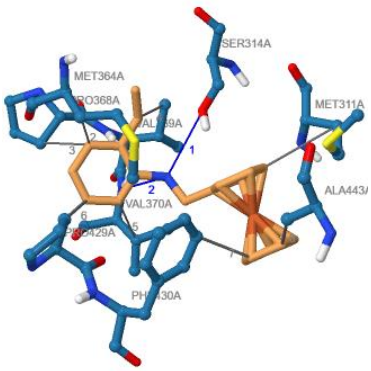
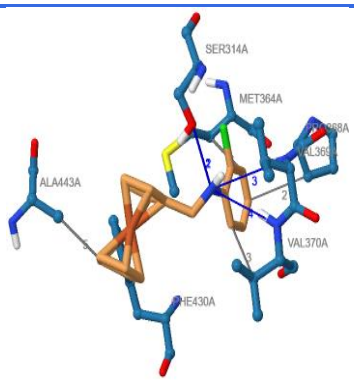
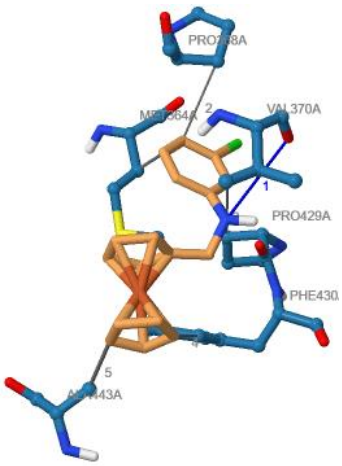
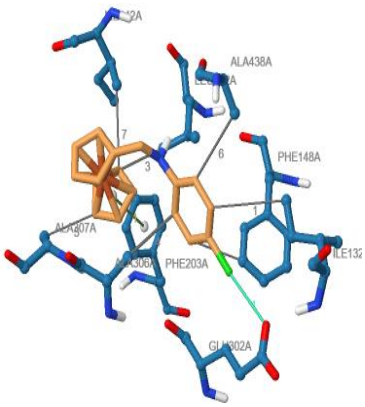
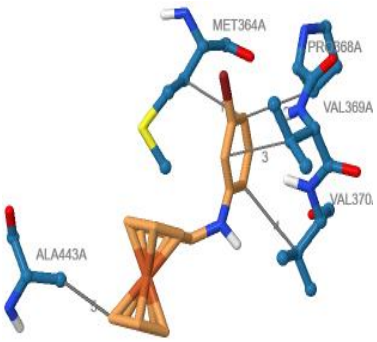
Aromatase														
Bonds type	A0		A1		A2		A3		A4		A5		A6	
	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist
Hydrophobic	MET364	3.28	ILE133	3.60	PHE134	3.47	ILE132	3.46	ILE132	3.30	THR310	3.56	MET364	3.49
	PRO368	3.58	PHE134	3.28	TRP224	3.37	ILE133	3.30	ILE133	3.45	MET364	3.48	PRO368	3.36
	VAL369	3.95	VAL370	3.68	THR310	3.17		3.23		2.94	PRO368	3.48	VAL370	3.72
		3.97	VAL373	3.31	LEU372	3.23	PHE148	3.54	VAL370	3.44	VAL369	3.90	ALA443	3.14
	PRO429	3.98	LEU477	3.45	VAL373	3.30		3.67	VAL373	3.04	ILE398	3.96		
	PHE430	3.90			LEU477	3.79	LEU152	3.20		3.67	PRO429	3.84		
	ALA443	3.15				3.30	ALA306	3.19	ALA438	3.07	PHE430	3.22		
							ALA307	3.50				3.62		
							ALA438	3.49			ALA443	3.11		
							ALA443	3.44						
Hydrogen Bonds	VAL270	3.50	ARG115	2.12	MET374	1.88			ARG115	2.92	PRO429	1.99		
				1.81						2.72				
			ARG375	2.72					ARG375	2.91				
				2.36						3.25				
									CYS437	4.00				
π-Stacking					PHE221	4.77	PHE203	4.83						
	A7		A8		A9		A10		A11		A12		A13	
	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist
Hydrophobic	MET311	3.82	MET364	3.09	MET364	3.26	ILE132	3.98	MET364	3.33	MET311	3.81	ILE132	3.66
	MET364	3.20	PRO368	3.43	PRO368	3.38	PHE148	3.40	PRO368	3.37	MET364	3.20		3.91
	PRO368	3.45	VAL370	3.22	PRO429	3.78	LEU152	2.96	VAL369	3.99	PRO368	3.41	ILE133	3.49
	VAL369	3.26	PHE430	3.08	PHE430	3.95	ALA306	3.38	VAL370	3.81	VAL370	3.25	PHE148	3.13
	VAL370	3.72	ALA443	3.13	ALA443	3.11	ALA307	3.13	ALA443	3.14		3.21	LEU152	2.99
	PRO429	3.87					ALA438	3.61			PHE430	3.07	PHE203	3.34
	PHE430	3.34					ILE42	3.75			ALA443	2.94	ALA306	3.25
	ALA443	2.91												3.14
													ALA307	3.28
													ILE442	3.69
Hydrogen Bonds	SER314	2.84	SER314	2.97	VAL370	3.37					SER314	2.79	ALA438	1.94
	VAL370	2.89	SER314	2.97							VAL369	3.81	ILE442	3.69
			VAL369	4.03							VAL370	3.74		
			VAL370	2.75										
Halogen bonds							GLU302	3.17						
π-Stacking							PHE203	5.04						
	A14		A15		A16		A17		A18		A19		A20	
	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist

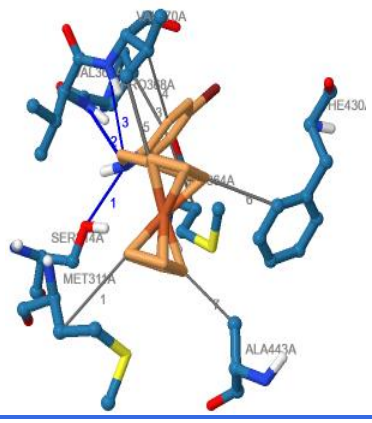
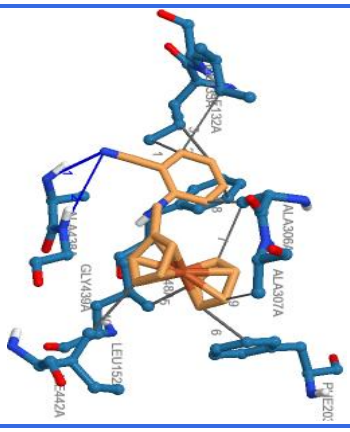
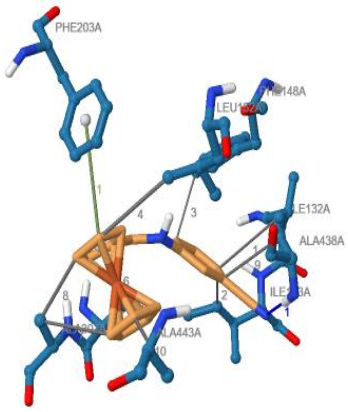
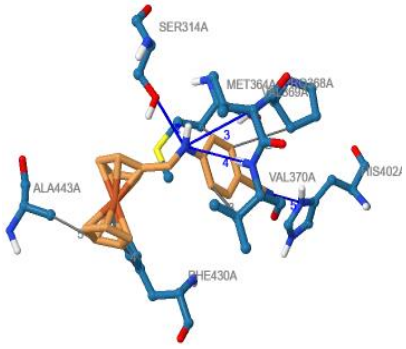
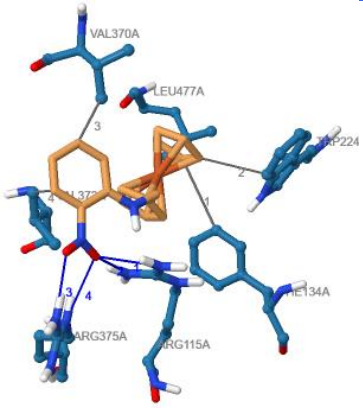
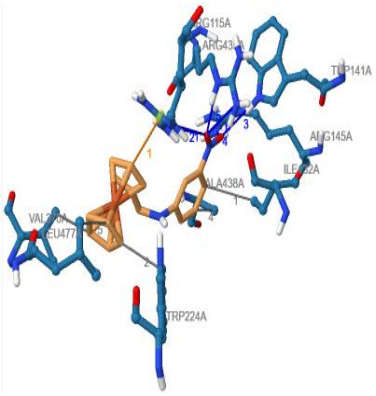
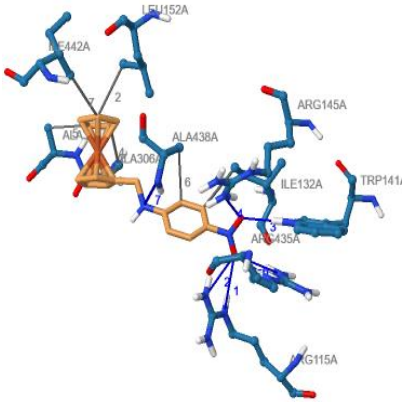
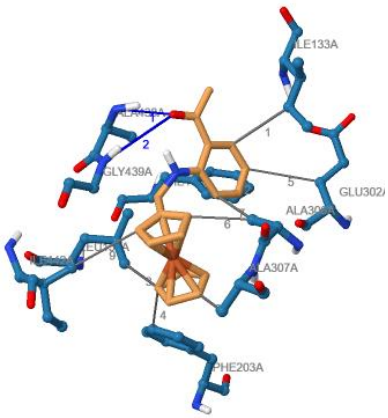
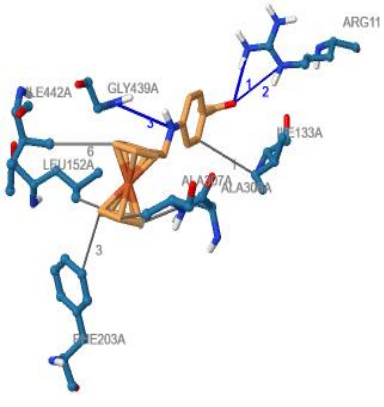
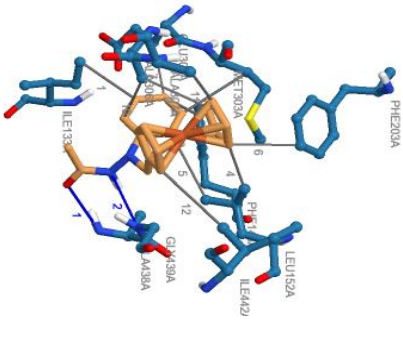
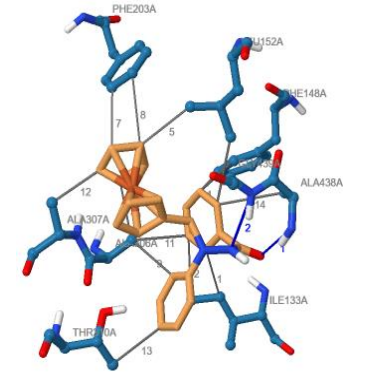
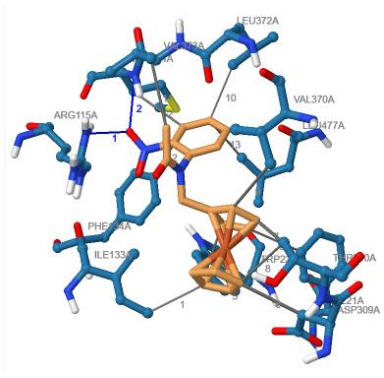
Hydrophobic	ILE132	3.84	MET364	3.29	PHE134	3.74	ILE132	3.43	ILE132	3.07	ILE133	3.39	ILE133	3.94
	ILE133	3.44	PRO368	3.47	TRP224	3.58	TRP224	3.40	LEU152	3.35	PHE148	3.18	LEU152	2.93
	PHE148	3.31	VAL370	3.24	VAL370	3.67	VAL370	3.49	ALA306	3.10	LEU152	3.14	PHE203	3.43
	LEU152	3.85	PHE430	3.18	VAL373	3.10	ALA438	3.50		3.14	PHE203	3.33	ALA306	3.38
	ALA306	3.69	ALA443	3.33	LEU477	3.20	LEU477	3.39	ALA307	3.86	GLU302	3.93	ALA307	3.10
		3.74							ALA438	3.94	ALA306	3.13	ILE442	3.98
	ALA307	3.60							ILE442	3.58		3.13		
		3.07									ALA307	2.94		
	ALA438	3.92									ILE442	3.80		
	ALA443	3.85												
Hydrogen Bonds	ALA438	1.98	SER314	2.38	ARG115	2.92	ARG115	3.20	ARG115	2.76	ALA307	1.83	ARG115	2.03
				2.18		2.74		2.85		2.04	GLY439	3.22		3.67
			VAL369	3.18	ARG375	2.97	TRP141	3.23	TRP141	2.09			GLY439	3.97
				2.91		3.26	ARG145	2.93	ARG145	2.07				
			HIS402	2.54			ARG435	2.66	ARG435	1.78				
								3.15		2.61				
									ALA438	2.78				
π-Stacking	PHE203	4.87												
π-Cation interaction							ARG115	1.29						

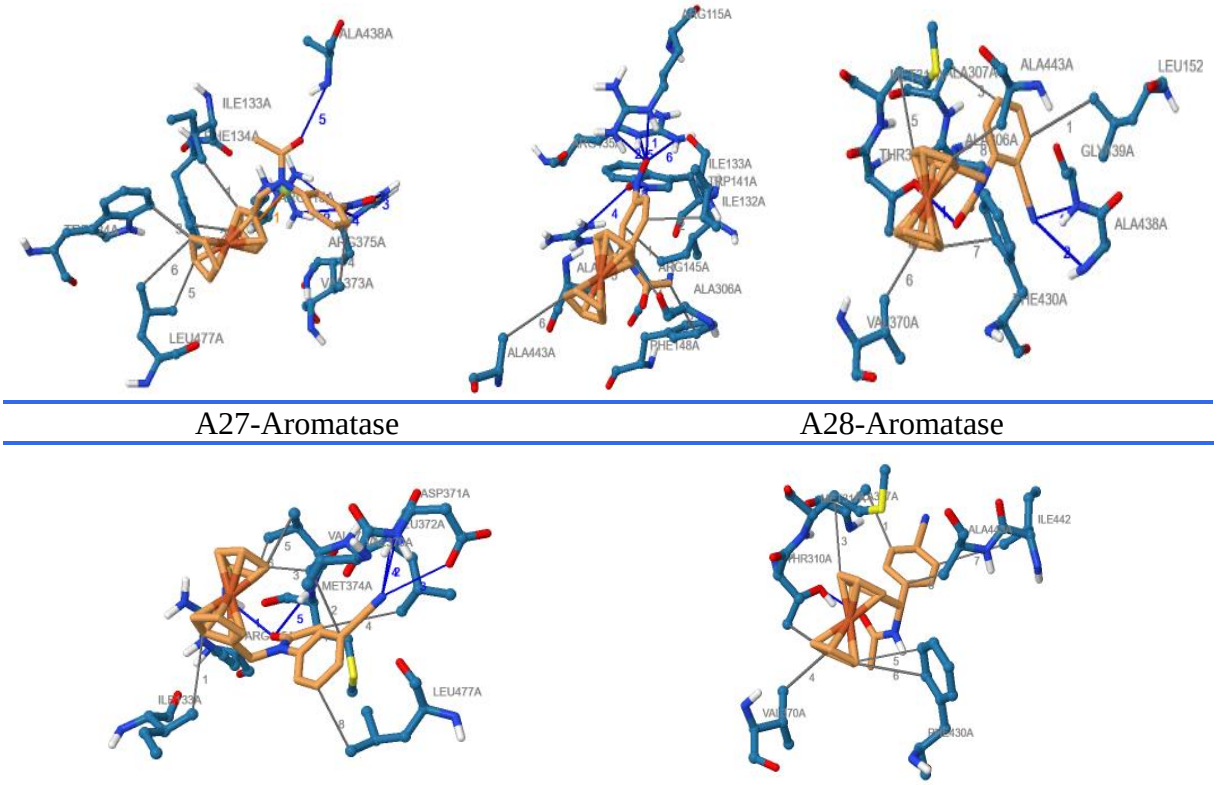
	A21		A22		A23		A24		A25		A26		A27	
	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist
Hydrophobic	ILE133	3.56	ILE133	3.73	ILE133	3.37	ILE133	3.46	ILE132	3.17	LEU152	3.99	ILE133	3.75
	PHE148	3.99		3.78	PHE134	3.89	PHE134	3.45	ILE133	3.38	ALA306	3.53	VAL370	3.39
		3.12	PHE148	3.25	PHE221	3.63	TRP224	3.33	PHE148	3.53	ALA307	3.56		3.33
	LEU152	3.11		3.37	TRP224	3.15	VAL373	3.45	ALA306	3.31	THR310	3.58	LEU372	3.94
		3.86	LEU152	3.07		3.63	LEU477	3.14	ALA438	2.93	MET311	3.89	VAL373	3.34
	PHE203	3.44		3.62	ASP309	3.57		3.87	ALA443	3.97	VAL370	3.04		3.16
	GLU302	3.94	PHE203	3.78	THR310	3.32					PHE430	3.86	MET374	3.63
	MET303	3.93		3.69		3.27					ALA443	3.02	LEU477	3.43
	ALA306	3.34	ALA306	3.10	VAL370	3.80								
		3.40		3.16	LEU372	3.16								
	ALA307	3.12	ALA306	3.29	VAL373	3.42								
	ILE442	3.98	ALA307	2.89	MET374	3.80								
			THR310	3.35	LEU477	3.29								
			ALA438	3.66										
Hydrogen Bonds	ALA307	1.94	ALA438	2.09	ARG115	2.05	ARG115	2.40	ARG115	2.16	THR310	2.29	ARG115	2.59
	GLY439	2.51	GLY439	2.32	MET374	1.98		1.90		1.96	ALA438	3.21	ASP371	2.43
							ARG375	1.83	TRP141	2.02	GLY439	2.37		2.56
								1.85	ARG145	2.57			LEU372	2.19
							ALA438	2.70	ARG435	1.82			MET374	2.31
										2.25				

π-Cation interaction							ARG115	3.95						

	A28	
	AA	Dist
Hydrophobic	ALA307	3.37
	THR310	3.60
	MET311	4.00
	VAL370	2.96
	PHE430	3.38
		3.94
	ILE442	3.78
	ALA443	3.57
Hydrogen Bonds	THR310	1.73

A0-Aromatase	A1-Aromatase	A2-Aromatase
		
A3-Aromatase	A4-Aromatase	A5-Aromatase
		
A6-Aromatase	A7-Aromatase	A8-Aromatase
		
A9-Aromatase	A10-Aromatase	A11-Aromatase
		

A12-Aromatase	A13-Aromatase	A14-Aromatase
		
A15-Aromatase	A16-Aromatase	A17-Aromatase
		
A18-Aromatase	A19-Aromatase	A20-Aromatase
		
A21-Aromatase	A22-Aromatase	A23-Aromatase
		
A24-Aromatase	A25-Aromatase	A26-Aromatase



A27-Aromatase A28-Aromatase

Figure III. 10. 3D interaction between studied ligands and the target Aromatase (PDB ID: 3S79) where the grey lines show hydrophobic bonds and the bleu lines show the H-bonds.

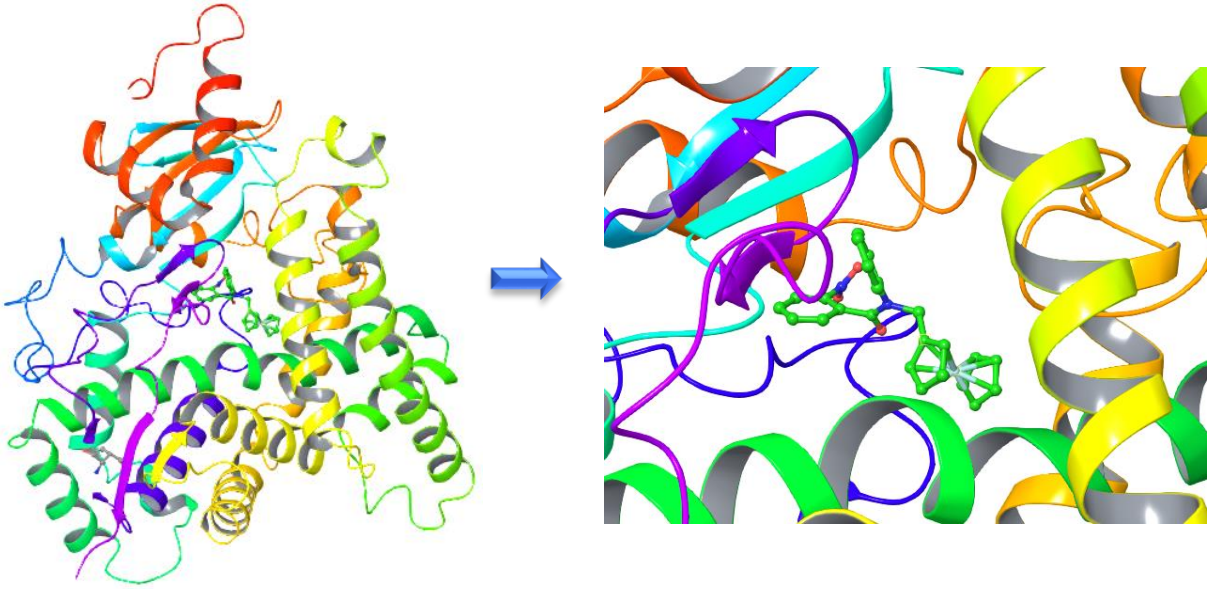


Figure III. 11. Docking pose of A4 with aromatase illustrating (a) the secondary structure view (b) zoomed view of the interactions.

The docking results of the studied molecules with 17β-HSD1 showed medium binding affinity with docking scores in the range from -6.76 to -9.83 Kcal.mol^{-1} comparing to the reference ligand E2B (-11.61 Kcal.mol^{-1}), the top three derivatives are A4, A24 and A25 possessing the lowest binding free energies equal to -9.12, -9.83 and -9.34 Kcal.mol^{-1}, respectively. However, the least potent derivative is A0 with ΔG equal to -6.76 Kcal.mol^{-1}.

Moreover, all compounds showed at least 3 hydrophobic bonds with different amino acids, some of them made one H-bond or more with for example Tyr155, Lys159, Gly186, thr190 and val188. Furthermore, another bond of the type π- π stacking was observed with the amino acids Tyr155 and Phe192.

It can be seen that most of the compounds are located at the 17β-HSD1 binding cavity, which is formed by the following amino acids: Ser 142, Val 143, Lys 149, Gly 185, Pro 187, Tyr 218, His 221, Ser 222, Phe 226, Phe 259, Met 279, and Glu 282 [53].

It was found that the natural substrate binds to17β-HSD1 by hydrogen bonds with the residues Tyr155, Ser142, His221 and Glu282 [53], also E2B binds to 17β-HSD1 via hydrogen and π–π bonds with the residues Tyr155[54]. Similar to these findings, 13 of our derivatives (A1, A2, A14, A15, A16, A17, A18, A21, A23, A24, A25, A26 and A27) have made similar bonds with the amino acid Tyr155. These indicate that the derivatives may act as competitive inhibitors (table III.9 and figure III.12).

Table III. 9. Distances of formed bonds between the ligands and 17β-HSD1 residues

17β-HSD1														
Bonds type	A0		A1		A2		A3		A4		A5		A6	
	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist
Hydrophobic	VAL188	3.48	LEU96	3.47	LEU96	3.33	THR140	3.12	ILE14	3.42	ILE14	3.68	VAL188	3.80
	PHE192	3.95		3.82	TYR155	3.71	PHE192	3.67	THR140	3.03	VAL188	3.97	THR190	3.94
		3.58	TYR155	3.90		3.59		3.80	VAL188	3.84		3.53	PHE192	3.84
		3.71	VAL188	3.17	PHE192	3.36			PHE192	3.26	THR190	3.93		3.49
	PHE226	3.32		3.31	VAL196	3.40					PHE192	3.95	PHE226	3.73
		3.48	PHE226	3.49	PHE226	3.37						3.33		
											PHE226	3.24		
Hydrogen Bonds	THR190	1.98			VAL188	1.66	ILE14	2.12	ILE14	1.84	THR190	2.06	THR190	1.83
									GLY15	3.14				
									ASN90	1.92				
π-Stacking			TYR155	4.88	TYR155	4.67					PHE192	4.28		
			TYR155	4.75										

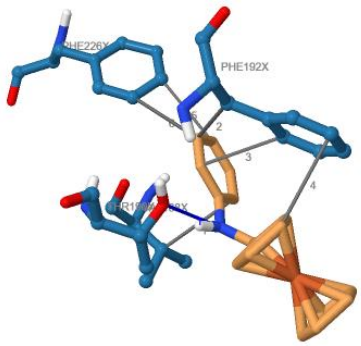
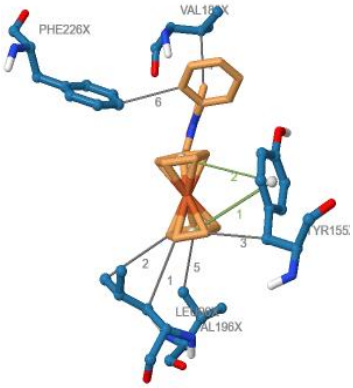
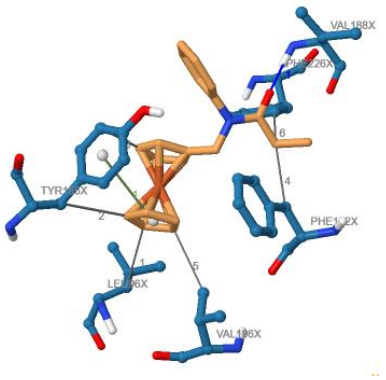
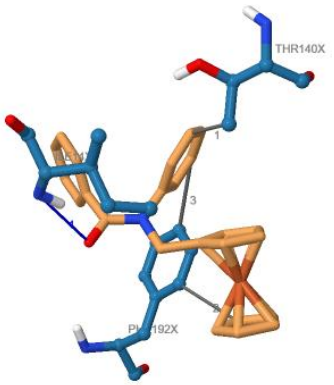
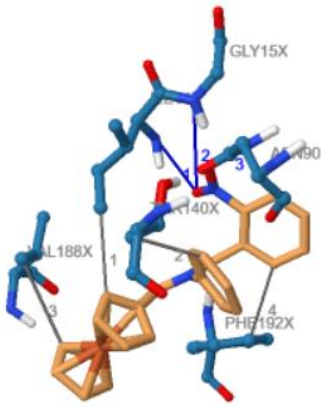
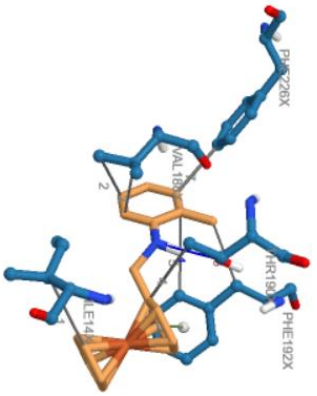
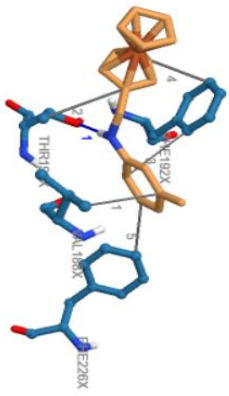
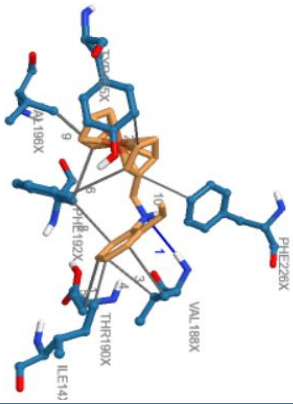
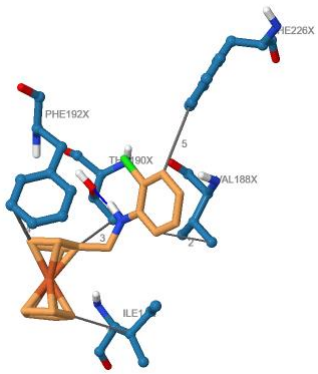
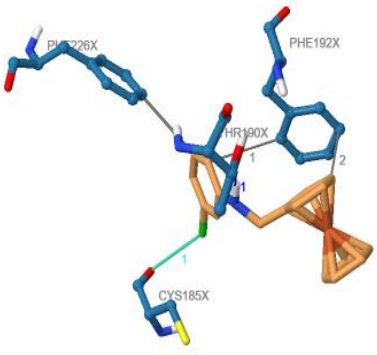
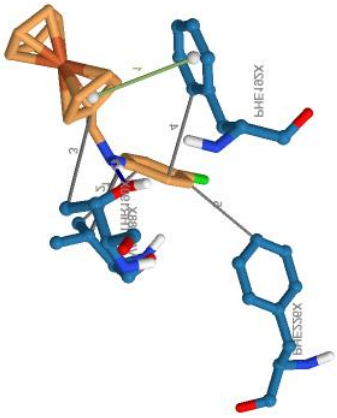
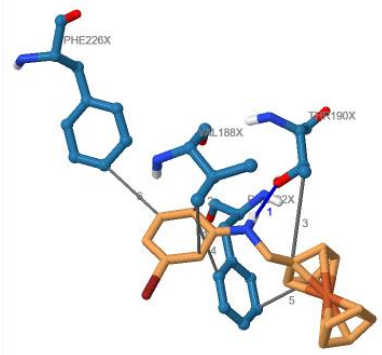
	A7		A8		A9		A10		A11		A12		A13	
	Aa	Dist	Aa	Dist	Aa	Dist	Aa	Dist	Aa	Dist	Aa	Dist	Aa	Dist
Hydrophobic	ILE14	3.35	ILE14	3.84	PHE192	3.77	VAL188	3.69	VAL188	3.8	VAL188	3.83	VAL188	3.99
	TYR155	3.58	VAL188	3.52	PHE192	3.44		3.41		3.88	THR190	3.98		3.64
	VAL188	3.43	THR190	3.85	PHE226	3.71	THR190	3.85	THR190	3.99	PHE192	3.75	THR190	3.68
		3.57	PHE192	3.47			PHE192	3.81	PHE192	3.97	PHE226	3.74	PHE192	3.68
	THR190	3.85	PHE226	3.61			PHE226	3.59		3.59				3.83
	PHE192	3.63							PHE226	3.55			PHE226	3.30
		3.44							THR190	2.97				
		3.67												
	VAL196	3.36												
	PHE226	3.71												
Hydrogen Bonds	VAL188	2.76	THR190	1.91	THR190	1.82	THR190	1.84			THR190	1.82	THR190	2.78
Halogen bonds					CYS185	3.96							THR190	3.65
													ALA191	3.46
π-Stacking							PHE192	4.32			PHE192	4.03	PHE192	4.40

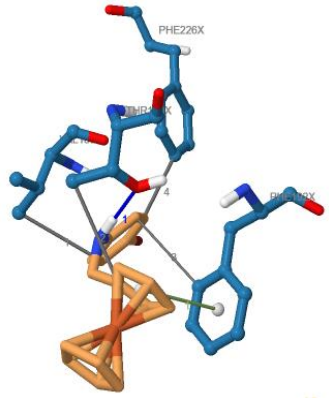
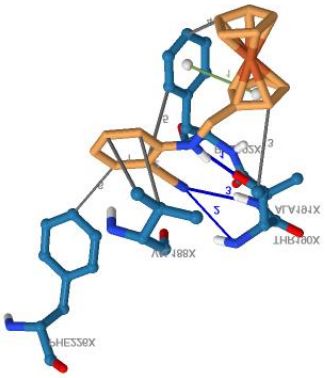
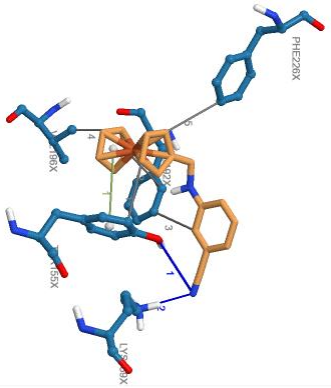
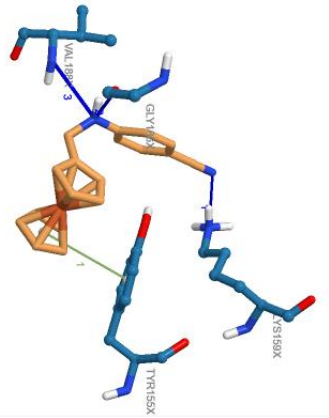
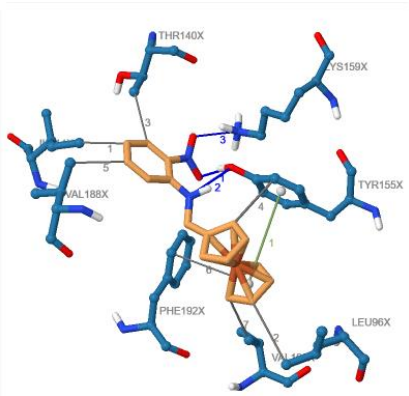
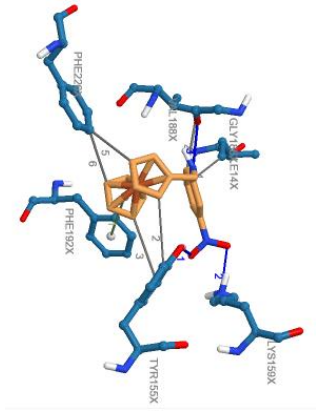
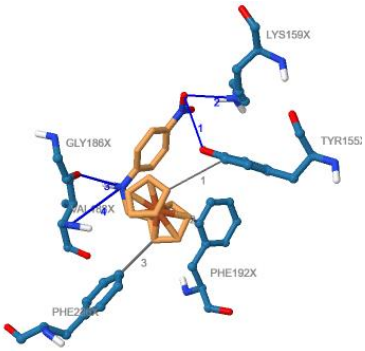
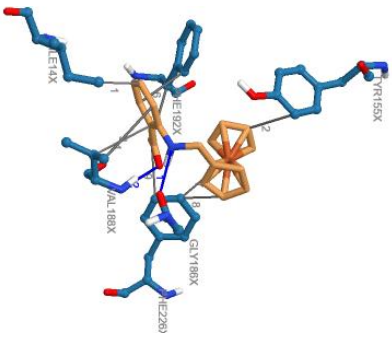
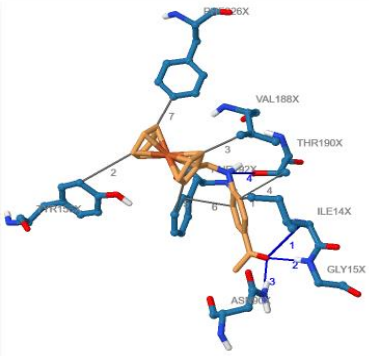
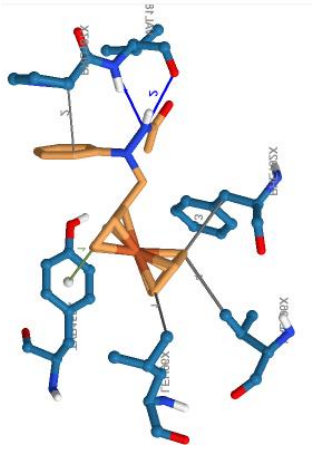
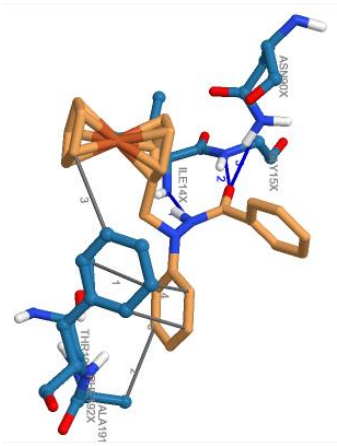
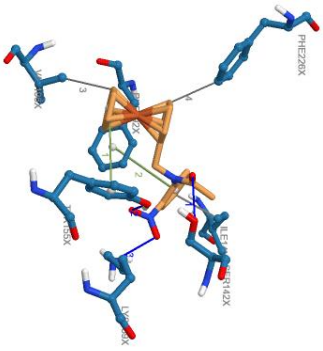
	A14		A15		A16		A17		A18		A19		A20	
	Aa	Dist	Aa	Dist	AA	Dist								
Hydrophobic	TYR155	3.41	LYS159	2.89	ILE14	3.19	ILE14	3.26	TYR155	3.60	ILE14	3.16	ILE14	3.15
	PHE192	3.39	GLY186	2.82	LEU96	3.83	TYR155	3.91	PHE192	3.36	TYR155	3.53	TYR155	3.88
		3.88	VAL188	2.69	THR140	3.72	TYR155	3.98	PHE226	3.02	VAL188	3.86	VAL188	3.07
	VAL196	3.70			TYR155	3.28	VAL188	3.55			VAL188	3.71	THR190	3.83
	PHE226	3.62			VAL188	3.02	PHE226	3.45			PHE192	3.60	PHE192	3.60

					PHE192	3.69		3.39			PHE192	3.35		3.35
					VAL196	3.38					PHE226	3.62	PHE226	3.69
												3.66		
												3.36		
Hydrogen Bonds	TYR155	3.33			TYR155	1.77	TYR155	2.10	TYR155	3.33	GLY186	2.28	ILE14	2.81
	LYS159	3.07				1.71	LYS159	1.75	LYS159	2.04	VAL188	1.66	GLY15	2.08
					LYS159	2.03	GLY186	1.95	GLY186	2.32			ASN90	2.16
									VAL188	3.00			THR190	1.85
π-Stacking	TYR155	5.15	TYR155	4.77	TYR	2.07	PHE192	4.87						

	A21		A22		A23		A24		A25		A26		A27	
	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist
Hydrophobic	LEU96	3.82	THR190	3.97	ILE14	3.29	LEU96	3.81	LEU96	3.76	ILE14	3.48	LEU96	3.65
	PRO187	4.00	ALA191	3.28	PHE192	3.45		3.46	LEU96	3.45	VAL188	3.49		3.40
	PHE192	3.39	PHE192	3.35	VAL196	3.33	TYR155	3.31	VAL188	3.67	THR190	3.78	VAL188	3.55
	VAL196	3.36		3.42	PHE226	3.17	PHE192	3.83	PHE192	3.32	PHE192	3.57	VAL196	3.26
				3.50				3.83		3.49	VAL196	3.32	PHE226	3.80
							VAL196	3.40	MET193	3.55	PHE226	3.40		
									VAL196	3.36				
									PHE226	3.20				
Hydrogen Bonds	VAL188	2.28	ILE14	2.77	SER142	2.83	TYR155	2.06	THR190	3.24	VAL188	1.93	LYS159	2.10
		2.15	GLY15	1.69	TYR155	1.77	LYS159	2.10		1.79				
			ASN90	2.07	LYS159	2.23	VAL188	1.98	ALA191	2.94				
									MET193	2.04				
π-Stacking	TYR155	4.67			TYR155	5.23	TYR155	5.07	TYR155	5.01	TYR155	4.67	TYR155	5.16
					PHE192	4.64				4.55	PHE192	4.66		4.77

	A28	
	AA	Dist
Hydrophobic	LEU149	3.30
	TYR155	3.07
	PHE192	3.20
	MET193	3.72
	PHE226	3.31
		3.46
Hydrogen Bonds	THR190	1.94
		3.14
	ALA191	2.35
π-Stacking	PHE192	4.88

A0-17β-HSD1	A1-17β-HSD1	A2-17β-HSD1
		
A3-17β-HSD1	A4-17β-HSD1	A5-17β-HSD1
		
A6-17β-HSD1	A7-17β-HSD1	A8-17β-HSD1
		
A9-17β-HSD1	A10-17β-HSD1	A11-17β-HSD1
		

A12-17β-HSD1	A13-17β-HSD1	A14-17β-HSD1
		
A15-17β-HSD1	A16-17β-HSD1	A17-17β-HSD1
		
A18-17β-HSD1	A19-17β-HSD1	A20-17β-HSD1
		
A21-17β-HSD1	A22-17β-HSD1	A23-17β-HSD1
		

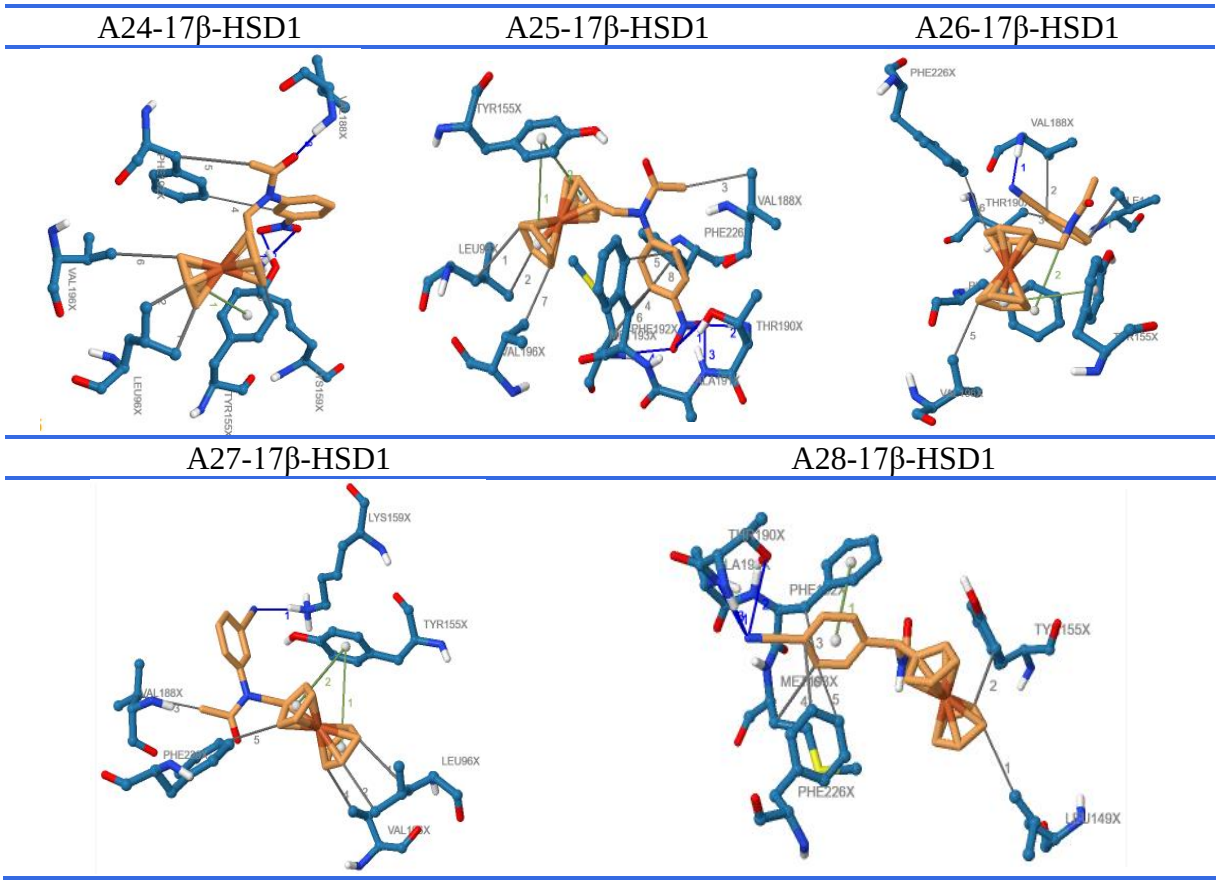


Figure III. 12. 3D interaction between studied ligands and the target 17β-HSD1 (PDB ID: 3HB5) where the grey lines show hydrophobic bonds and the bleu lines show the H-bonds.

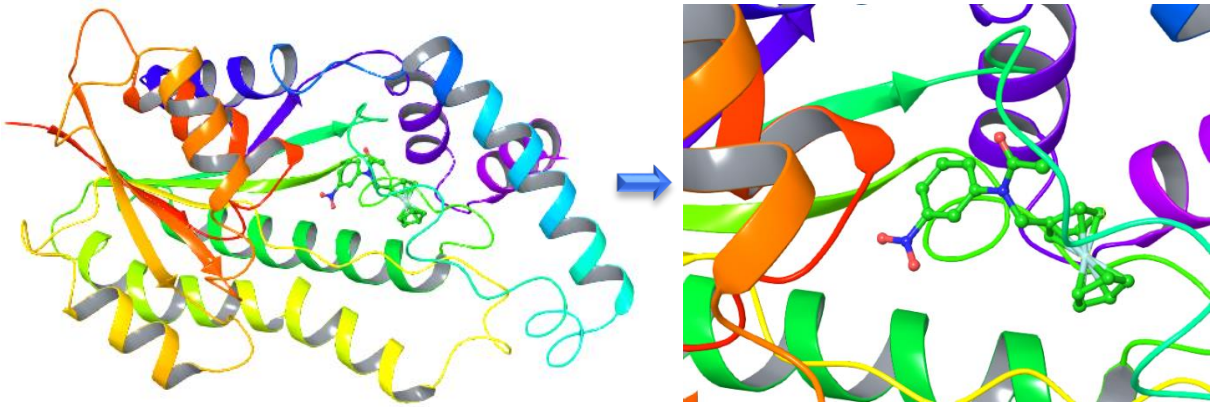


Figure III. 13. Docking pose of A24 with 17β-HSD1 illustrating (a) the secondary structure view (b) zoomed view of the interactions.

III.5. Conclusion

In summary, the ferrocenylmethylanilines derivatives showed an important anti-proliferative effect against MCF-7 cell lines *in vitro*, which was explained using molecular docking by the interaction of these compounds with the ERα and PR. These findings are worth to be followed by *in* vitro study of the inhibitory effect of the studied compounds against these receptors. Furthermore, docking study against aromatase shows that A4 and A25 were the most active with the lowest binding free energies and more than three hydrogen bonds compared with the rest of studied compounds and ASD. However, moderate affinity was predicted against 17β-HSD1 for all compounds compared to E2B.

References

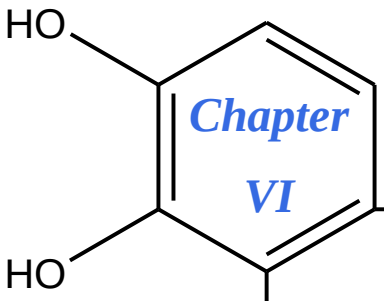
1. Belgacem, T., synthèse, cyclisation, étude electrochimique et structurale de quelques N-acyl-n'alkylférrocenyl méthyle hydrazide. 2007, university Elhadj Lakhder of batna.
2. Oumelkhair, R., Synthèse et étude électrochimique et structurale de quelques N-ferrocényl méthyle anilines substitués Applications à l'inhibition de la corrosion aqueuse. 2014, University Kasdi Merbah of Ouergla.
3. Hadia, H., Synthèse, caractérisation et évaluation in vitro de l'activité biologique et du pouvoir mutagène de quelques N-(ferrocenylmethyl)nitroaniline. 2017, University of El Oued.
4. Meriem, H., Synthèse, caractérisation et évaluation in vitro de l'activité biologique et du pouvoir mutagène de quelques N-(ferrocenylmethyl) aminobenzonitrile, in chemistry departement 2017, university of Eloued.
5. Elhafnaoui, L., Synthesis, characterization and in vitro and in silico study of biological activities of some ferrocenylmethyl-nucleic bases derivatives. 2020, University Kasdi Merbah of Ouargla.
6. Langdon, S.P., Cancer cell culture. 2010: Springer.
7. Kasper, C., V. Charwat, and A. Lavrentieva, Cell Culture Technology. 2018: Springer.
8. Soule, H., et al., A human cell line from a pleural effusion derived from a breast carcinoma. Journal of the national cancer institute, 1973. 51(5): p. 1409-1416.
9. Lee, A.V., S. Oesterreich, and N.E. Davidson, MCF-7 cells—changing the course of breast cancer research and care for 45 years. JNCI: Journal of the National Cancer Institute, 2015. 107(7).
10. Horwitz, K., M. Costlow, and W. McGuire, MCF-7: a human breast cancer cell line with estrogen, androgen, progesterone, and glucocorticoid receptors. Steroids, 1975. 26(6): p. 785-795.
11. Lippman, M.E. and G. Bolan, Oestrogen-responsive human breast cancer in long term tissue culture. Nature, 1975. 256(5518): p. 592-593.
12. Brooks, S.C., E.R. Locke, and H.D. Soule, Estrogen receptor in a human cell line (MCF-7) from breast carcinoma. Journal of Biological Chemistry, 1973. 248(17): p. 6251-6253.
13. Huang, Y., et al., Estrogen increases intracellular p26Bcl-2 to p21Bax ratios and inhibits taxol-induced apoptosis of human breast cancer MCF-7 cells. Breast cancer research and treatment, 1997. 42(1): p. 73-81.
14. Perillo, B., et al., 17β-Estradiol inhibits apoptosis in MCF-7 cells, inducing bcl-2 expression via two estrogen-responsive elements present in the coding sequence. Molecular and cellular biology, 2000. 20(8): p. 2890-2901.
15. Zheng, A., A. Kallio, and P. Harkonen, Tamoxifen-induced rapid death of MCF-7 breast cancer cells is mediated via extracellularly signal-regulated kinase signaling and can be abrogated by estrogen. Endocrinology, 2007. 148(6): p. 2764-2777.

16. Kallio, A., et al., Role of mitochondria in tamoxifen-induced rapid death of MCF-7 breast cancer cells. Apoptosis, 2005. 10(6): p. 1395-410.
17. Cailleau, R., M. Olive, and Q.V. Cruciger, Long-term human breast carcinoma cell lines of metastatic origin: preliminary characterization. In vitro, 1978. 14(11): p. 911-915.
18. Welsh, J., Chapter 40 - Animal Models for Studying Prevention and Treatment of Breast Cancer, in Animal Models for the Study of Human Disease, P.M. Conn, Editor. 2013, Academic Press: Boston. p. 997-1018.
19. Mersin, H., et al., The prognostic importance of triple negative breast carcinoma. The Breast, 2008. 17(4): p. 341-346.
20. Lønne, G.K., et al., Protein kinase Cδ supports survival of MDA-MB-231 breast cancer cells by suppressing the ERK1/2 pathway. Journal of Biological Chemistry, 2009. 284(48): p. 33456-33465.
21. Yang, W., et al., Akt isoform-specific inhibition of MDA-MB-231 cell proliferation. Cellular Signalling, 2011. 23(1): p. 19-26.
22. Tan, A.R. and S.M. Swain, Therapeutic strategies for triple-negative breast cancer. The Cancer Journal, 2008. 14(6): p. 343-351.
23. Dadashpour, S., et al., Synthesis and anticancer activity of Benzimidazole/Benzoxazole substituted triazolotriazines in hepatocellular carcinoma. Anti-Cancer Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry-Anti-Cancer Agents), 2019. 19(17): p. 2120-2129.
24. Vessieres, A., et al., Modification of the estrogenic properties of diphenols by the incorporation of ferrocene. Generation of antiproliferative effects in vitro. Journal of medicinal chemistry, 2005. 48(12): p. 3937-3940.
25. Vladusic, E.A., et al., Expression of estrogen receptor β messenger RNA variant in breast cancer. Cancer research, 1998. 58(2): p. 210-214.
26. Paterni, I., et al., Estrogen receptors alpha (ERα) and beta (ERβ): subtype-selective ligands and clinical potential. Steroids, 2014. 90: p. 13-29.
27. Scarpin, K.M., et al., Progesterone action in human tissues: regulation by progesterone receptor (PR) isoform expression, nuclear positioning and coregulator expression. Nuclear receptor signaling, 2009. 7: p. e009-e009.
28. Spiering, M.J., On the trail of steroid aromatase: The work of Kenneth J. Ryan. Journal of Biological Chemistry, 2019. 294(28): p. 10743-10745.
29. Palmieri, C., et al., Breast cancer: current and future endocrine therapies. Molecular and cellular endocrinology, 2014. 382(1): p. 695-723.
30. Sainsbury, R., The development of endocrine therapy for women with breast cancer. Cancer treatment reviews, 2013. 39(5): p. 507-517.

31. Liehr, J.G., Is estradiol a genotoxic mutagenic carcinogen? Endocrine reviews, 2000. 21(1): p. 40-54.
32. Aka, J.A., et al., 17β-hydroxysteroid dehydrogenase Type 1 stimulates breast cancer by dihydrotestosterone inactivation in addition to estradiol production. Molecular endocrinology, 2010. 24(4): p. 832-845.
33. Delvoux, B., et al., Inhibition of Type 1 17β-Hydroxysteroid Dehydrogenase Impairs the Synthesis of 17β-Estradiol in Endometriosis Lesions. The Journal of Clinical Endocrinology & Metabolism, 2014. 99(1): p. 276-284.
34. Penning, T.M., Molecular endocrinology of hydroxysteroid dehydrogenases. Endocrine reviews, 1997. 18(3): p. 281-305.
35. Cavalli, A., et al., Enantioselective nonsteroidal aromatase inhibitors identified through a multidisciplinary medicinal chemistry approach. Journal of medicinal chemistry, 2005. 48(23): p. 7282-7289.
36. Borgne, M.L., et al., Synthesis and In Vitro Evaluation of 3-(1-Azolylmethyl)-1H-indoles and 3-(1-Azolyl-1-phenylmethyl)-1H-indoles as Inhibitors of P450 arom. Archiv der Pharmazie, 1997. 330(5): p. 141-145.
37. Jacobs, C., et al., 1-Imidazolyl (alkyl)-substituted di-and tetrahydroquinolines and analogues: syntheses and evaluation of dual inhibitors of thromboxane A2 synthase and aromatase. Journal of Medicinal Chemistry, 2000. 43(9): p. 1841-1851.
38. Tan, Z., et al., Theory of binless multi-state free energy estimation with applications to protein-ligand binding. The Journal of chemical physics, 2012. 136(14): p. 04B608.
39. Fortuné, A., Techniques de Modélisation Moléculaire appliquées à l'Etude et à l'Optimisation de Molécules Immunogènes et de Modulateurs de la Chimiorésistance. 2006, University of Joseph Fourier- Grenouble1.
40. Forli, W.E.H., et al., AutoDock Version 4.2, User Guide. 2012, Citeseer.
41. Traore, M.D.M., Synthèse et études de modélisation moléculaire dans l'optimisation de la sélectivité de nouveaux agents antiparasitaires inspirés de produits naturels. 2016, University Grenoble Alpes.
42. Morris, G.M., et al., AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of computational chemistry, 2009. 30(16): p. 2785-2791.
43. Systèmes, D., BIOVIA Discovery Studio Modeling Environment. Dassault Systèmes Biovia: San Diego, CA, USA, 2016.
44. Frisch, M., et al., Gaussian, Inc., Wallingford CT. Gaussian 09, 2009.
45. Darwin, K.H., et al., The proteasome of Mycobacterium tuberculosis is required for resistance to nitric oxide. Science, 2003. 302(5652): p. 1963-1966.
46. Bandyopadhyay, N., et al., Synthesis, structure, DFT calculations, electrochemistry, fluorescence, DNA binding and molecular docking aspects of a novel oxime based ligand and

its palladium (II) complex. Journal of Photochemistry and Photobiology B: Biology, 2016. 160: p. 336-346.

47. Guhathakurta, B., et al., Synthesis, characterization, structure, DNA binding aspects and molecular docking study of a novel Schiff base ligand and its bis (μ-chloro) bridged Cu (II) dimer. Polyhedron, 2017. 126: p. 195-204.
48. Salentin, S., et al., PLIP: fully automated protein–ligand interaction profiler. Nucleic acids research, 2015. 43(W1): p. W443-W447.
49. Brzozowski, A.M., et al., Molecular basis of agonism and antagonism in the oestrogen receptor. Nature, 1997. 389(6652): p. 753-758.
50. Williams, S.P. and P.B. Sigler, Atomic structure of progesterone complexed with its receptor. Nature, 1998. 393(6683): p. 392-396.
51. Tanenbaum, D.M., et al., Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains. Proceedings of the National Academy of Sciences, 1998. 95(11): p. 5998-6003.
52. Ghosh, D., et al., Structural basis for androgen specificity and oestrogen synthesis in human aromatase. Nature, 2009. 457(7226): p. 219-223.
53. Lin, S.X., et al., 3D-structure of human estrogenic 17β-HSD1: binding with various steroids. The Journal of Steroid Biochemistry and Molecular Biology, 1999. 69(1): p. 425-429.
54. Mazumdar, M., et al., Binary and ternary crystal structure analyses of a novel inhibitor with 17β-HSD type 1: a lead compound for breast cancer therapy. Biochemical Journal, 2009. 424(3): p. 357-366.



Chapter VI

Development of Quantitative structure–activity relationship (QSAR) model for the anti-proliferative effect against MCF-7 cell lines

IV.1. Introduction

Quantitative structure-activity relationship (QSAR) analysis is based on the general concept of medicinal chemistry that a ligand or compound's biological activity is related to its molecular structure or properties, and structurally similar molecules may have similar biological activities[1]. QSAR model defines mathematical relationships between descriptors and biological activities of known ligands to predict unknown ligands' activities. Nowadays, a large amount of experimental and predicted data about the structure of organic molecules and biomolecules is available. Advanced computational methods and high-performance computers enable us to get wide sets of descriptors. Therefore, it becomes easy to predict and classify biological activities of virtual or newly-synthesized compounds in several scientific studies including chemistry, biology, toxicology and drug discovery [2-5]. 1D, 2D, 3D and multidimensional QSAR approaches have been developed during the past several decades [6-8]. The major differences in these methods include the chemical descriptors and mathematical approaches that are used to establish the correlation between the target properties and the descriptors.

Many statistical methods have been employed to generate QSAR models from descriptive variables. Multiple linear regressions is one of the most successful techniques use by many researcher in construct of QSAR models [9, 10].

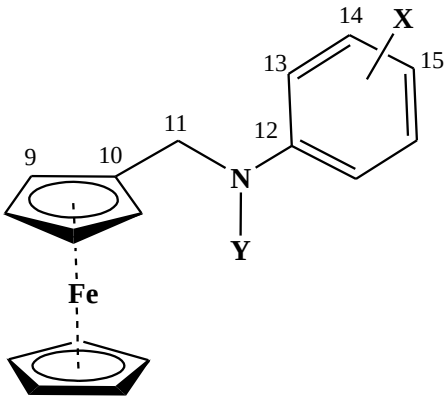
IV.2. Materials and methods

In this part, the anti-proliferative activity of 29 ferrocenylmethylaniline derivatives against MCF-7 cell lines was investigated to predict a QSAR model using molecular descriptors and MLR analysis.

IV.2.1. Training set

To develop QSAR a series of compounds called a training set was used. The chemical structure of the training set used in this study are presented in table IV.1 where X and Y are the substituting group. These compounds were optimized by DFT/B3LYP with 6-311G++(d,p) basis set using Gaussian 09.

Table. IV. 1. Chemical structures of the molecules under study. X and Y are the substituting groups



N°:	X	Y	N°:	X	Y	N°:	X	Y
A0	H	H	**A10**	4-Cl	H	**A20**	4-COMe	H
A1	H	COMe	**A11**	3-Br	H	**A21**	H	NHCOMe
A2	H	COEt	**A12**	4-Br	H	**A22**	H	NHCOPh
A3	H	COPh	**A13**	2-CN	H	**A23**	2-NO2	COMe
A4	H	CO-Ph-2-NO2	**A14**	3-CN	H	**A24**	3-NO2	COMe
A5	2-Me	H	**A15**	4-CN	H	**A25**	4-NO2	COMe
A6	3-Me	H	**A16**	2-NO2	H	**A26**	2-CN	COMe
A7	2-Et	H	**A17**	3-NO2	H	**A27**	3-CN	COMe
A8	2-Cl	H	**A18**	4-NO2	H	**A28**	4-CN	COMe
A9	3-Cl	H	**A19**	2-COMe	H			

IV.2.2. Biological data

In the present study a set of ferrocenylmethylaniline derivatives used as anti-proliferative agents against MCF-7 cell line. The data set have been prepared by converting the obtained IC_{50} values (M) (Chapter IV) to the logarithmic scale [pIC_{50}=-Log IC_{50}] to guarantee linearity, and to achieve normality.

IV.2.3. Molecular descriptors

The molecular descriptors have been calculated using Gaussian 09W[11] and HyperChem (8.03)[12] softwares and the website SwissADME[13, 14].

The QSAR properties module from HyperChem (8.03)[12] was used to calculate: molar polarizability (Pol), molar volume (MV) and molar weight (MW). The semi empirical PM3 optimized geometry was used to calculate the heat of formation (HF).

- **Molar polarizability (Pol): is** the ease with which an external electric field deforms the electron cloud. A molecule's polarizability can improve its aqueous solubility. This feature plays a very important role in many molecular properties modeling [15].
- **Molar volume (MV):** bounded by solvent accessible surfaces, using a grid method[16].
- **Molar weight (MW):** is an essential criterion in the passive diffusion of drugs across the tightly packed aliphatic side chains of the bilayer membrane. It is related to the size

of the molecule, as molecular size increases, a larger cavity must be formed in water in order to solubilize the compound [17]

- **Heat of formation (HF):** The heat of Formation is known as the change in enthalpy accompanying the formation of one mole of a compound from its elements in their natural and stable states, under standard conditions at a given temperature of one atmosphere[18].

The Quantum Chemical descriptors: dipole moment (μ), Energy of frontier orbital's HOMO, LUMO, energetic gap (GapE), atomic net charges (qFe1, qN1, qC9, qC10, qC11, qC12, qC13, qC14 and qC15) and total energy (E) were computed using Gaussian 09W software [11] by using DFT/B3LYP with 6-311G++(d,p) basis set.

- **Dipole moment (μ):** can be measured from the partial atomic charges. Probably the most basic experimental measure of the charge density in a molecule is the molecular dipole moment. The accuracy of the overall distribution of electrons in a molecule is difficult to quantify, since it involves all of the multipole moments. Experimental measures of accuracy are required to determine the findings [16].
- **Energy of frontier orbital's HOMO, LUMO:** The energies of the frontier orbitals HOMO (highest occupied molecular orbital) and LUMO (lowest unoccupied molecular orbital) are commonly used descriptors in QSAR analysis. They reflect the reactivity of a molecule. A higher HOMO energy suggests higher affinity of a molecule to react as a nucleophile, a lower LUMO energy suggests stronger electrophilic nature of a molecule. Additionally, electrophilic and nucleophilic attack will most likely occur at atoms where the coefficients of the corresponding atomic orbitals in HOMO and LUMO, respectively, are large [19].
- **Energetic gap (GapE):** translates the energy between the highest occupied and lowest vacant molecular orbital, it's an important stability index. This energy difference is a measure of the excitability of a molecule. Thus, the narrower the energy range, the more the molecule will be able to interact with the environment.

  A large HOMO-LUMO deviation implies high stability for the molecule in the sense of its low reactivity in chemical reactions, and similarly, a small deviation implies high reactivity of the molecule. [20, 21].
- **Atomic net charges:** It plays an essential role in determining the atoms' vibrational properties. Indeed, it figures out which atom was electron donor and which one was

electron acceptor. So it can determine the dipole moment and polarizability of the molecule [22, 23]

- **Total energy (E):** For a molecule isolated in the ground state, the calculated total energy can be used as a quantum molecular descriptor. This approximate energy has been calculated for an optimized conformation of the most stable geometry with minimal energy structure[24]

Moreover, octanol/water partition coefficient (log P), the molar refractivity (MR), Topological polar surface area (TPSA) and water solubility (logS) were obtained using the free web-server SwissADME [13, 14] by entering the SMILES (Simplified Molecular Input Line Entry System).

- **Octanol/water Partition coefficient (logP):** Lipophilicity is an important property that has a significant impact on the solubility, absorption, distribution, metabolism and excretion of drugs. The partition coefficient P, defined as the ratio of molar concentration of a chemical dissolved at equilibrium in octanol phase to its molar concentration in aqueous phase [25-27].
- **Molar refractivity (MR):** The molar refractivity (MR) is a steric parameter which depends on the aromatic ring spatial array in the synthesized compounds [28]. The refractivity is a special case of the molecular volume, it is a refractivity reduced to a quantity of material, as well as the refractivity is the quality of the refractive power of a body, it is used in the radio-electricity, biology, etc. In chemistry, the molecular refractivity is an important criterion to measure the steric factor, it is important in the case where the substitute has π electrons or lone pairs [15, 29].
- **Topological polar surface area (TPSA**): is a very useful parameter for predicting drug transport properties. It is defined as the sum of the surface areas of the polar atoms (usually oxygen, nitrogen, sulphur, chlorine and hydrogen attached) in a molecule [30].
- **Water solubility (log S):** Aqueous solubility of a chemical is defined as the amount of solute dissolved under equilibrium conditions in a saturated aqueous solution. Solubility is not only a property of interest to many academic research areas, but also a key parameter in the pharmaceutical industry and drug design. Solubility can have a significant effect on intestinal absorption. [31]

IV.2.4. Multiple Linear Regressions

Multiple linear regression (MLR) is one of the most popular modeling methods due to its simplicity of use and ease of interpretation. The important advantage of multiple linear

regression is that it is very transparent, since the algorithm is available, and predictions can be made easily.

The MLR method is based on the assumption that the property Y depends linearly on the different variables (the descriptors) X_1, $X_2 \ldots X_I$, according to the equation:

$$y = a_0 + \sum_{i=1}^{n} a_i x_i \qquad (Eq\ IV.1)$$

Where: Y is the dependent variable, Xi are the independent variables (descriptor), n is the number of explanatory variables, a_0 is the constant in the model equation; a_i are the descriptor coefficients in the model equation.

The size of these coefficients indicates the degree of influence of the corresponding molecular descriptors on the target activity. A positive coefficient indicates that the corresponding molecular descriptor contributes positively to the target, while a negative coefficient suggests the negative contribution[32].

Multiple linear regression analysis was carried out using the stepwise strategy in SPSS 22 software for Windows[33] . Our regression analysis generates an equation for explaining the statistical relationship between the predictor variables (descriptors) and the response variable (bioactivity), as well as the statistical parameters such as correlation coefficient, squared correlation coefficient, standard error of estimate and F- test values. These parameters explain the significance of the obtained equation.

The Stepwise Method Search selects a model based on their statistical significance by inserting or removing individual descriptors, one step at a time. The final result of this method is a single regression model, making it nice and simple.

The p-value for each term checks the null hypothesis that the coefficient is equal to zero (no effect). A low p-value (< 0.05) indicates that the null hypothesis is rejectable. In other words, a descriptor with a low p-value is likely to be a meaningful addition to the model because changes in the value of the descriptor are linked to changes in the bioactivity.

Conversely, a larger (insignificant) p-value indicates that changes in the descriptor are not correlated with bioactivity changes.

In our result, we found that the nine equation descriptors are statistically important, since their p-values are less than 0.05.

IV.2.5. QSAR model validation: leave –one–out (LOO) cross–validation method

The final step in QSAR model development is model validation. In order to test the validity of the predictive power of selected MLR model, as opposed to traditional regression methods,

the cross-validation method estimates a model's trustworthiness. This method is a special case of "k-packet" cross-validation where k=n. i.e. one takes (n-1) observations to build the QSAR model, then validates it on the n^{th} observation and repeats this operation n times so that in the end each observation has been used exactly once as a validation set. The predictive ability of the model is quantified in terms of the corresponding leave-one-out cross-validated parameters, PRESS (predicted residual sum squares), SSY (sum of the squares of response value), R^2_{cv} (overall predictive ability), adjusted R^2 and Q^2, S_{PRESS} (uncertainty of prediction), and PSE (predictive square error). See equations IV.2 – IV.7.

IV.2.6. Statistical Parameters

- **Fischer Statistic (F):** is the ratio of two mean squares. The null hypothesis can be rejected when the F value is large and the significance level small. In other words, F is a statistical parameter used to evaluate the null hypothesis that inclusion of an additional variable does not result in a significant increase in R2[34].
- **Correlation Coefficient (R):** R is such that -1 < R < +1. The + and – signs are used for positive linear correlations and negative linear correlations, respectively. If the predicted and observed values have a strong linear correlation r is close to 1, however if there is no linear correlation or a weak linear correlation r is close to 0. The value of the correlation coefficient can be strongly influenced by one outlying point.
- **Coefficient of Determination (R^2) :** is such that 0 < R2 < 1, and the stronger the correlation (R is closer to 1). It is found by squaring the correlation coefficient and is used as a more precise way to interpret the correlation coefficient. [35]
- **Standard Error of the Estimate (SE):** is a measure of how much the value of a test statistic varies from sample to sample. It is the standard deviation of a statistical's sampling distribution.
- **Factor of the quality Q:** it suggests the power of prediction. Positive value for QSAR model suggests its high predictive power and lack of over fitting[36].
- **PRESS:** The predicted residual sum of squares is the difference between the observed value and the model's expected value.
- **SSY:** is the sum of the squares of the distances found for a variable compared to the average of this variable, the sum of squares allows the overall difference in a variable to be measured
- **SPRESS:** the predictive ability of the models is evaluated by the root mean square error. PRESS is smaller than SSY, the model can be considered statistically significant and

predicts better than chance. The ratio **PRESS/SSY** can be used to measure approximate confidence intervals for predicting new observations (compounds) [34, 37]..

- **R^2 cv (Q^2):** Change in R^2 statistic created by adding or deleting an independent variable. When the R^2 change associated with a variable is high, that means the variable is a strong predictor of the dependent variable. the R^2_{cv} (or Q2) values should be >0.6. The model is considered to be excellent, if R^2_{cv} (or q2) is >0.9.
- **R^2adj:** Adjusted R^2, The R2 sample attempts to improve the calculation of how well these models match the population. The model does not generally match the population, as it fits the sample from which it is derived. Adjusted R2 attempts to correct R2 to represent more closely the goodness of model fit within the population[38].
- **PSE:** predictive square error, is used to determine the predictive power of the models proposed.

These parameters are defined in the following equations:

$$PSE = \sqrt{\frac{PRESS}{n}} \quad .....(EqVI.2)$$

$$S_{PRESS} = \sqrt{\frac{PRESS}{n-p-1}} \quad .....(Eq.VI.3)$$

$$R^2_{CV} = 1 - \frac{PRESS}{SSY} = Q^2 \quad .....(Eq.VI.4)$$

$$S = \sqrt{\frac{SSY}{n-1}} \quad .....(Eq.VI.5)$$

$$R^2_{adj} = (1-r^2)\left(\frac{n-1}{n-p-1}\right) \quad ....(\ Eq.VI.6)$$

$$Q = \frac{R}{SE} \quad .....(Eq.VI.7)$$

Where: n is the number of compounds used, p is the number of parameters (descriptors) used in the model.

IV.3. Results and discussion

Table IV.2. shows the values of the calculated descriptors used in order to establish the 2D-QSAR model. After multiple regression analysis using SPSS statistics 22 software [33] the best statistically significant correlations generated along with relevant statistical parameters are given in the model below, the model was generated after removing 4 compounds (A7, A18, A1 and A20) which behaved as outliers with high residual (0.32, 0.29, 0.32 and 0.28, respectively) comparing to the standard error of estimation.

pIC50= -13.85 - 2.179 HOMO +0.0003 E +0.107 MR + 0.022 MV + 0.005 HF - 0.617 Pol - 0.029 TPSA - 0.836 qC11 – 0.829 qN1

n=25, R=0.92, R^2= 0.84, SEE=0.14, F=9.33, Q= 6.75, S= 0.302.

Table. IV. 2: Values of molecular descriptors used in the multiple regression analysis

Comps	pIC50	HOMO (eV)	LUMO (eV)	ΔE (eV)	Total E (a.u)	μ (Debay)	MR ($Å^3$)	log P	log S	MV ($Å^3$)	HF (kcal/mol)	MW (amu)
A0	4.298	-5.442	-0.49	4.952	-1976.69	1.796	78.79	3.03	-4.76	798.72	-148.52	291.18
A1	4.603	-5.469	-0.925	4.544	-2129.40	3.87	88.7	2.85	-4.28	894.27	-188.402	333.21
A2	4.747	-5.687	-0.626	5.061	-2168.71	3.8103	93.51	3.15	-4.77	941.97	-191.884	347.24
A3	4.514	-5.442	-1.252	4.191	-2321.17	3.4562	108.6	3.9	-6.01	1047.36	-152.411	395.28
A4	4.555	-5.448	-2.765	2.683	-2525.73	6.1846	117.43	3.19	-6.79	1100.54	-156.861	440.2
A5	4.328	-5.496	-0.462	5.034	-2016.02	1.229	83.76	3.3	-5.14	841.92	-155.7	305.2
a6	4.478	-5.388	-0.489	4.898	-2016.02	1.606	83.76	3.3	-5.14	848.88	-157.865	305.2
A7	4.418	-5.496	-0.462	5.034	-2055.34	1.222	88.57	3.56	-5.58	890.31	-158.701	319.23
A8	4.201	-5.586	-0.582	5.004	-2436.32	2.138	83.8	3.5	-5.41	837.18	-155.037	325.62
A9	4.850	-5.632	-0.634	4.998	-2436.32	3.418	83.8	3.5	-5.41	842.77	-155.491	325.62
A10	4.465	-5.551	-0.653	4.898	-2436.32	3.874	83.8	3.5	-5.41	842.09	-155.44	325.62
A11	4.327	-5.633	-0.653	4.979	-4550.24	3.48	86.49	3.57	-5.48	860.72	-141.27	370.07
A12	4.144	-5.551	-0.68	4.871	-4550.24	3.96	86.49	3.57	-5.48	861.34	-141.25	370.07
A13	4.431	-5.769	-1.415	4.354	-2068.97	4.4866	83.51	2.89	-5.54	857.83	-111.68	316.19
A14	4.852	-5.714	-1.524	4.191	-2068.97	6.0922	83.51	2.78	-4.97	855.64	-113.73	316.19
A15	4.569	-5.905	-1.143	4.762	-2068.97	8.2504	83.51	2.78	-4.97	851.21	-118.746	316.19
A16	3.696	-5.606	-2.64	2.966	-2181.27	5.1478	87.62	2.56	-6.12	837.99	-159.441	336.17
A17	4.714	-5.823	-2.612	3.211	-2181.27	6.7529	87.62	2.45	-5.55	866.96	-156.37	336.17
A18	4.253	-5.987	-2.259	3.728	-2181.27	9.4489	87.62	2.45	-5.55	865.39	-158.208	336.17
A20	4.327	-5.741	-1.279	4.462	-2129.38	6.326	88.99	2.96	-4.79	910.78	-189.498	333.21
A19	4.389	-5.632	-1.768	3.864	-2129.38	5.052	88.99	3	-5	900.48	-187.028	333.21
A21	4.069	-5.742	-0.653	5.089	-2184.74	3.9532	91.5	2.64	-4.82	926.6	-164.587	348.23
A22	4.365	-5.66	-1.252	4.408	-2376.52	3.4243	96.31	2.93	-5.32	1072.71	-128.642	410.3
A23	3.644	-5.578	-2.912	2.667	-2333.96	6.3106	97.52	2.15	-5.07	944.33	-194.722	378.21
A24	4.292	-5.959	-2.721	3.238	-2333.96	8.8934	97.52	2.15	-5.07	930.05	-201.07	378.21
A25	4.386	-5.687	-3.075	2.612	-2333.96	2.1933	97.52	2.15	-5.07	953.21	-196.525	378.21
A26	4.351	-5.551	-2.095	3.456	-2221.67	6.395	93.42	2.63	-4.49	943.26	-151.403	358.22
A27	4.799	-5.606	-2.095	3.51	-2221.67	4.0356	93.42	2.63	-4.49	949.31	-152.688	358.22
A28	4.504	-5.608	-1.983	3.624	-2221.68	4.9661	93.42	2.63	-4.49	949.61	-152.842	358.22

Table. IV. 2. Continued.

Comp	Pol (Å^3)	TPSA (Å^2)	qFe (C)	qC9 (C)	qC10 (C)	qC11 (C)	qC12 (C)	qC13 (C)	qC14 (C)	qC15 (C)	qN1 (C)
A0	30.25	12.03	0.561	-0.095	0.662	-1.14	-0.324	-0.124	-0.272	-0.373	0.206
A1	34.01	20.31	0.736	-0.455	0.913	-1.132	-0.66	0.152	-0.278	-0.382	0.55
A2	35.84	20.31	0.548	-0.4	-0.316	-0.539	-0.745	-0.465	-0.129	-0.29	0.574
A3	41.83	20.31	0.454	-0.167	0.749	-0.8	-0.281	-0.381	-0.397	0.018	0.546
A4	43.55	66.13	0.457	-0.392	0.842	-0.895	-0.171	0.073	-0.402	-0.274	0.521
A5	32.09	12.03	0.639	-0.18	0.345	-0.929	-0.313	0.355	-0.081	-0.409	0.184
A6	32.09	12.03	0.59	-0.099	0.736	-1.189	-0.46	0.396	0.603	-0.749	0.219
A7	33.92	12.03	0.629	-0.264	0.563	-0.975	-0.062	0.572	-0.216	-0.509	0.185
A8	32.18	12.03	0.555	-0.199	0.692	-1.085	-1.39	1.111	-0.198	-0.524	0.303
A9	32.18	12.03	0.544	-0.137	0.668	-1.087	-0.661	-0.339	0.527	-0.717	0.205
A10	32.18	12.03	0.551	-0.155	0.682	-1.121	-0.646	-0.24	-0.649	0.554	0.218
A11	32.88	12.03	0.559	-0.122	0.682	-1.126	-0.248	-0.339	0.091	-0.218	0.199
A12	32.88	12.03	0.556	-0.115	0.645	-1.115	-0.464	-0.224	-0.253	0.081	0.211
A13	32.1	35.82	0.868	-0.324	-0.22	-0.451	-0.579	1.479	0.151	-0.367	0.108
A14	32.1	35.82	0.593	-0.368	-0.832	-1.267	-0.407	0.849	1.339	-0.703	0.192
A15	32.1	35.82	0.527	-0.2	-0.026	-0.304	-0.058	-0.069	-0.259	1.84	0.078
A16	31.97	57.85	0.429	-0.176	0.731	-0.998	-0.257	-0.032	0.066	-0.395	0.093
A17	31.97	57.85	0.71	-0.355	-0.185	-0.707	-0.455	0.338	0.203	0.286	0.121
A18	31.97	57.85	0.821	-0.373	-0.152	-0.468	-0.066	0.466	-0.002	-0.451	0.066
A20	34.01	29.1	0.842	-0.417	-0.116	-0.553	-0.404	0.703	-0.005	-0.498	0.235
A19	34.01	29.1	0.778	-0.405	-0.029	-0.535	-0.405	-0.766	-1.217	1.596	0.102
A21	35.36	32.34	0.802	-0.385	0.361	-0.905	-0.584	0.007	-0.086	-0.301	0.901
A22	43.18	32.34	0.539	-0.423	0.768	-0.936	-0.112	-0.433	-0.226	0.074	-0.573
A23	35.72	66.13	0.696	-0.31	-0.772	-0.986	-0.199	-0.199	0.071	-0.429	0.855
A24	35.72	66.13	0.784	-0.302	0.893	-0.85	-0.035	-0.543	-0.217	0.104	0.536
A25	35.72	66.13	-0.146	-0.36	0.468	-0.541	0.465	-0.305	-0.347	0.367	-0.244
A26	35.86	44.1	0.704	-0.198	1.048	-1.302	-0.954	1.198	-0.078	-0.374	0.703
A27	35.86	44.1	0.79	-0.47	-0.913	-1.26	-0.788	-0.383	1.445	-0.316	0.591
A28	35.86	44.1	0.651	-0.419	1.003	-1.355	-0.821	0.134	-0.277	1.749	0.665

It can be seen that the obtained model shows R= 0.92 and R^2= 0.84, which allowed us to indicate firmly the correlation between the descriptors (HOMO, E, MR, MV, HF, POL, TPSA, qC11 and qN1) and the anti-proliferative activity against MCF-7 cell line. The calculated F value of the obtained model has been found greater than the tabulated one by a wide margin as required for a meaningful regression. The positive value of Q equal to 6.75 for QSAR's model suggests its high predictive power. Moreover, the accuracy in the analysis is indicated by the low standard error of estimate equal to 0.14. All the values of t-statistics are significant which confirms the significance of each descriptor.

It can be seen from the equation that the negative coefficients of HOMO, Pol, TPSA, qC11 and qN1 explain that any increase in these descriptors causes a decrease in the biological activity. Moreover, the positive coefficients of E, MR, MV and HF indicates that any increase in these parameters causes an increase in biological activity.

The correlation matrix between the descriptors obtained by the MLR analysis and the biological activity is reported in Table IV.3, where the descriptors obtained afterwards are

correlated with each other and with the biological activity. Table IV.3 shows that the heat of formation and TPSA are important parameters in the correlation between the selected descriptor and the anti-proliferative activity against MCF-7 cell lines.

Table. IV. 3. Correlation Matrix for the obtained Model

	pIC50	HOMO	E	MR	MV	HF	pol	TPSA	qC11	qN1
pIC50	1.000									
HOMO	-0.125	1.000								
E	0.168	-0.086	1.000							
MR	-0.018	0.135	0.01	1.000						
MV	0.076	0.102	0.053	0.926	1.000					
HF	0.283	0.008	-0.117	-0.351	-0.191	1.000				
pol	0.051	0.155	0.019	0.906	0.988	-0.14	1.000			
TPSA	-0.204	-0.393	0.264	0.55	0.442	-0.344	0.35	1.000		
qC11	0.053	-0.502	0.214	0.113	0.095	-0.146	0.066	0.206	1.000	
qN1	-0.134	0.108	0.068	0.303	0.145	-0.325	0.098	0.151	-0.285	1.000

The generated model was validated for predictive ability inside the model using Leave One Out method (LOO). From the results depicted in table IV.4, the model is statistically significant. The PRESS value is lower than SSY which points out that the model predicts better than chance and can be considered statically significant. Furthermore, The PRESS/SSY ratio is equal to 0.151 which should be lower than 0.4[39] . The high R^2_{cv} and low S_{PRESS} (equal to 0.848 and 0.148, respectively) reflect its good predictive potential. Also, the high value of R^2_{adj} is important criteria for our QSAR model to have the best qualification.

The predictive error of the coefficient correlation PE is another parameter used to evaluate the predictive power of the obtained model. If R < PE, then correlation is not significant; if R > PE, several times (at least three times), then correlation is indicated; if R > 6 PE, then the correlation is definitely good [40]. For the developed model the condition R > 6 PE is satisfied and hence it has a good predictive power.

Table IV. 4. Cross-validation parameters.

R^2_{adj}	R^2_{cv}	PE	6PE	SSY	PRESS	S_{PRESS}	PRESS/SSY
0.758	0.848	0.021	0.129	2.191	0.332	0.148	0.151

Table. IV. 5. Observed vs. calculated activity of N-ferrocenylmethylaniline derivatives.

Compounds	pIC50Obs	pIC50Pred	Residuals
A0	4.298	4.154	0.144
A2	4.747	4.586	0.161
A3	4.514	4.671	-0.157
A4	4.555	4.386	0.169
A5	4.328	4.400	-0.072
A6	4.478	4.492	-0.014
A8	4.201	4.343	-0.142
A9	4.850	4.645	0.205
A10	4.465	4.471	-0.006
A11	4.327	4.332	-0.005
A12	4.144	4.148	-0.004
A13	4.431	4.487	-0.056
A14	4.852	4.921	-0.069
A15	4.569	4.504	0.065
A16	3.696	3.752	-0.056
A17	4.714	4.604	0.110
A19	4.389	4.376	0.013
A21	4.069	4.273	-0.204
A22	4.365	4.340	0.026
A23	3.644	3.626	0.018
A24	4.292	4.264	0.028
A25	4.386	4.587	-0.201
A26	4.351	4.325	0.026
A27	4.799	4.627	0.172
A28	4.504	4.655	-0.151

Figure IV.1 shows the plot of linear regression of predicted versus experimental values of the biological activity. The plot for the obtained model shows much correlation with experimentally reported data having $R^2 = 0.848$. The plot of residuals of predicted values of the biological activity pIC_{50} against the experimental values does not show any systematic error (figure IV.2) positive and negative residuals are randomly distributed.

Thus, the current QSAR research shows that in the generation of these molecules, this model can be successfully applied to predict anti-proliferative activity against MCF-7 cell line.

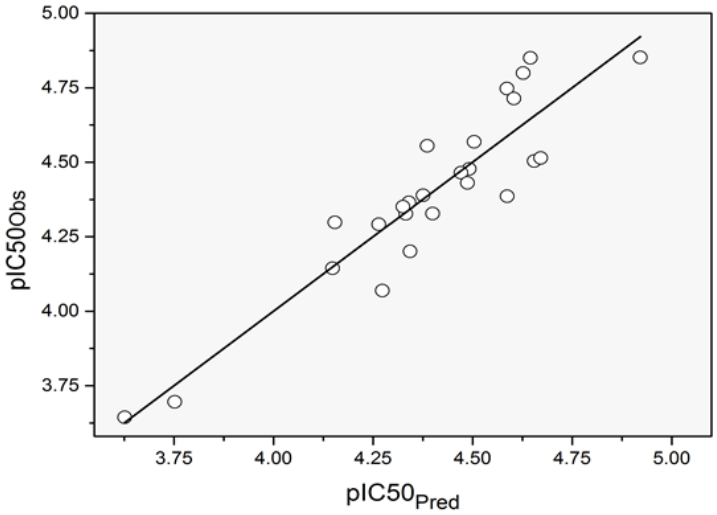


Figure. IV. 1. Scatter Plot between the Observed and Predicted Activities of the Model.

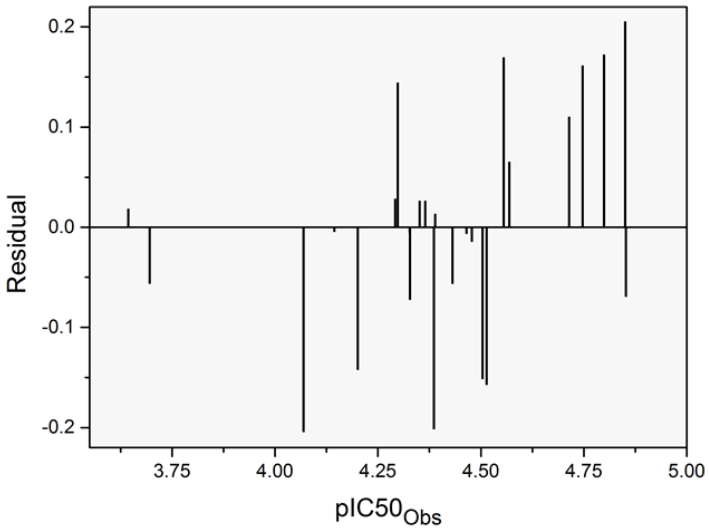


Figure. IV. 2 . Plots of the residual values against the experimentally observe.

IV. 4. Conclusion

QSAR study has been performed. A significant regression equation was obtained by the MLR method with respect to their experimental biological activities. The best equation of regression contains HOMO, E, MR, MV, HF, POL, TPSA, qC11 and qN1 descriptors. QSAR model indicates that these descriptors have significant relationships with the observed bioactivities.

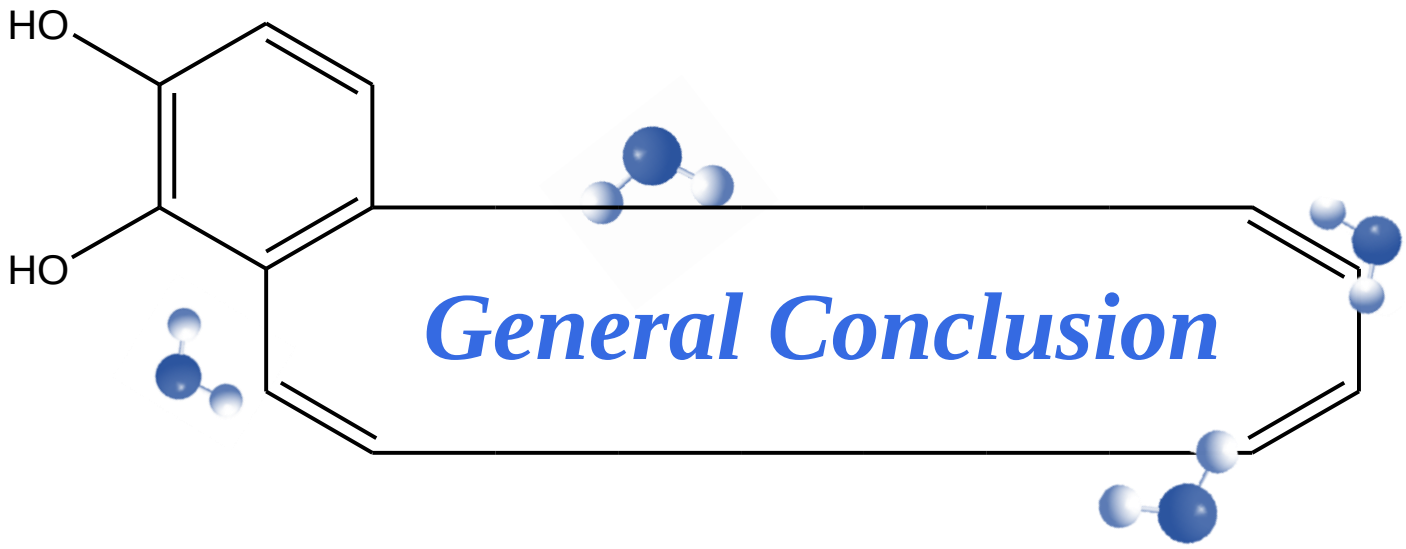
This study will help the researchers to predict a better anti-proliferative activity of ferrocenylmethylanilines derivatives before the experimental tests.

References

1. Esposito, E.X., A.J. Hopfinger, and J.D. Madura, Methods for applying the quantitative structure-activity relationship paradigm, in Chemoinformatics. 2004, Springer. p. 131-213.
2. Bradbury, S.P., Quantitative structure-activity relationships and ecological risk assessment: an overview of predictive aquatic toxicology research. Toxicol Lett, 1995. **79**(1-3): p. 229-37.
3. Hansen, C., B.R. Telzer, and L. Zhang, Comparative QSAR in toxicology: examples from teratology and cancer chemotherapy of aniline mustards. Critical reviews in toxicology, 1995. **25**(1): p. 67-89.
4. Perkins, R., et al., Quantitative structure-activity relationship methods: Perspectives on drug discovery and toxicology. Environmental Toxicology and Chemistry: An International Journal, 2003. **22**(8): p. 1666-1679.
5. Chen, J.-Z., et al., 3D-QSAR studies of arylpyrazole antagonists of cannabinoid receptor subtypes CB1 and CB2. A combined NMR and CoMFA approach. Journal of medicinal chemistry, 2006. **49**(2): p. 625-636.
6. Lill, M.A., Multi-dimensional QSAR in drug discovery. Drug Discovery Today, 2007. **12**(23-24): p. 1013-1017.
7. Verma, J., V.M. Khedkar, and E.C. Coutinho, 3D-QSAR in drug design-a review. Current topics in medicinal chemistry, 2010. **10**(1): p. 95-115.
8. Fischer, P.M., Computational chemistry approaches to drug discovery in signal transduction. Biotechnology Journal: Healthcare Nutrition Technology, 2008. **3**(4): p. 452-470.
9. Zahouily, M., et al., QSAR for anti-malarial activity of 2-aziridinyl and 2, 3-bis (aziridinyl)-1, 4-naphthoquinonyl sulfonate and acylate derivatives. Journal of molecular modeling, 2006. **12**(4): p. 398-405.
10. Liang, G., et al., Study on quantitative structure-activity relationship by 3D holographic vector of atomic interaction field. Acta Physico-Chimica Sinica, 2006. **22**(3): p. 388-390.
11. Frisch, M., et al., Gaussian, Inc., Wallingford CT. Gaussian 09, 2009.
12. Teppen, B.J., HyperChem, release 2: molecular modeling for the personal computer. Journal of chemical information and computer sciences, 1992. **32**(6): p. 757-759.
13. Daina, A., O. Michielin, and V. Zoete, iLOGP: a simple, robust, and efficient description of n-octanol/water partition coefficient for drug design using the GB/SA approach. Journal of chemical information and modeling, 2014. **54**(12): p. 3284-3301.

14. Daina, A., O. Michielin, and V. Zoete, SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Scientific reports, 2017. **7**: p. 42717.
15. Belaidi, S., N. Melkemi, and D. Bouzidi, Molecular geometry and structure-property relationships for 1, 2-dithiole-3-thione derivatives. Int. J. Chem. Res, 2012. **4**: p. 134-139.
16. Imane, B., Study by Quantum and Empirical Methods of SAR and QSAR of Heterocyclic with Medicinal Interest. 2017, University of Mohamed Khider BISKRA.
17. Hirshfeld, F., Electron density distributions in molecules. Crystallography reviews, 1991. **2**(4): p. 169-200.
18. Aicha, K., Etude par la modélisation moléculaire des relations structures-propriétés de quelques séries hétérocycliques bioactives. 2015, university of biskra.
19. Lessigiarska, I., Development of structure-activity relationships for pharmacotoxicological endpoints relevant to European Union legislation. 2006, Liverpool John Moores University.
20. Lewis, D., C. Ioannides, and D. Parke, Interaction of a series of nitriles with the alcohol-inducible isoform of P450: computer analysis of structure—activity relationships. Xenobiotica, 1994. **24**(5): p. 401-408.
21. Zhou, Z. and R.G. Parr, Activation hardness: new index for describing the orientation of electrophilic aromatic substitution. Journal of the American Chemical Society, 1990. **112**(15): p. 5720-5724.
22. Ghosh, K., et al., Three mononuclear octahedral cobalt (III) complexes with salicylaldimine Schiff bases: Synthesis, characterization, phenoxazinone synthase mimicking activity and DFT study on supramolecular interactions. Polyhedron, 2016. **112**: p. 6-17.
23. Tomberg, A., GAUSSIAN 09W TUTORIAL AN INTRODUCTION TO COMPUTATIONAL CHEMISTRY USING G09W AND AVOGADRO SOFTWARE.
24. CHTITA, S., Modélisation de molécules organiques hétérocycliques biologiquement actives par des méthodes QSAR/QSPR. Recherche de nouveaux médicaments. 2017, university Moulay Ismail.
25. Luque, F., et al., SCRF calculation of the effect of water on the topology of the molecular electrostatic potential. The Journal of Physical Chemistry, 1993. **97**(37): p. 9380-9384.
26. Murray, J.S. and K. Sen, Molecular electrostatic potentials: concepts and applications. Vol. 3. 1996: Elsevier.

27. Pathak, R.K. and S.R. Gadre, Maximal and minimal characteristics of molecular electrostatic potentials. The Journal of chemical physics, 1990. **93**(3): p. 1770-1773.
28. Bazooyar, F., et al., Molecular modelling of cellulose dissolution. Journal of Computational and Theoretical Nanoscience, 2013. **10**(11): p. 2639-2646.
29. Salah, T., et al., Molecular Geometry, Electronic Properties, MPO Methods and Structure Activity/Property Relationship Studies of 1, 3, 4-Thiadiazole Derivatives by Theoretical Calculations. Reviews in Theoretical Science, 2015. **3**(4): p. 355-364.
30. Saunders, R.A. and J.A. Platts, Scaled polar surface area descriptors: development and application to three sets of partition coefficients. New Journal of Chemistry, 2004. **28**(1): p. 166-172.
31. Du-Cuny, L., et al., Computational aqueous solubility prediction for drug-like compounds in congeneric series. European Journal of Medicinal Chemistry, 2008. **43**(3): p. 501-512.
32. Roy, K., S. Kar, and R.N. Das, Understanding the basics of QSAR for applications in pharmaceutical sciences and risk assessment. 2015: Academic press.
33. Nie, N.H., D.H. Bent, and C.H. Hull, SPSS: Statistical package for the social sciences. Vol. 227. 1975: McGraw-Hill New York.
34. Podunavac-Kuzmanović, S.O., D.D. Cvetković, and D.J. Barna, QSAR analysis of 2-amino or 2-methyl-1-substituted benzimidazoles against Pseudomonas aeruginosa. International journal of molecular sciences, 2009. **10**(4): p. 1670-1682.
35. Rosner, B., Fundamentals of biostatistics. 2015: Nelson Education.
36. Verma, R.P. and C. Hansch, QSAR modeling of taxane analogues against colon cancer. European journal of medicinal chemistry, 2010. **45**(4): p. 1470-1477.
37. Purkayastha, S., et al., Possible antineoplastic agents: Part XIII. Synthesis, biological evaluation and QSAR studies of some 1-(substituted benzenesulphonyl)-5-oxopyrrolidine-2-carboxylic acid derivatives. Anti-cancer drug design, 1993. **8**(2): p. 95-100.
38. SALAH, T., Etude par la chimie computationnelle des corrélations 2D-QSAR et 3D-QSAR de quelques composés bioactifs. 2017, University Mohammed khidher of Biskra. p. 130.
39. Podunavac-Kuzmanović, S.O., D.D. Cvetković, and D.J. Barna, QSAR analysis of 2-amino or 2-methyl-1-substituted benzimidazoles against Pseudomonas aeruginosa. Int J Mol Sci, 2009. **10**(4): p. 1670-82.
40. Srivastava, A. and N. Shukla, Quantitative structure activity relationship (QSAR) studies on a series of imidazole derivatives as novel ORL1 receptor antagonists. Journal of Saudi Chemical Society, 2013. **17**(3): p. 321-328.



General Conclusion

General conclusion

In the present work, the structures of 29 ferrocene derivatives of type ferrocenylmethylaniline were optimised with Gaussian 09 software using the theoretical methods of DFT/B3LYP with 6-311G++(d,p) and LanL2DZ basis sets. The optimised structures were validated by comparison of their bond lengths and angles of their constituent atoms with the experimental Xray data. A good agreement between calculated and experimental parameters was found which made these compounds suitable for *in silico* study. Moreover, a vibrational analysis has been studied and functional groups were determined using both experimental and theoretical methods. Computational ^{1}H and ^{13}C chemical shift values were calculated and compared with experimental data, showing good agreement for both ^{1}H and ^{13}C NMR. Based on HOMO and LUMO energy gaps, the compounds A23 and A25 were predicted to be the most reactive. The MESP at the 6-311G++ (d, p) optimized geometry was calculated to predict the reactive sites for electrophilic and nucleophilic attack for all the studied derivatives.

Additionally, all derivatives showed an important cytotoxicity effect against MCF-7 cell lines, which was validated using molecular docking between the compounds and the ERα and PR.

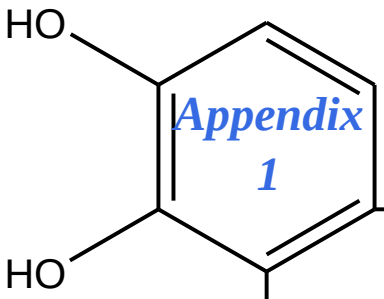
These results indicate that the studied derivates could be drug candidate against MCF7 cells, however, its mechanism of action is not clear yet since the docking study showed high affinity with both ERα and PR, it might be mediated by ERα, ERβ or PR. Thus, this work should be further followed by in vitro studies to better investigate their inhibitory effect against ERs and PR, also other MTT assays with prostate cancer cells (PC-3, devoid of ERα but with ERβ) and healthy cell line L929 should be carried out to evaluate the selectivity of compounds on cancer cells over normal cells.

Furthermore, docking study against aromatase shows that compounds A4 and A25 were the most active with the lowest binding free energies and more than three hydrogen bonds compared to other compounds and ASD. However, moderate affinity was predicted against 17β-HSD1 for all compounds compared to E2B. This indicate that these derivatives might be strong drug inhibitors of aromatase, which should also be followed by further *in vitro* study to investigate their inhibitory effect against this enzyme, especially for the molecules A4 and A25 that have higher affinity than ASD.

Finally, QSAR study has been performed, and a significant regression equation was obtained by the MLR method with respect to their experimental biological activities. The best equation of regression contains HOMO, E, MR, MV, HF, POL, TPSA, qC11 and qN1

descriptors. The QSAR model indicates that these descriptors have significant relationships with the observed bioactivity. The model was validated for predictive ability inside the model using Leave One Out method (LOO), which proves that the model can be successfully applied to predict anti-proliferative activity against MCF-7 cell line.

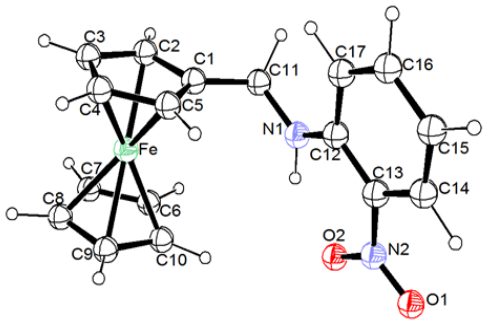
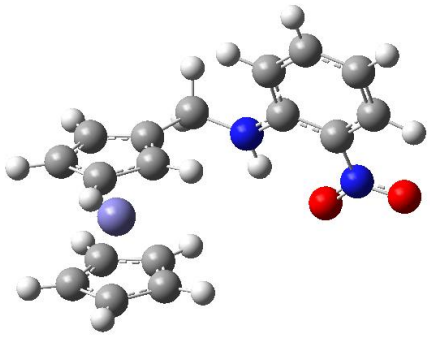
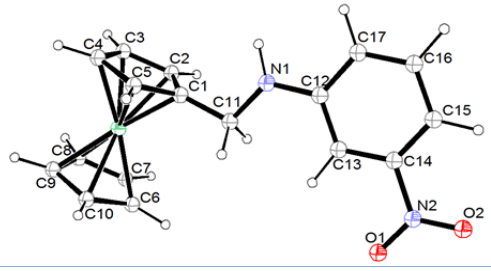
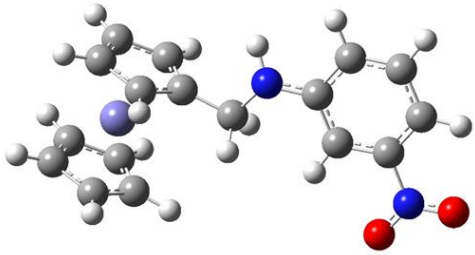
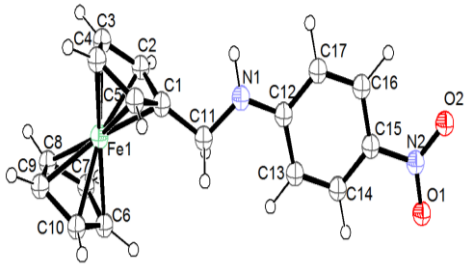
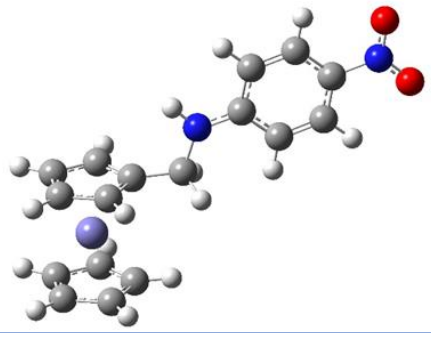
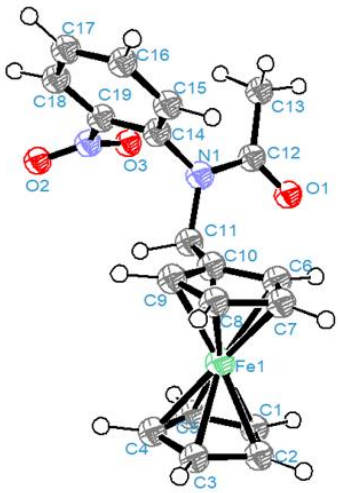
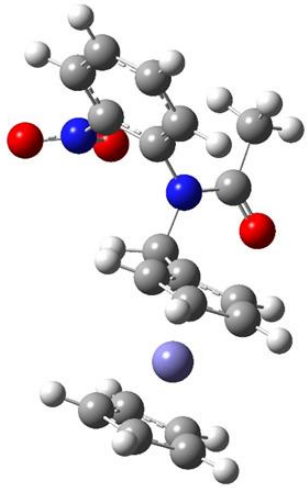
The obtained results make it easy to predict the anti-proliferative activity against MCF7 cell line for other substituted ferrocenylmethylaniline derivatives before carrying any experimental assays.



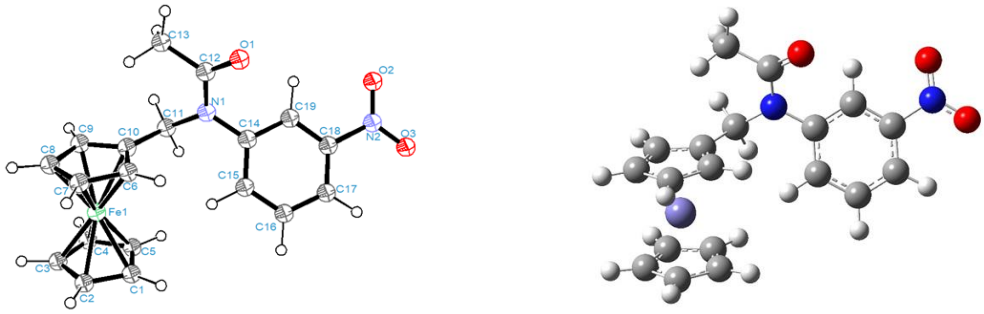
Appendix 1

Molecular Geometry, Structural and electronic property studies

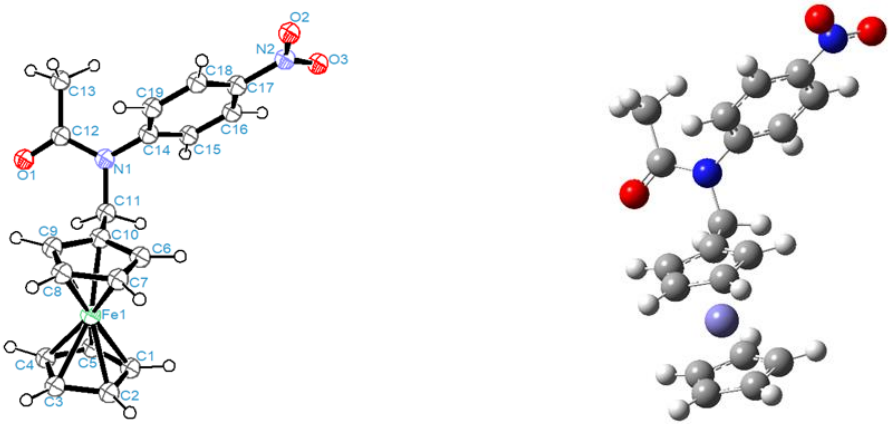
1. Optimized structures of studied ferrocenic derivatives by Ortep view and GaussView

Ortep view	GaussView
N-(ferrocènylméthyl)-2-nitrobenzènamine (A16)	
	
N-(ferrocènylméthyl)-3 -nitrobenzènamine (A17)	
	
N-(ferrocènylméthyl)-4 -nitrobenzènamine (A18)	
	
N- ferrocenylmethyl-N-(2-nitrophenyl)acetamide (A23)	
	

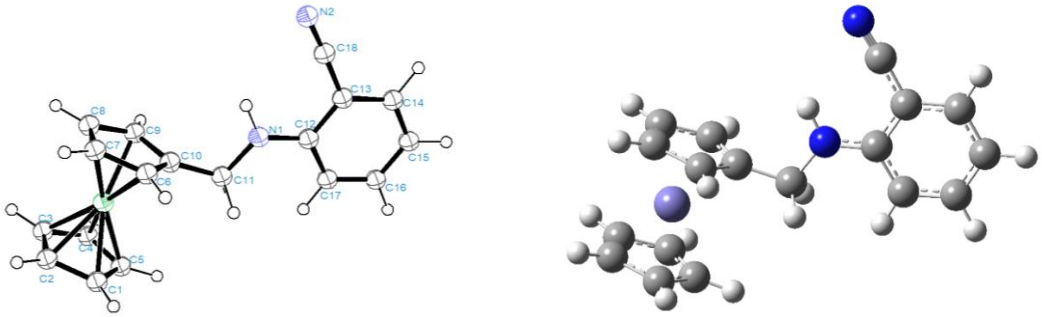
N- ferrocenylmethyl-N-(3-nitrophenyl)acetamide (A24)



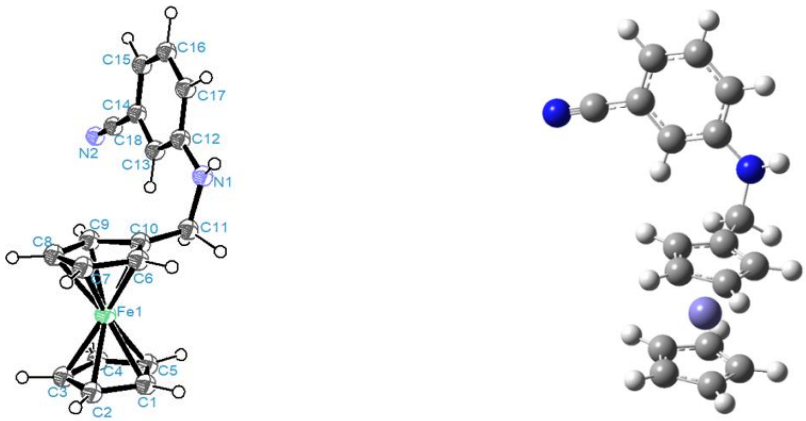
N- ferrocenylmethyl-N-(4-nitrophenyl)acetamide (A25)



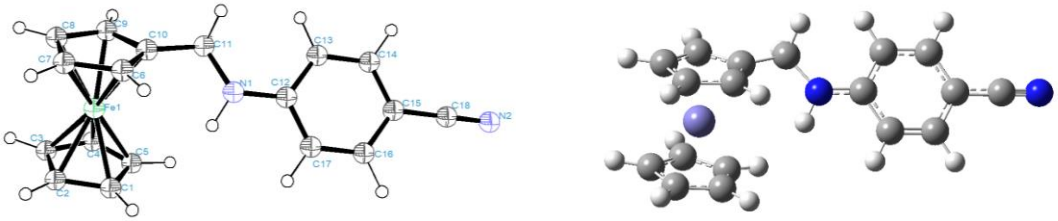
2-(ferrocenylmethylamino) benzonitrile (A13)



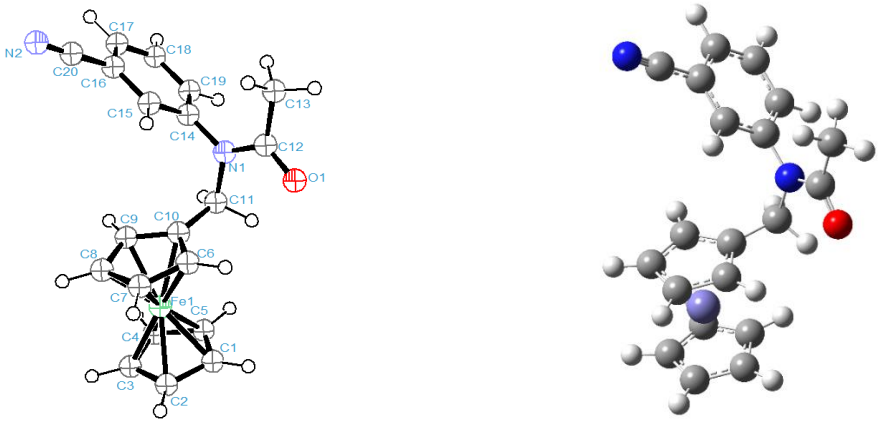
3-(ferrocenylmethylamino)benzonitrile (A14)



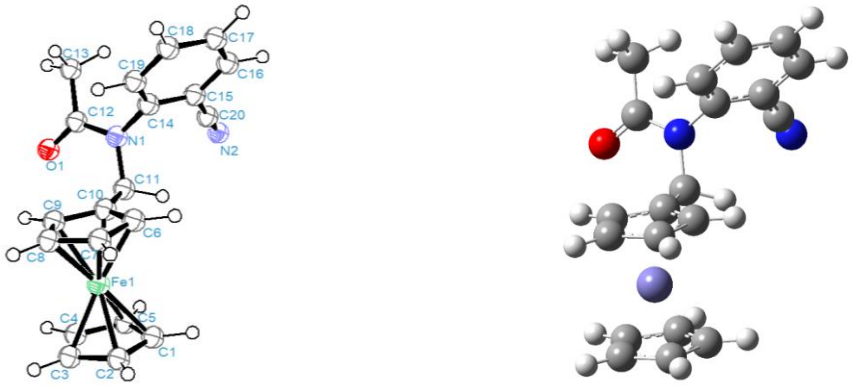
4-(ferrocenylmethylamino)benzonitrile (A15)



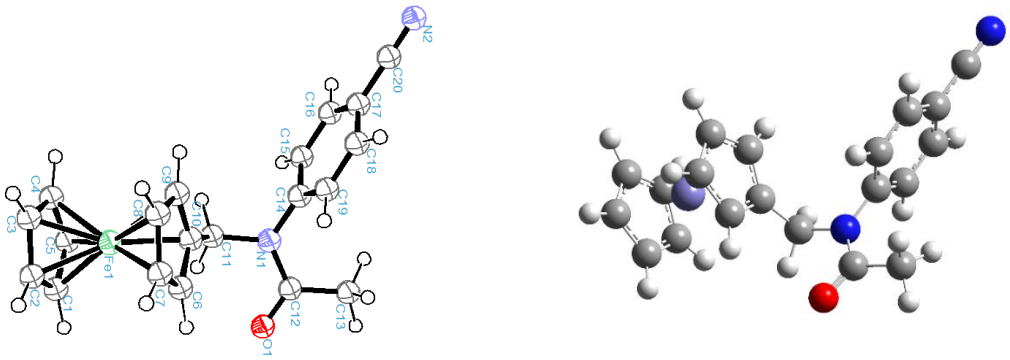
N-ferrocénylméthyl-N-(3-cyanophenyl) benzamide (A27)



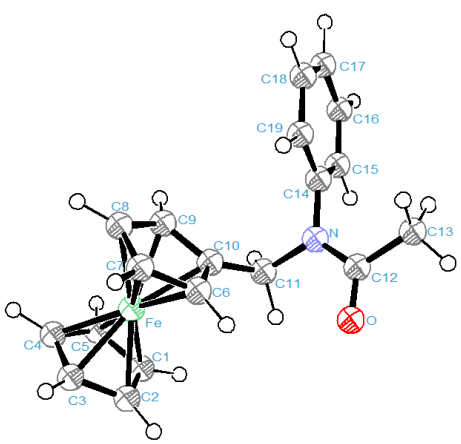
N-ferrocénylméthyl-N-(2-cyanophenyl) benzamide (A26)

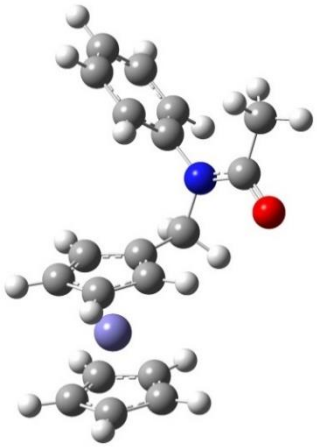


N-ferrocénylméthyl-N-(4-cyanophenyl) benzamide (A28)

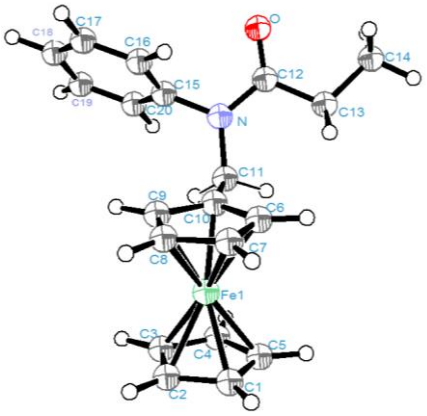


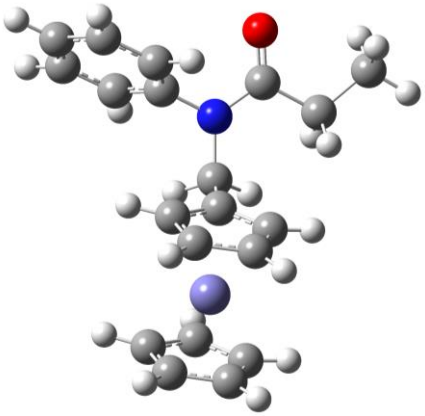
N-ferrocénylméthyl-N-phénylacétamide (A1)



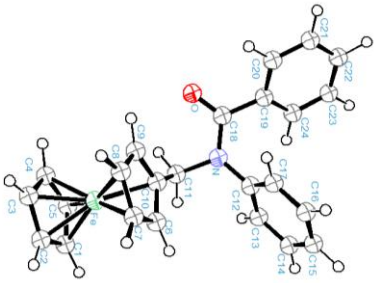


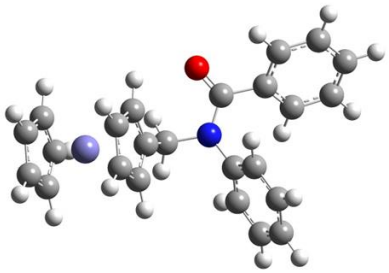
N-ferrocénylméthyl-N-phénylpropionamide (A2)



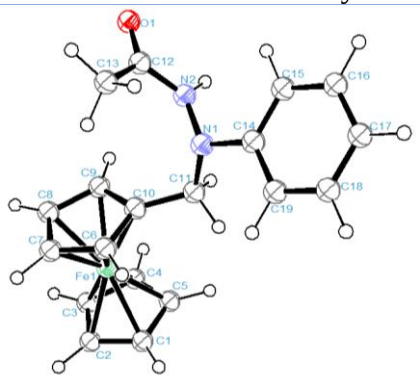


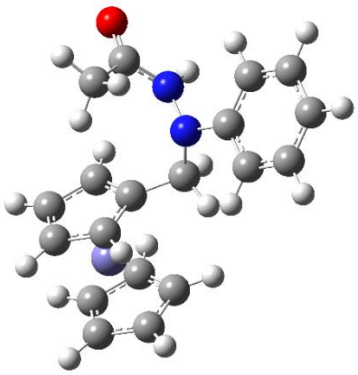
N-ferrocénylméthyl-N-phénylbenzamide (A3)



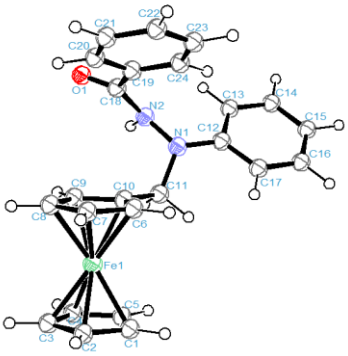


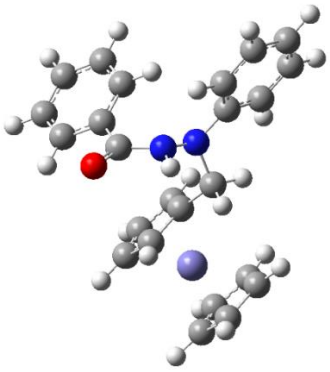
N-ferrocenylmethyl-N-phenylacetohydrazide (A21)





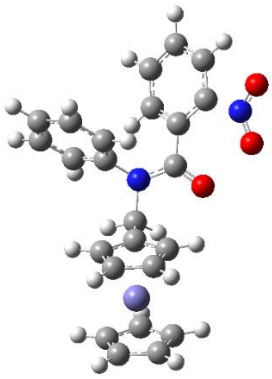
N-ferrocenylmethyl-N-pheylbenzohydrazide (A22)

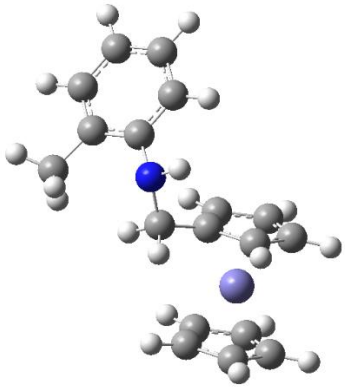




N-ferrocenylmethyl-2-nitro-N-phenylbenzamide (A4)

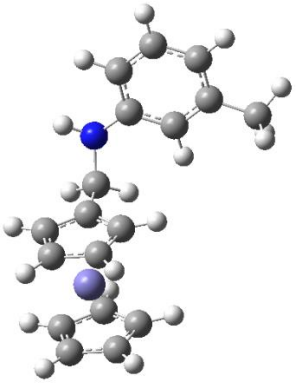
N-ferrocenylmethyl-2-methylaniline (A5)

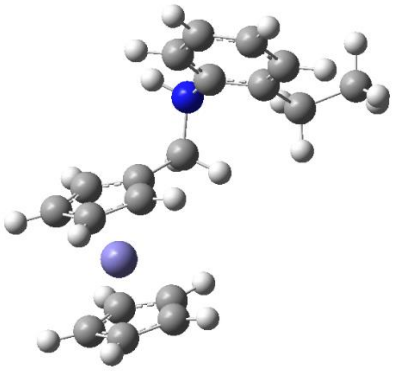




N-Ferrocenylmethyl-3-methylaniline (A6)

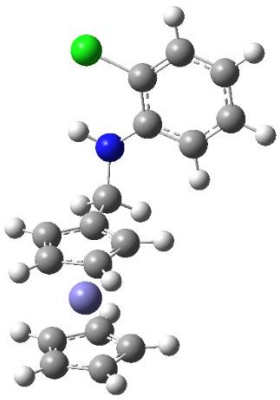
N-Ferrocenylmethyl-2-ethylaniline (A7)

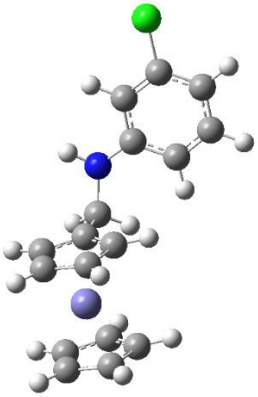




N-Ferrocenylmethyl-2-Chloroaniline (A8)

N-Ferrocenylmethyl-3-Chloroaniline (A9)





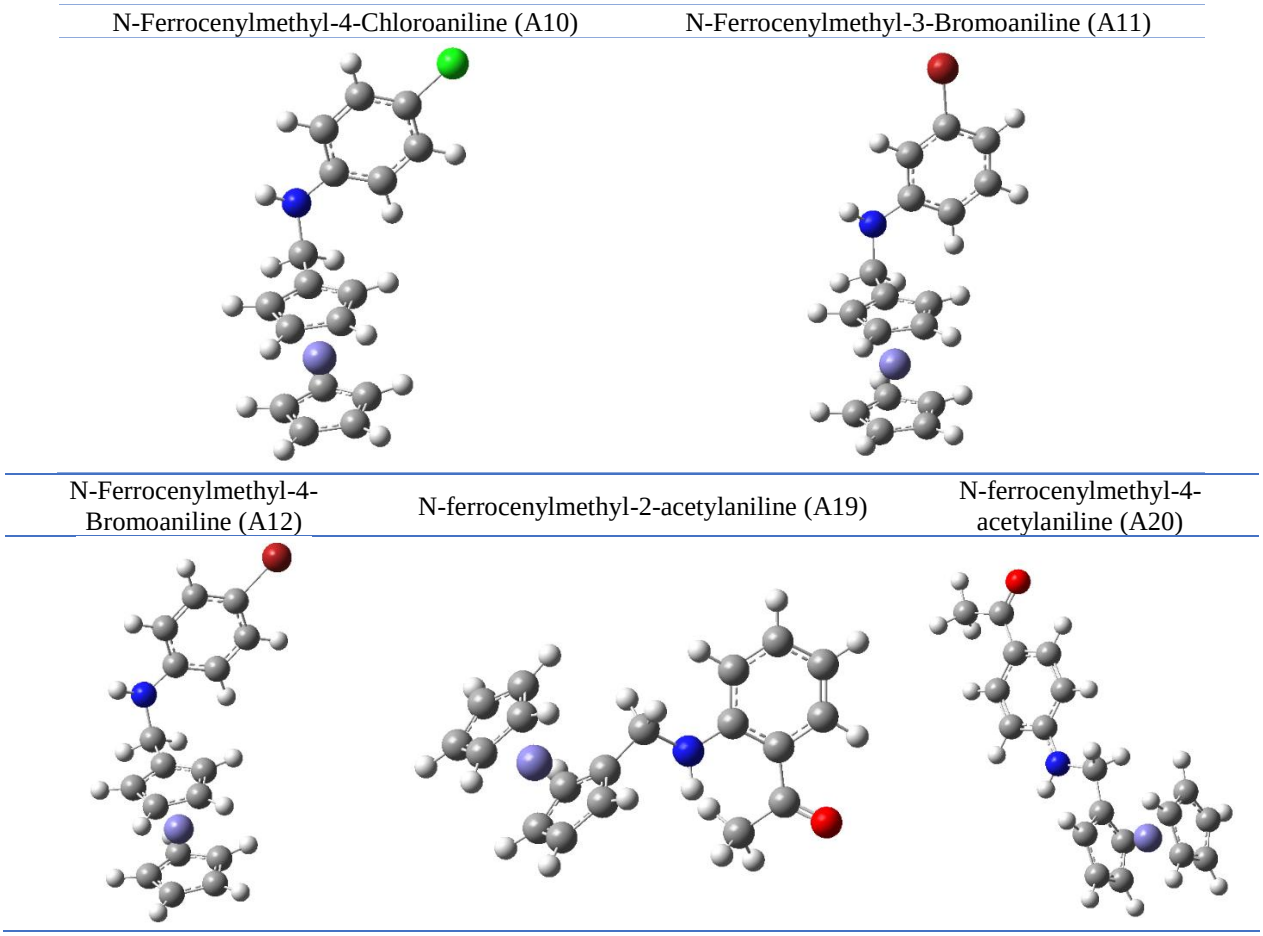


Figure 1. optimized structures of studied ferrocenic derivatives by Ortep view and GaussView.

1. Molecular Geometry

Table 1. Selected optimized and experimental bond distances of studied compound.

	Bond distances(Å)		X-ray	a	b				X-ray	a	b
A16	O1	N2	1.242	1.227	1.245	A1	C12	O	1.225	1.2237	1.261
	O2	N2	1.254	1.245	1.305		C12	N	1.359	1.3794	1.3899
	N2	C13	1.4471	1.449	1.447		C12	C13	1.509	1.5185	1.5234
A17	O1	N2	1,227	1,226	1,282		C13	H13a	0.96	1.0885	1.0921
	O2	N2	1,229	1,225	1,279		C13	H13b	0.959	1.0908	1.095
	N2	C14	1,474	1,483	1,479		C13	H13c	0.96	1.0935	1.0972
A18	O1	N2	-	1,23	1,286	A2	C12	O	1.229	1.22	1.2597
	O2	N2	-	1,230	1,286		C12	N	1.368	1.384	1.3959
	N2	C15	-	1,456	1,455		C12	C13	1.513	1.5300	1.5334
A23	N1	C12	-	1.386	1.396		C13	H13a	0.97	1.0940	1.0965
	N2	C19	-	1.478	1.478		C13	H13b	0.97	1.0962	1.0947
	N2	O2	-	1.225	1.281		C13	C14	1.523	1.5276	1.5496
	O1	C12	-	1.22	1.257		C14	H14b	0.959	1.0912	1.0970
	N2	O3	-	1.222	1.277		C14	H14c	0.96	1.0908	1.0961
A24	O1	C12	1.227	1.217	1.256		C14	H14a	0.96	1.0933	1.0968
	N2	O3	1.222	1.225	1.281	A3	C19	C20	1.389	1.3995	1.4123
	N2	C18	1.464	1.484	1.48		C19	C24	1.393	1.3987	1.4122

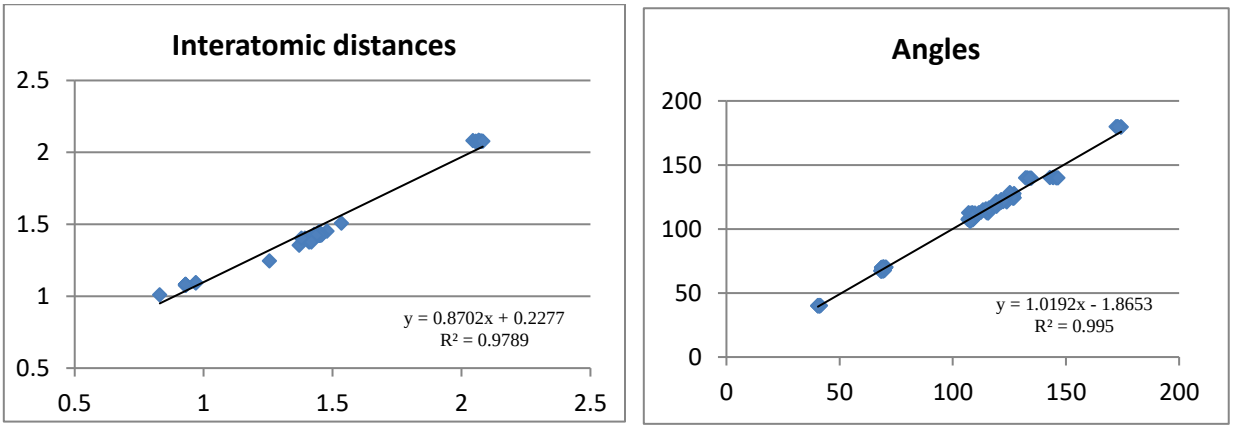
	C12	C13	1.502	1.516	1.525		C20	C21	1.385	1.3913	1.4047
	N1	C12	1.365	1.396	1.401		C21	C22	1.375	1.3947	1.409
	N2	O2	1.226	1.223	1.279		C22	C23	1.388	1.3938	1.4084
A25	C12	C13	-	1.519	1.524		C23	C24	1.383	1.3929	1.4059
	N1	C12	-	1.387	1.398		O	C18	1.224	1.2254	1.2650
	N2	C17	-	1.478	1.473		C18	C19	1.51	1.5052	1.5054
	N2	O2	-	1.225	1.28		N	C18	1.371	1.3827	1.3944
	N2	O3	-	1.225	1.281	A21	N1	N2	-	1.386	1.397
	O1	C12	-	1.22	1.259		N2	H12	-	1.018	1.021
A13	C13	C18	-	1.424	1.433		N2	C12	-	1.379	1.396
	C18	N2	-	1.158	1.186		O1	C12	-	1.219	1.255
A14	C14	C18	-	1.433	1.443		C12	C13	-	1.512	1.517
	C18	N2	-	1.156	1.183		C13	H13b	-	1.088	1.093
A15	N2	C18	-	1.157	1.185	A22	C18	C19	1.5	1.497	1.5
	C15	C18	-	1.426	1.435		O1	C18	1.23	1.221	1.261
A27	C12	C13	1.506	1.518	1.523		N1	C12	1.42	1.424	1.421
	C20	C16	1.442	1.432	1.441		N2	C18	1.35	1.388	1.397
	O1	C12	1.233	1.222	1.259		N2	N1	1.41	1.392	1.402
	N2	C20	1.151	1.156	1.182		C22	C23	1.383	1.394	1.408
	N1	C12	1.364	1.383	1.394		C24	C23	1.381	1.394	1.406
	H13a	C13	0.888	1.089	1.093		C24	C19	1.39	1.399	1.413
	H13b	C13	0.862	1.091	1.096		C20	C21	1.381	1.391	1.404
	H13c	C13	0.904	1.093	1.097		C20	C19	1.39	1.399	1.414
A28	C12	C13	-	1,518	1.524		C22	C21	1.386	1.394	1.409
	C13	H13a	-	1,089	1.096	A26	C20	C15	-	1.431	1.44
	C20	C17	-	1,431	1.440		O1	C12	-	1.22	1.258
	N1	C12	-	1,384	1.395		C12	C13	-	1.519	1.523
	N2	C20	-	1,156	1.183		N1	C12	-	1.386	1.396
	O1	C12	-	1,221	1.259		N2	C20	-	1.155	1.182
							H13a	C13	-	1.09	1.096

Table 2. Selected optimized and experimental bond angles of studied compound.

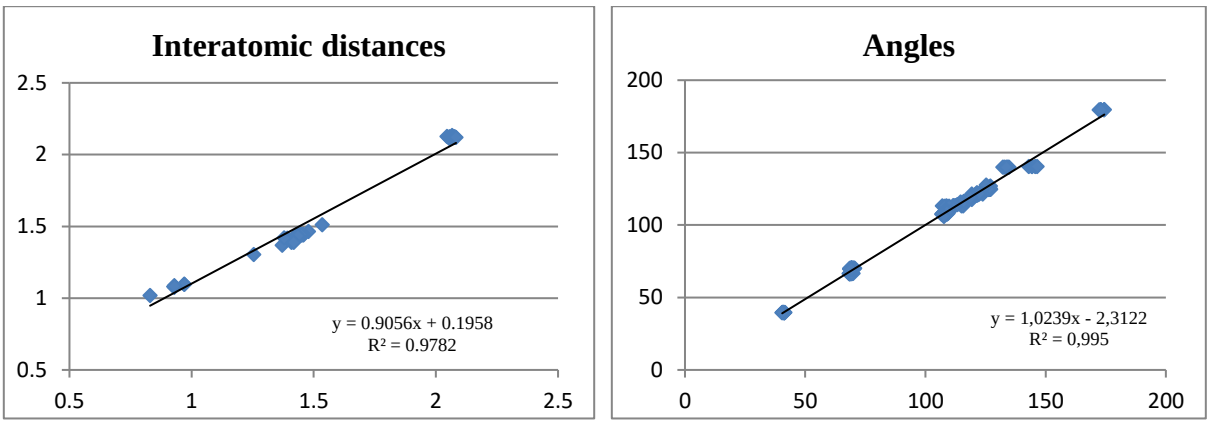
				X-ray	a	b					X-ray	a	b
	Bond angles (°)												
A16	O1	N2	O2	121,7	122,18	121,08	A1	O	C12	N	120.99	121.565	121.081
	O1	N2	C13	118,8	118,94	119,52		O	C12	C13	121.32	121.248	121.354
	O2	N2	C13	119,3	118,87	119,38		N	C12	C13	117.69	117.187	117.564
	C14	C13	N2	116,03	116,359	116,486		H13b	C13	H13a	109.49	110.154	109.922
	C12	C13	N2	121,7	122,297	122,07	A2	O	C12	N	121.49	121.384	121.423
A17	O1	N2	O2	123,27	123,23	124,326		O	C12	C13	121.92	119.194	121.12
	O1	N2	C14	118,34	118,34	117,876		N	C12	C13	116.58	119.422	117.448
	O2	N2	C14	118,39	118,42	117,798		H13a	C13	H13b	107.88	107.247	106.13
	C15	C14	N2	118,69	118,58	118,698		C12	C13	C14	112.28	112.704	112.254
	C13	C14	N2	117,28	117,91	117,908		H14b	C14	H14c	109.52	108.126	107.276
A18	N2	C15	C14	-	119,697	119,63	A3	O	C18	N	121.8	121.246	120.383
	N2	C15	C16	-	119,662	119,542		O	C18	C19	121.2	119.642	119.673
	O1	N2	O2	-	123,947	123,024		N	C18	C19	117	119.064	119.9
	O1	N2	C15	-	118,068	118,528		C18	C19	C20	118.5	117.316	117.229
	O2	N2	C15	-	117,985	118,448		C18	C19	C24	122.1	123.238	123.267
A23	C18	C19	N2	-	116.63	116.36		C20	C19	C24	119.3	119.228	119.318
	H13a	C13	C12	-	107.77	108.14		C19	C20	C21	119.9	120.44	120.445
	H13a	C13	H13b	-	109.15	109.38		C20	C21	C22	120.8	120.051	120.028
	O1	C12	C13	-	122.1	122.12		C21	C22	C23	119.5	119.789	119.745
	O1	C12	N1	-	121.06	120.51		C22	C23	C24	120.3	120.207	120.257

	O2	N2	C19	-	116.9	117.35		C19	C24	C23	120.2	120.255	120.185
	O3	N2	C19	-	117.92	118.86	A21	N2	C12	C13	-	116.78	117.19
	O3	N2	O2	-	125.15	123.77		N2	N1	C14	-	116.08	118.67
A24	N1	C12	C13	117.82	116.88	117.79		O1	C12	C13	-	123.8	123.94
	O1	C12	C13	120.79	121.33	120.87		O1	C12	N2	-	119.42	118.87
	O1	C12	N1	121.37	121.78	121.32	A22	C18	N2	N1	121.12	124.95	127.72
	C19	C18	N2	118.12	118.28	118.55		C20	C19	C18	117.09	116.85	116.16
	O2	N2	C18	118.62	117.9	118.34		C20	C19	C24	118.8	119.41	119.4
	O2	N2	O3	123.29	124.67	123.67		C20	C21	C22	119.93	120.03	119.92
	O3	N2	C18	118.09	117.45	117.99		C21	C20	C19	120.9	120.4	120.51
	H13c	C13	C12	109.54	112.21	111.67		C21	C22	C23	119.36	119.85	119.77
	H13c	C13	H13b	109.46	107.91	108.51		C23	C24	C19	120.41	120.04	119.95
A25	O1	C12	N1	-	121.15	120.45		C24	C19	C18	123.98	123.61	124.33
	O1	C12	C13	-	121.1	120.88		C24	C23	C22	120.56	120.26	120.42
	N1	C12	C13	-	117.72	118.63		N2	C18	C19	116.09	119.61	121.7
	H13c	C13	H13a	-	110.1	109.78		N2	N1	C11	112.43	112.04	115.32
	O2	N2	O3	-	124.74	123.7		N2	N1	C12	114.54	115.02	117.28
	O2	N2	C17	-	117.65	118.17		O1	C18	C19	121.9	121.81	121.4
	O3	N2	C17	-	117.61	118.13		O1	C18	N2	121.9	118.52	116.9
	C18	C17	N2	-	119.07	119.07		C24	C23	C22	120.56	120.26	120.42
	C16	C17	N2	-	119.02	119.02	A26	N1	C12	C13	-	117.2	117.61
A13	C18	C13	C12	-	118.92	119.1		N2	C20	C15	-	178.64	179
	N2	C18	C13	-	177.53	178.33		O1	C12	C13	-	121.45	121.59
	H12	N1	C11	-	117	117.04		O1	C12	N1	-	121.33	120.8
	C14	C13	C18	-	120.29	120.25	A27	N2	C20	C16	178.9	179.72	179.85
A14	N1	C12	C13	-	122.81	122.85		O1	C12	C13	120.59	121.4	121.54
	N1	C12	C17	-	119.29	119.13		O1	C12	N1	121.5	121.31	120.87
	N2	C18	C14	-	179.74	179.83		N1	C12	C13	117.9	117.29	117.59
	H12	N1	C12	-	113.75	116.61	A28	O1	C12	N1	-	121,28	120.78
A15	N1	C12	C13	-	122.44	122.05		O1	C12	C13	-	121,3	121.28
	N1	C12	C17	-	119.47	119.74		N1	C12	C13	-	117,42	117.93
	N2	C18	C15	-	179.92	179.97		N2	C20	C17	-	179,9	179.95

2. Distribution graphs for geometric parameters

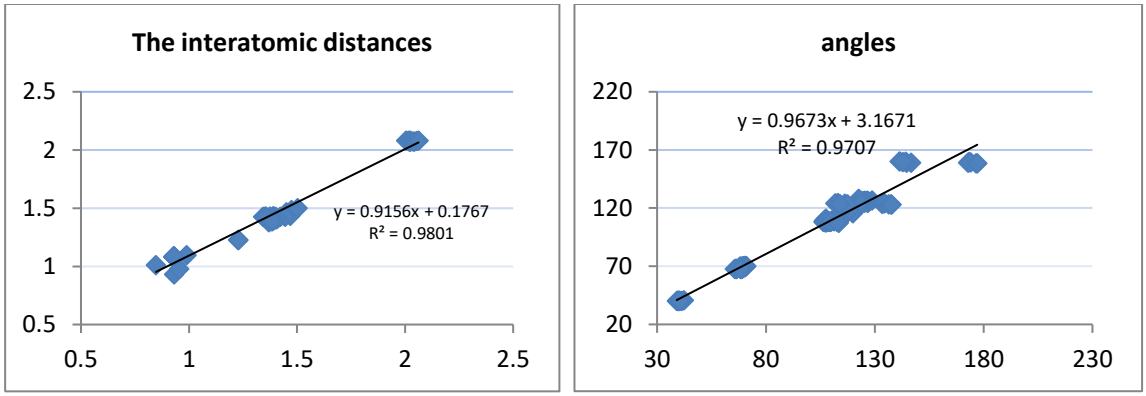


(a)

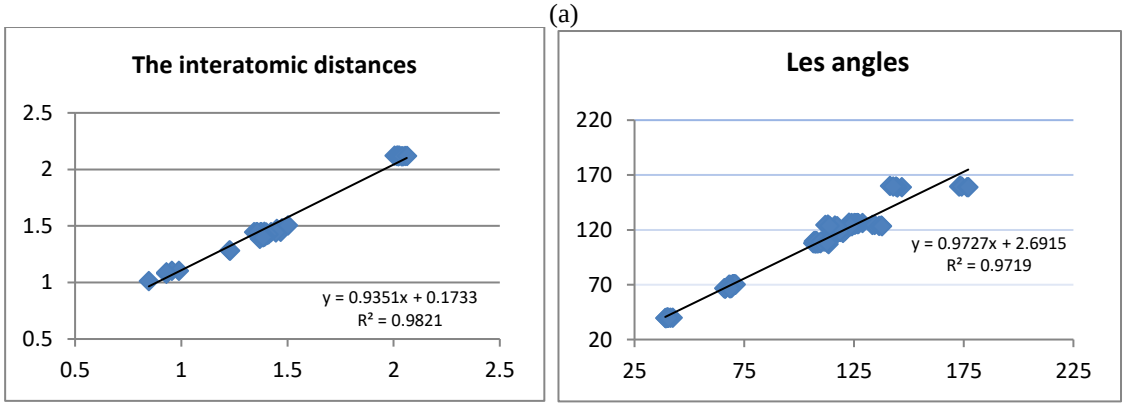


(b)

Graph 1. Distribution graphs between the experimental and calculated structural parameters for A16 .(a): 6-311G++ (d,p), (b): LanL2DZ.

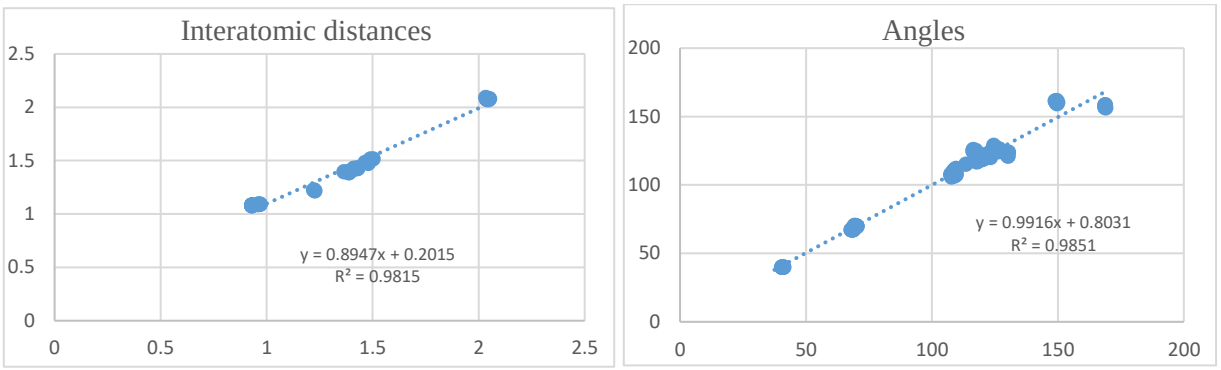


(a)

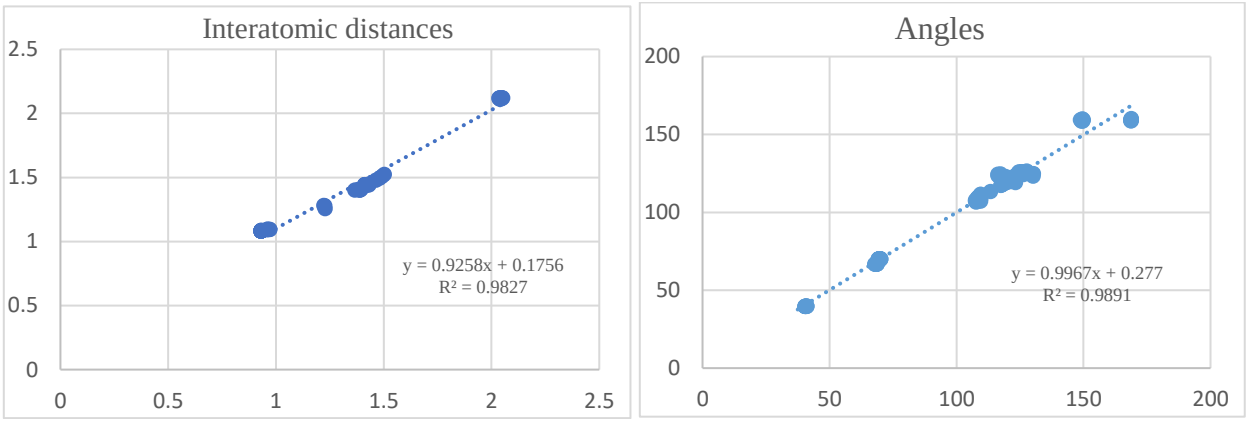


(b)

Graph 2. Distribution graphs between the experimental and calculated structural parameters for A17 .(a): 6-311G++ (d,p), (b): LanL2DZ.

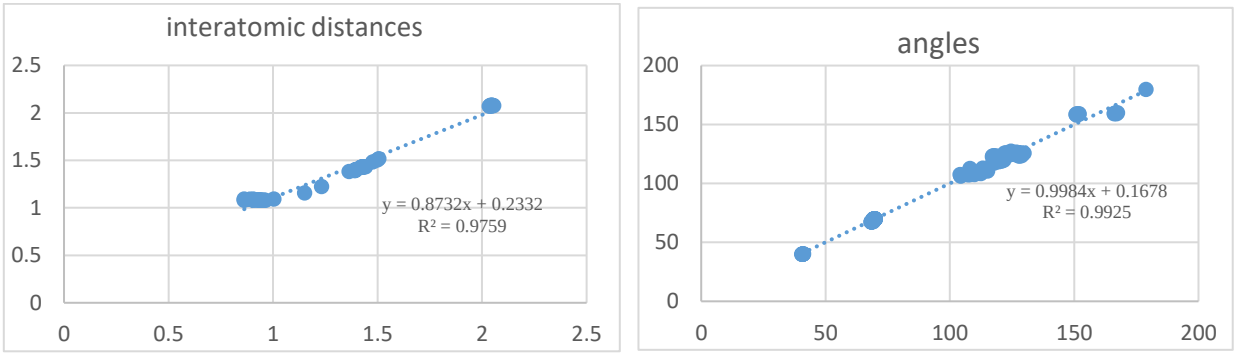


(a)

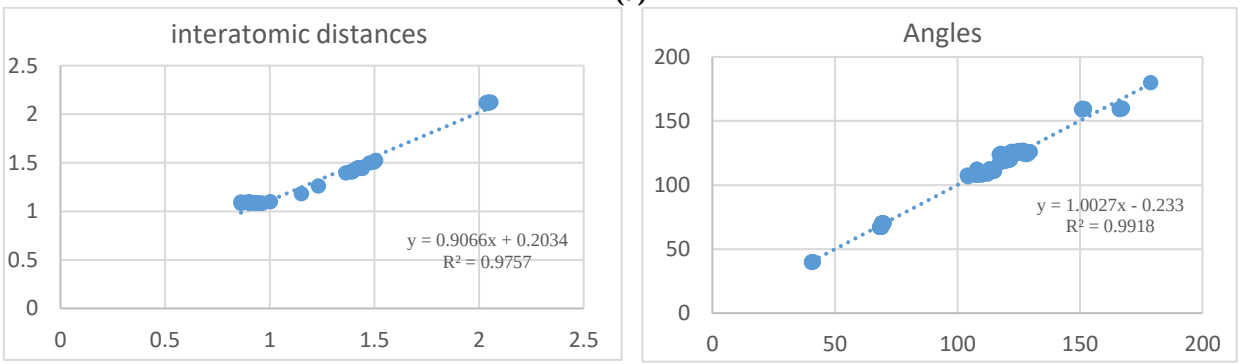


(b)

Graph 3. Distribution graphs between the experimental and calculated structural parameters for A18.(a): 6-311G++ (d,p), (b): LanL2DZ.

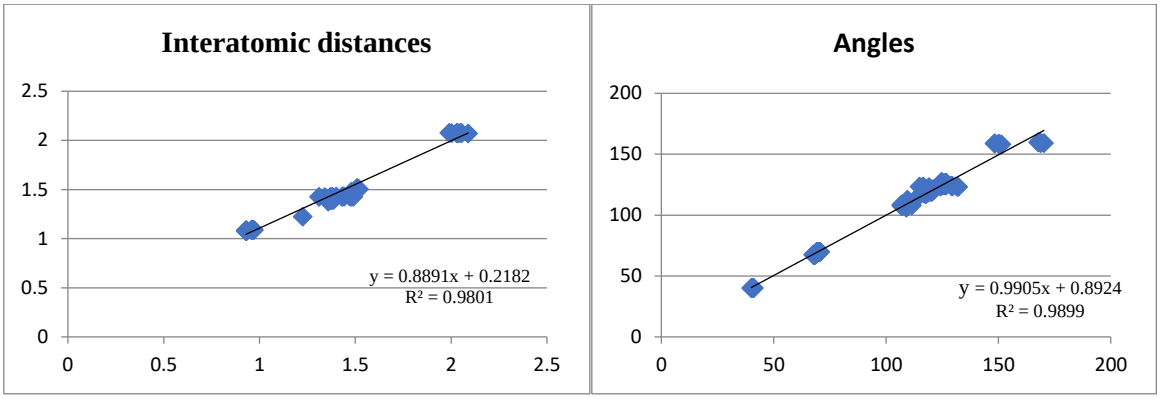


(a)

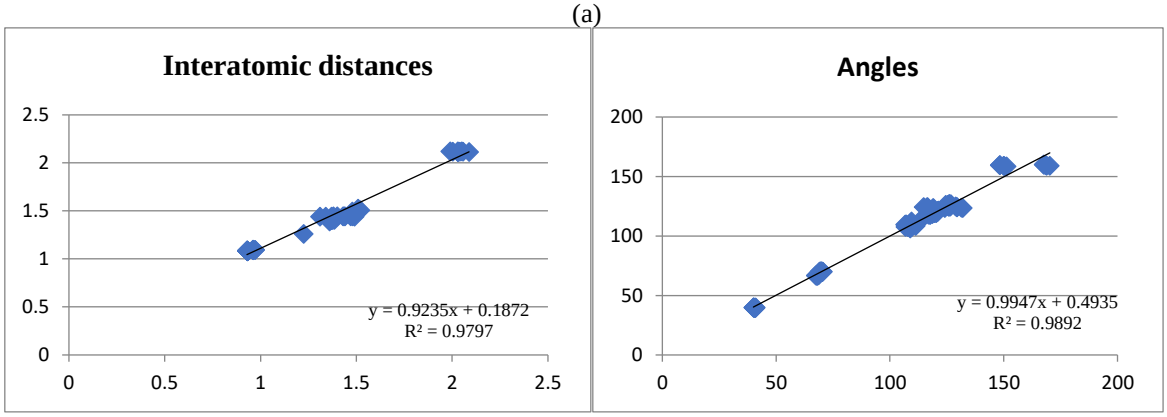


(b)

Graph 4. Distribution graphs between the experimental and calculated structural parameters for A27 .(a): 6-311G++ (d,p), (b): LanL2DZ

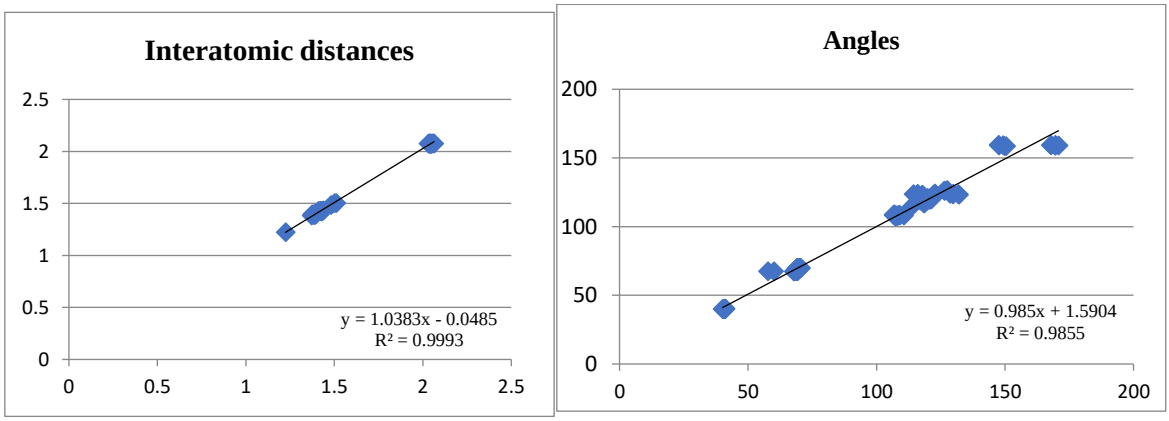


(a)

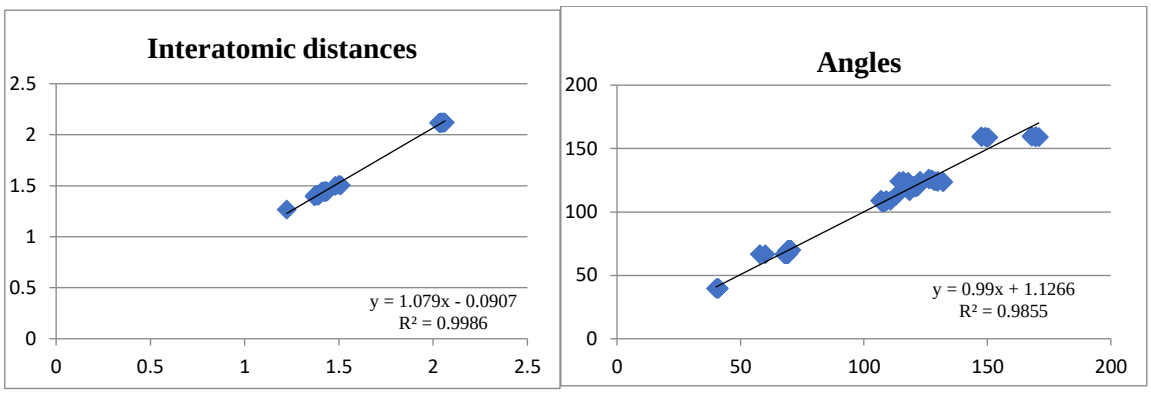


(b)

Graph 5. Distribution graphs between the experimental and calculated structural parameters for A1.(a): 6-311G++ (d,p), (b): LanL2DZ

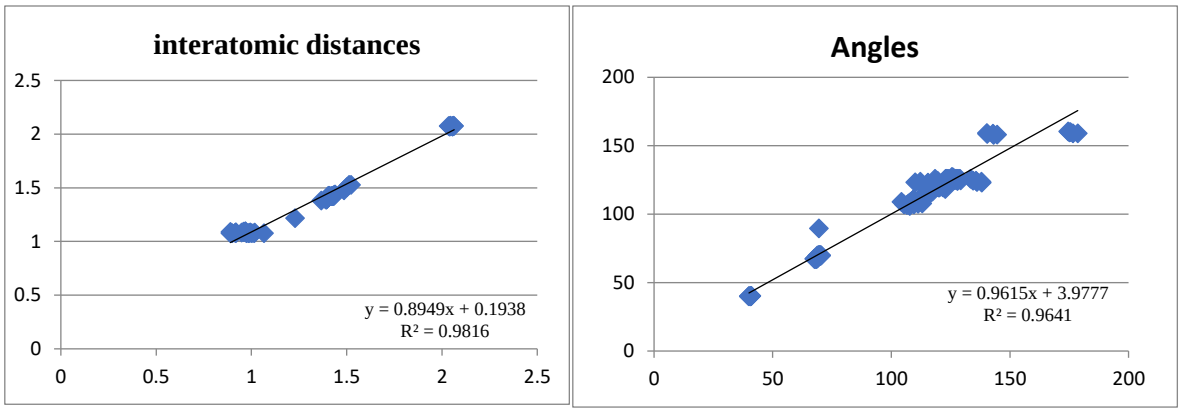


(a)

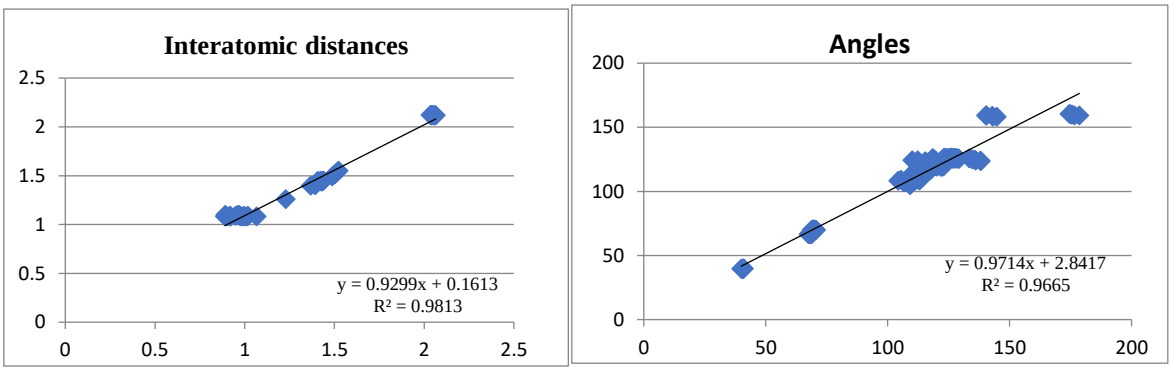


(b)

Graph 6. Distribution graphs between the experimental and calculated structural parameters for A3.(a): 6-311++ G (d,p), (b): LanL2DZ

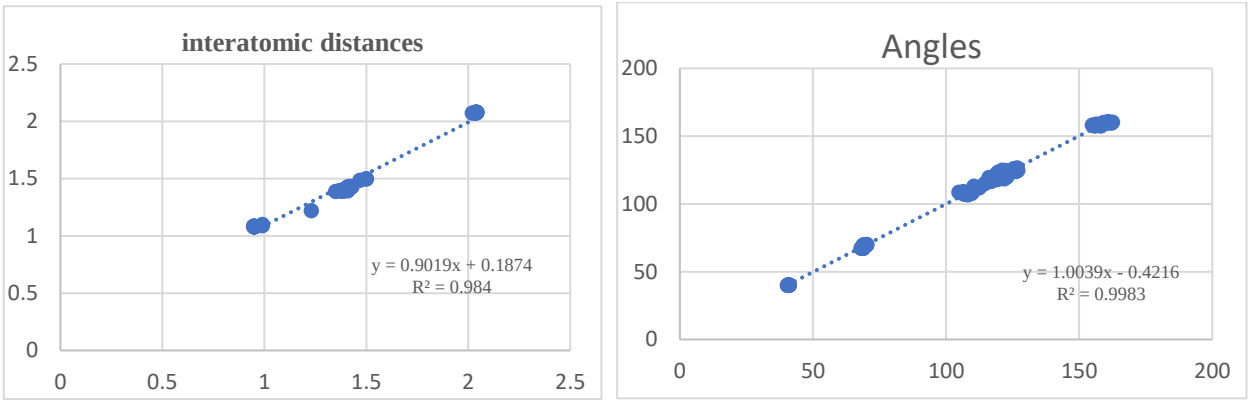


(a)

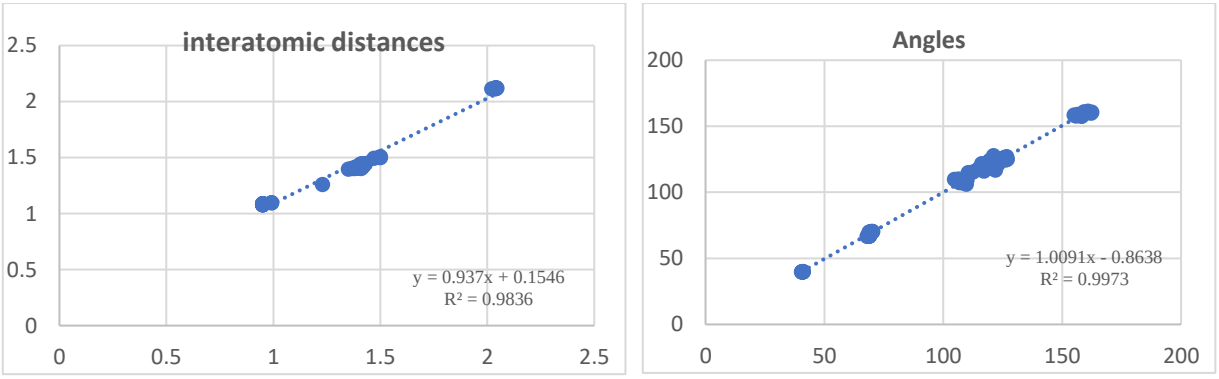


(b)

Graph 7. Distribution graphs between the experimental and calculated structural parameters for A2.(a): 6-311G++ (d,p), (b): LanL2DZ.



(a)



(b)

Graph 8. Distribution graphs between the experimental and calculated structural parameters for A22.(a): 6-311G++ (d,p), (b): LanL2DZ

4. Atomic charges

Table 3. Mulliken atomic charges for A16, A17 and A18 by DFT/B3LYP method with 6-311++ G (d,p) and LanL2DZ basis sets.

A16			A17			A18		
Atoms	a	b	Atoms	a	b	atoms	a	b
Fe	0,429	-0,136	Fe	0,71	-0,154	Fe	0,821	-0,152
C1	0,731	0,358	C1	-0,185	0,302	C1	-0,152	0,298
C2	-0,176	-0,358	C2	-0,355	-0,355	C2	-0,373	-0,353
C3	-0,164	-0,22	C3	-0,131	-0,229	C3	-0,232	-0,227
C4	-0,345	-0,234	C4	-0,32	-0,229	C4	-0,232	-0,227
C5	-0,251	-0,336	C5	-0,061	-0,355	C5	-0,374	-0,353
C6	-0,164	-0,273	C6	-0,081	-0,244	C6	-0,102	-0,244
C7	-0,175	-0,232	C7	-0,327	-0,229	C7	-0,327	-0,229
C8	-0,23	-0,228	C8	-0,162	-0,234	C8	-0,132	-0,233
C9	-0,16	-0,231	C9	-0,114	-0,235	C9	-0,133	-0,233
C10	-0,327	-0,24	C10	-0,394	-0,229	C10	-0,327	-0,229
C11	-0,998	-0,519	C11	-0,707	-0,448	C11	-0,468	-0,459
N1	0,093	-0,394	C12	-0,455	0,485	C12	-0,066	0,46
C12	-0,257	0,336	C13	0,338	-0,451	C13	0,466	-0,304
C13	-0,032	0,172	C14	0,203	0,358	C14	-0,002	-0,36
C14	0,066	-0,312	C15	0,286	-0,39	C15	-0,451	0,358
C15	-0,395	-0,227	C16	-0,289	-0,144	C16	-0,316	-0,33
C16	-0,553	-0,221	C17	-0,876	-0,417	C17	-0,391	-0,379

C17	0,309	-0,294	N1	0,121	-0,451	N1	0,066	-0,425
N2	-0,327	0,071	N2	-0,194	0,063	N2	-0,218	0,065
O1	0,019	-0,248	O1	-0,004	-0,245	O1	-0,023	-0,263
O2	-0,025	-0,315	O2	0,01	-0,238	O2	-0,021	-0,264

Table 4. Mulliken atomic charges for A23, A24 and A25 by DFT/B3LYP method with 6-311++ G (d,p) and LanL2DZ basis sets.

A23			A24			A25		
Atoms	a	b	Atoms	a	b	atoms	a	b
C1	-0.298	-0.227	C1	-0.064	-0.229	C1	-0.233	-0.233
C2	-0.128	-0.239	C2	-0.241	-0.234	C2	-0.237	-0.237
C3	-0.188	-0.237	C3	-0.113	-0.234	C3	-0.237	-0.237
C4	-0.307	-0.233	C4	-0.202	-0.231	C4	-0.228	-0.228
C5	-0.155	-0.246	C5	-0.186	-0.247	C5	-0.247	-0.247
C6	-0.299	-0.364	C6	-0.661	-0.364	C6	-0.387	-0.387
C7	-0.308	-0.239	C7	-0.301	-0.223	C7	-0.228	-0.228
C8	-0.397	-0.232	C8	-0.286	-0.229	C8	-0.234	-0.234
C9	-0.31	-0.394	C9	-0.302	-0.38	C9	-0.36	-0.36
C10	-0.772	0.464	C10	0.893	-0.478	C10	0.468	0.458
C11	-0.986	-0.54	C11	-0.85	-0.541	C11	-0.541	-0.541
C12	0.283	0.246	C12	0.177	0.214	C12	0.214	0.214
C13	0.52	-0.735	C13	-0.574	-0.744	C13	-0.711	-0.711
C14	-0.199	0.308	C14	-0.035	0.54	C14	0.465	0.465
C15	0.197	-0.281	C15	0.303	-0.374	C15	-0.339	-0.339
C16	-0.273	-0.213	C16	-0.444	-0.174	C16	-0.341	-0.341
C17	-0.429	-0.202	C17	0.104	-0.37	C17	0.367	0.367
C18	0.071	-0.336	C18	-0.217	0.351	C18	-0.347	-0.347
C19	-0.199	0.258	C19	-0.543	-0.422	C19	-0.305	-0.305
Fe	0.696	-0.137	Fe	0.784	-0.155	Fe	-0.146	-0.146
N1	0.855	-0.242	N1	0.536	-0.266	N1	-0.244	-0.244
N2	-0.255	0.047	N2	-0.231	0.065	N2	0.069	0.069
O1	0.282	-0.312	O1	-0.265	-0.287	O1	-0.305	-0.305
O2	0.044	-0.223	O2	0.006	-0.226	O2	-0.235	-0.235
O3	0.043	-0.21	O3	0.003	-0.238	O3	-0.236	-0.236

Table 5**.** Mulliken atomic charges for A13, A14 and A15 by DFT/B3LYP method with 6-311++ G (d,p) and LanL2DZ basis sets.

A13			A14			A15		
Atoms	a	b	Atoms	a	b	atoms	a	b
C1	-0.301	-0.245	C1	-0.26	-0.231	C1	-0.185	-0.238
C2	-0.145	-0.231	C2	-0.181	-0.236	C2	-0.218	-0.233
C3	-0.146	-0.235	C3	-0.23	-0.236	C3	-0.181	-0.228
C4	-0.3	-0.235	C4	-0.195	-0.228	C4	-0.251	-0.23
C5	-0.131	-0.231	C5	-0.182	-0.248	C5	-0.248	-0.254
C6	-0.325	-0.35	C6	-0.141	-0.372	C6	-0.474	0.371
C7	-0.205	-0.224	C7	-0.22	-0.227	C7	-0.141	-0.22
C8	-0.204	-0.224	C8	-0.429	-0.234	C8	-0.192	-0.211
C9	-0.324	-0.35	C9	-0.368	-0.326	C9	-0.2	-0.391
C10	-0.22	-0.305	C10	-0.832	0.417	C10	-0.026	0.372
C11	-0.451	-0.455	C11	-1.267	-0.551	C11	-0.304	-0.486
C12	-0.579	0.275	C12	-0.407	0.479	C12	-0.058	0.464
C13	1.479	0.243	C13	0.849	-0.481	C13	-0.069	-0.325

C14	0.151	-0.33	C14	1.339	0.426	C14	-0.259	-0.376
C15	-0.367	-0.248	C15	-0.703	-0.413	C15	1.84	0.406
C16	-0.217	-0.213	C16	-0.272	-0.171	C16	-0.434	-0.349
C17	0.222	-0.327	C17	-0.223	-0.412	C17	0.025	-0.399
C18	-1.555	-0.284	C18	-1.259	-0.274	C18	-1.713	-0.273
Fe	0.868	-0.16	Fe	0.593	-0.152	Fe	0.527	-0.188
N1	0.108	-0.436	N1	0.192	-0.408	N1	0.078	-0.448
N2	-0.196	-0.046	N2	-0.181	-0.036	N2	-0.178	-0.046

Table 6. Mulliken atomic charges for A26, A27 and A28 by DFT/B3LYP method with 6-311++ G (d,p) and LanL2DZ basis sets.

A26			**A27**			**A28**		
Atoms	a	b	Atoms	a	b	atoms	a	b
C1	-0.307	-0.234	C1	-0.218	-0.229	C1	-0.233	-0.233
C2	-0.121	-0.238	C2	-0.219	-0.238	C2	-0.208	-0.237
C3	-0.192	-0.239	C3	-0.102	-0.238	C3	-0.112	-0.238
C4	-0.215	-0.228	C4	-0.343	-0.234	C4	-0.316	-0.228
C5	-0.204	-0.246	C5	-0.153	-0.245	C5	-0.179	-0.245
C6	-0.483	-0.394	C6	-0.232	-0.36	C6	-0.227	-0.387
C7	-0.348	-0.231	C7	-0.258	-0.235	C7	-0.278	-0.229
C8	-0.251	-0.238	C8	-0.389	-0.227	C8	-0.336	-0.235
C9	-0.198	-0.365	C9	-0.47	-0.387	C9	-0.419	-0.361
C10	1.048	0.472	C10	-0.913	0.459	C10	1.003	0.464
C11	-1.302	-0.541	C11	-1.26	-0.542	C11	-1.355	-0.543
C12	0.421	0.221	C12	0.334	0.213	C12	0.368	0.211
C13	-0.603	-0.721	C13	-0.567	-0.72	C13	-0.563	-0.715
C14	-0.954	0.284	C14	-0.788	0.447	C14	-0.821	0.450
C15	1.198	0.308	C15	-0.383	-0.476	C15	0.134	-0.351
C16	-0.078	-0.362	C16	1.445	0.42	C16	-0.277	-0.355
C17	-0.374	-0.212	C17	-0.316	-0.374	C17	1.749	0.420
C18	-0.286	-0.212	C18	-0.215	-0.184	C18	-0.314	-0.363
C19	0.4	-0.323	C19	0.218	-0.353	C19	-0.027	-0.331
C20	-1.453	-0.258	C20	-1.296	0.264	C20	-1.813	-0.258
Fe	0.704	-0.142	Fe	0.79	-0.141	Fe	0.651	-0.143
N1	0.703	-0.248	N1	0.591	-0.256	N1	0.665	-0.247
N2	-0.139	-0.02	N2	-0.165	-0.018	N2	-0.152	-0.020
O	-0.285	-0.318	O	-0.3	-0.316	O	-0.299	-0.313

a : 6-311++ G(d, p) , b: LanL2DZ

Table 7. Mulliken atomic charges for A1, A2 and A3 by DFT/B3LYP method with 6-311++ G (d,p) and LanL2DZ basis sets.

A1			A2			A3		
Atoms	a	b	Atoms	a	b	Atoms	a	b
C1	-0.182	-0.247	C1	-0.222	-0.231	C1	-0.258	-0.235
C2	-0.22	-0.229	C2	-0.188	-0.236	C2	-0.178	-0.239
C3	-0.203	-0.24	C3	-0.231	-0.238	C3	-0.211	-0.24
C4	-0.126	-0.239	C4	-0.215	-0.231	C4	-0.202	-0.23
C5	-0.331	-0.236	C5	-0.271	-0.247	C5	-0.229	-0.247
C6	-0.282	-0.36	C6	-0.133	-0.374	C6	-0.38	0.378
C7	-0.278	-0.238	C7	-0.243	-0.228	C7	-0.323	-0.228
C8	-0.356	-0.23	C8	-0.355	-0.288	C8	-0.294	-0.235
C9	-0.455	-0.386	C9	-0.4	-0.349	C9	-0.167	-0.355

C10	0.913	0.461	C10	-0.316	0.477	C10	0.749	0.465
C11	-1.132	-0.538	C11	-0.539	-0.544	C11	-0.8	-0.531
C12	0.286	0.212	C12	-0.024	0.173	C12	-0.282	0.438
C13	-0.553	-0.72	C13	-0.043	-0.436	C13	-0.054	0.388
C14	-0.66	0.42	C14	-0.745	-0.639	C14	-0.281	-0.207
C15	0.152	-0.359	C15	-0.465	0.486	C15	-0.381	-0.234
C16	-0.278	-0.202	C16	-0.129	-0.302	C16	-0.397	-0.205
C17	-0.382	-0.227	C17	-0.29	-0.213	C17	0.018	-0.324
C18	-0.271	-0.211	C18	-0.206	-0.218	C18	-0.623	0.096
C19	-0.028	-0.338	C19	-0.407	-0.232	C19	-0.647	0.283
N	0.55	-0.25	C20	-0.031	-0.365	C20	0.042	-0.362
O	-0.309	-0.327	Fe	0.548	-0.157	C21	-0.359	-0.221
Fe	0.736	-0.142	N	0.574	-0.257	C22	-0.222	-0.222
-	-	-	O	-0.24	-0.292	C23	-0.299	-0.232
-	-	-	-	-	-	C24	-0.066	-0.315
-	-	-	-	-	-	Fe	0.454	-0.151
-	-	-	-	-	-	N	0.546	-0.253
-	-	-	-	-	-	O	-0.242	-0.313

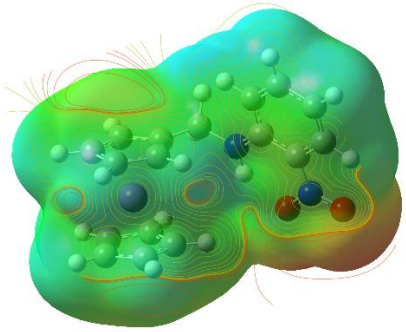
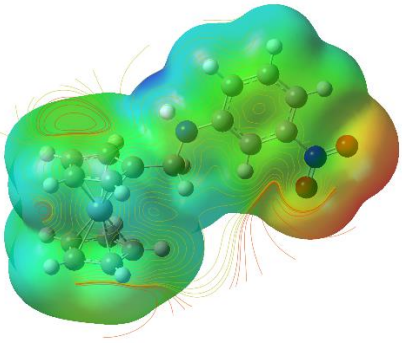
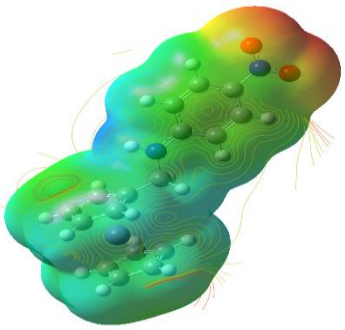
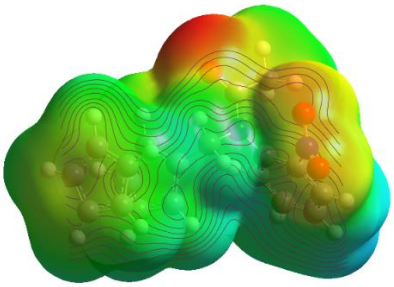
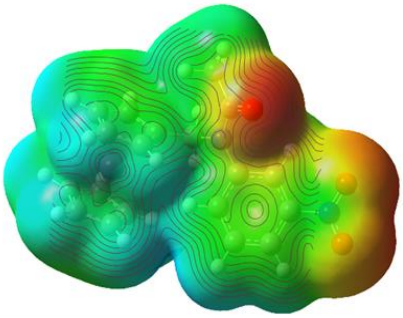
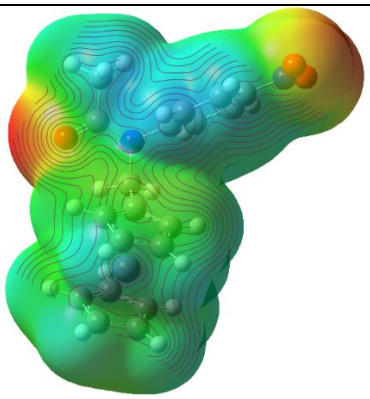
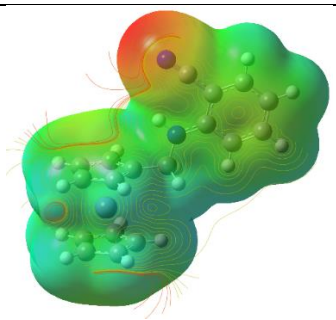
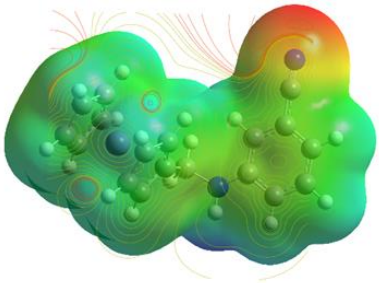
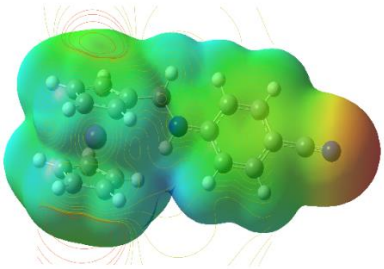
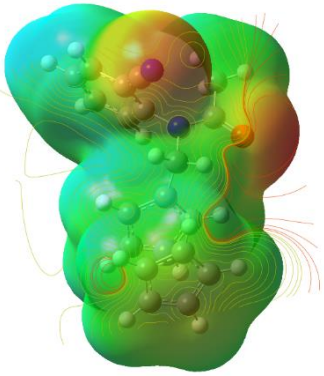
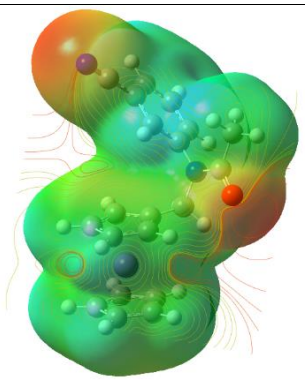
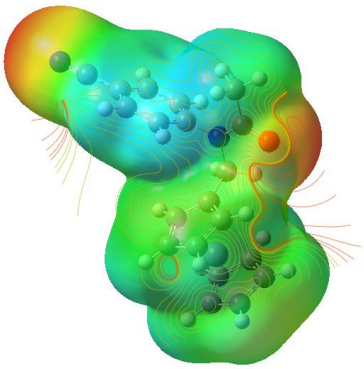
Table 8. Mulliken atomic charges for A21 and A22 by DFT/B3LYP method with 6-311++ G (d,p) and LanL2DZ basis sets.

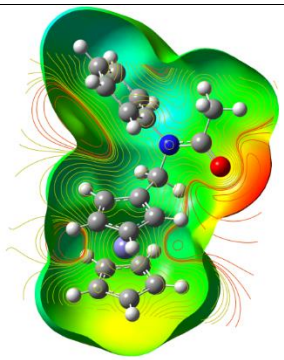
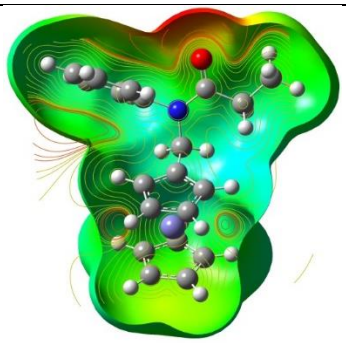
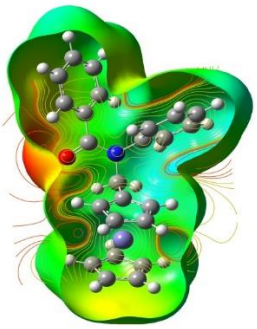
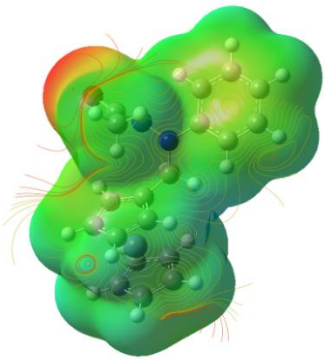
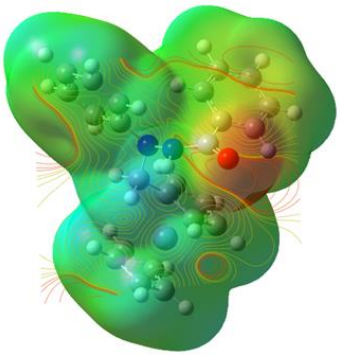
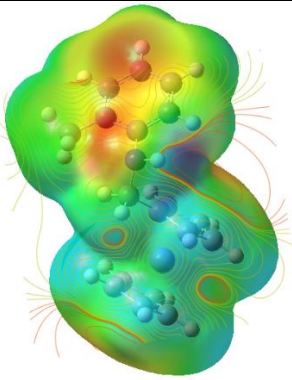
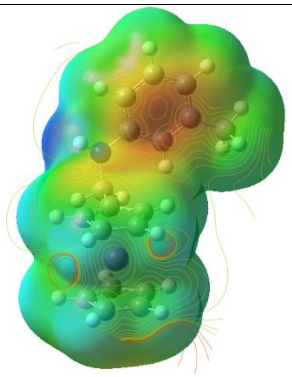
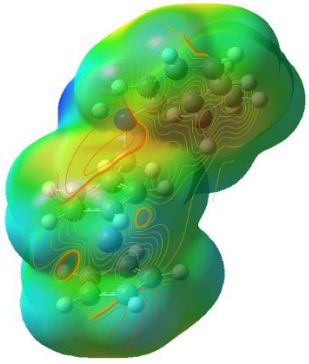
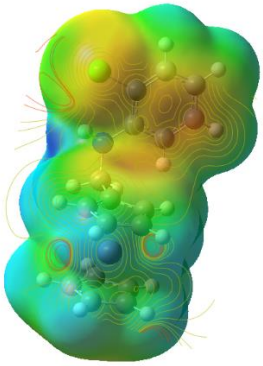
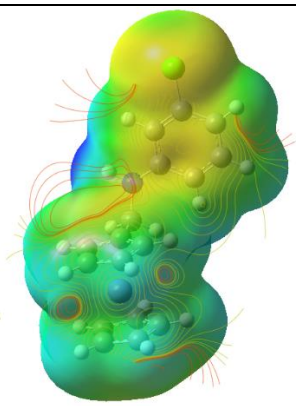
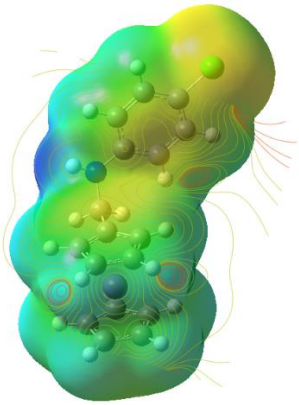
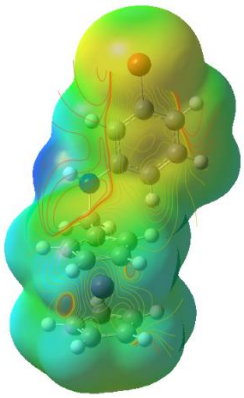
A21			A22		
Atoms	a	b	atoms	a	b
C1	-0.24	-0.23	C1	-0.221	-0.232
C2	-0.202	-0.236	C2	-0.216	-0.236
C3	-0.151	-0.236	C3	-0.126	-0.235
C4	-0.243	-0.231	C4	-0.288	-0.23
C5	-0.179	-0.246	C5	-0.144	-0.247
C6	-0.272	-0.356	C6	-0.385	-0.351
C7	-0.352	-0.235	C7	-0.237	-0.234
C8	-0.262	-0.232	C8	-0.238	-0.225
C9	-0.385	-0.372	C9	-0.423	-0.355
C10	0.361	0.445	C10	0.768	0.466
C11	-0.905	-0.535	C11	-0.936	-0.558
C12	-0.104	0.243	C12	-0.642	0.387
C13	-0.343	-0.699	C13	0.205	-0.371
C14	-0.584	0.405	C14	-0.112	-0.201
C15	0.007	-0.375	C15	-0.433	-0.255
C16	-0.086	-0.197	C16	-0.226	-0.219
C17	-0.301	-0.256	C17	0.074	-0.336
C18	-0.231	-0.22	C18	-1.109	0.113
C19	0.232	-0.332	C19	1.244	0.317
Fe	0.802	-0.149	C20	0.498	-0.356
N1	0.901	-0.154	C21	-0.442	-0.221
N2	-0.573	-0.38	C22	-0.283	-0.217
O	-0.267	-0.297	C23	-0.332	-0.244
			C24	-0.63	-0.278
			Fe	0.539	-0.167
			N1	-0.573	-0.146

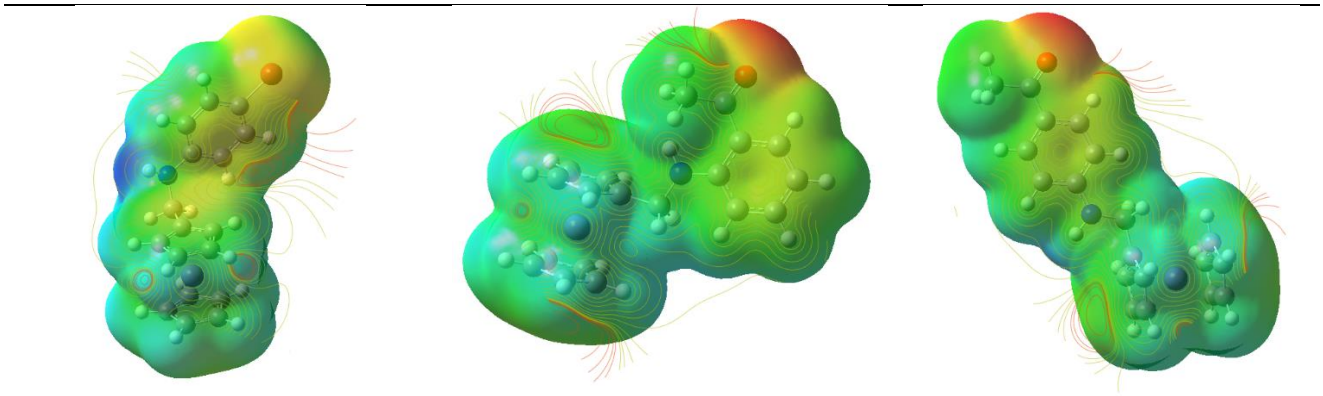
			N2	0.964	-0.405
			O	-0.174	-0.304

5. Molecular electrostatic potential map (MESP)

Figure .2 Molecular electrostatic potential map (MESP)

A16	A17	A18
		
A23	A24	A25
		
A13	A14	A15
		
A26	A27	A28
		

A1	A2	A3
		
A21	A22	A5
		
A6	A7	A8
		
A9	A10	A11
		
A12	A19	A20



6. NMR Spectrum Analysis

6.1. ^{1}H NMR

Table 9. Calculated ^{1}H NMR chemical shift (ppm) of A16, A17 and A18 $CDCl_3$ by experimental method and DFT/B3LYP methods

A16				A17				A18			
Atom	Exp.	a	b	Atom	Exp.	a	b	Atom	Exp.	a	b
H5	4.24	3.25	2.76	H5	4.18	3.98	3.30	H5	4.20	3.74	3.32
H2	4.3	3.53	3.00	H2	4.24	3.84	3.30	H2	4.24	3.74	3.32
H4	4.24	3.53	3.00	H3	4.18	3.84	3.30	H3	4.20	3.93	3.32
H3	4.24	3.69	3.13	H4	4.18	3.84	3.30	H4	4.20	3.93	3.32
H7	4.29	4.02	3.13	H7	4.19	3.84	3.30	H7	4.21	3.93	3.23
H6	4.29	3.53	3.64	H6	4.19	3.68	2.98	H6	4.21	3.59	2.85
H10	4.29	3.53	3.00	H10	4.19	3.84	3.30	H10	4.21	3.93	3.23
H9	4.29	3.53	3.13	H9	4.19	3.84	3.30	H9	4.21	3.81	3.32
H8	4.29	3.69	3.13	H8	4.19	3.84	3.30	H8	4.21	3.81	3.32
H11	4.16	4.23	3.25	H11	4.05	4.11	3.30	H11	4.08	4.08	3.32
H12	8.25	9.53	9.48	H12	4.16	3.14	2.34	H12	4.68	3.47	2.75
H14	8.22	8.50	8.10	H17	7.56	6.61	5.93	H17	6.57	6.28	5.61
H17	6.9	6.69	6.14	H16	7.28	7.20	6.65	H16	8.13	8.09	7.75
H16	7.46	7.30	6.81	H13	7.28	7.53	6.99	H13	8.10	6.45	5.79
H15	6.67	6.62	6.26	H15	7.35	7.74	7.31	H14	6.57	8.51	7.89

a : 6-311++ G(d, p) , b: LanL2DZ

Table 10 Calculated 1H NMR chemical shift (ppm) of A23, A24 and A25 CDCl3 by experimental method and DFT/B3LYP methods

A23			A24				A25			
Atom	a	b	Atom	Exp.	a	b	Atom	Exp.	a	b
H13a	1.92	0.39	H13a	1.8	1.74	1.31	H13a	1.89	1.22	0.46
H13b	1.70	1.10	H13b	1.8	1.86	1.31	H13b	1.89	1.71	0.97
H13c	1.48	1.35	H13c	1.8	1.57	1.22	H13c	1.89	1.71	1.21
H9	2.98	2.35	H9	4.19	3.73	3.15	H9	4.19	4.21	3.95
H5	3.52	2.87	H5	3.95	3.24	2.74	H5	4.09	3.46	2.93
H2	3.91	3.20	H2	3.95	3.83	3.15	H2	4.09	3.53	3.18
H4	3.59	2.95	H4	3.95	3.83	3.15	H4	4.09	3.83	3.39
H3	3.91	3.11	H3	3.95	3.73	3.24	H3	4.09	3.53	3.26
H7	3.91	3.20	H7	4.19	3.73	3.15	H7	3.96	3.40	2.93
H1	3.91	3.41	H1	3.95	3.99	3.15	H1	4.09	3.53	3.01
H8	3.52	2.95	H8	4.19	3.83	3.15	H8	3.96	3.62	3.18
H6	4.27	3.97	H6	4.19	3.83	2.48	H6	3.96	2.91	2.42

H11a	5.90	5.15	H11a	4.67	4.81	3.95	H11a	4.69	3.53	2.93
H11b	2.93	2.09	H11b	4.67	4.33	3.61	H11b	4.69	5.58	5.26
H15	6.73	6.11	H15	7.29	8.18	6.90	H15	7.10	7.42	6.79
H18	7.93	755	H19	8.18	9.00	7.78	H19	7.23	6.76	6.06
H17	7.49	7.00	H17	7.90	8.18	7.93	H18	8.23	8.28	7.80
H16	7.29	6.94	H16	7.53	7.56	7.11	H16	8.21	8.55	8.06

a : 6-311++ G(d, p) , b: LanL2DZ

Table 11 Calculated ^{1}H NMR chemical shift (ppm) of A13, A14 and A15 $CDCl_3$ by experimental method and DFT/B3LYP methods

A13				A14				A15			
Atom	Exp.	a	B	Atom	Exp.	a	b	Atom	Exp.	a	B
H9	4.03	3.81	3.36	H9	4.22	3.82	3.37	H9	4.03	3.93	3.39
H5	4.3	3.59	2.84	H5	4.27	3.53	2.90	H5	4.23	3.79	3.39
H2	4.3	3.81	3.28	H2	4.27	3.64	3.28	H2	4.23	3.93	3.39
H4	4.3	3.81	3.21	H4	4.27	3.82	3.28	H4	4.23	3.93	3.39
H3	4.3	3.81	3.28	H3	4.27	3.72	3.28	H3	4.23	3.93	3.57
H7	4.23	3.91	3.28	H7	4.20	3.64	3.16	H7	4.13	3.93	3.39
H1	4.3	3.81	3.21	H1	4.27	3.72	3.16	H1	4.23	4.06	3.39
H8	4.23	3.91	3.28	H8	4.20	3.64	3.16	H8	4.13	3.93	3.39
H6	4.23	3.81	3.36	H6	4.20	3.64	2.96	H6	4.13	3.93	3.39
H11a	7.42	3.91	3.28	H11a	4.52	3.82	2.96	H11a	4.01	4.14	3.05
H11b	7.42	3.91	3.28	H11b	4.52	4.45	3.52	H11b	4.01	3.64	2.96
H15	6.71	6.54	6.15	H15	7.23	6.88	6.41	H14	6.71	7.66	7.14
H14	6.73	7.44	6.85	H13	6.87	6.94	6.26	H13	7.42	6.69	6.01
H17	6.95	6.54	5.93	H17	7.26	6.42	5.71	H17	7.44	6.96	6.38
H16	6.75	7.44	6.79	H16	7.28	7.06	6.54	H16	7.47	7.66	7.14
H12	5.03	4.09	3.21	H12	3.98	3.30	2.68	H12	4.53	4.68	4.57

a : 6-311++ G(d, p) , b: LanL2DZ

Table 12. Calculated ^{1}H NMR chemical shift (ppm) of A26, A27 and A28 by experimental method and DFT/B3LYP methods

A26			A27				A28			
Atom	a	b	Atom	Exp.	a	b	Atom	Exp.	a	B
H13a	1.14	0.43	H13a	1.77	1.58	1.03	H13a	1.82	1.59	0.98
H13b	1.45	0.59	H13b	1.77	1.52	0.62	H13b	1.82	1.71	0.98
H13c	1.66	1.10	H13c	1.77	1.08	0.33	H13c	1.82	1.29	0.51
H9	4.36	3.94	H9	3.96	2.91	2.41	H9	4.09	3.00	3.80
H5	3.57	3.01	H5	3.98	3.47	2.86	H5	4.12	3.60	3.04
H2	3.57	3.16	H2	3.98	3.82	3.28	H2	4.12	3.60	3.30
H4	3.84	3.40	H4	3.98	3.61	2.97	H4	4.12	3.60	3.47
H3	3.57	3.25	H3	3.98	3.61	3.16	H3	4.12	3.60	3.37
H7	3.38	2.91	H7	3.99	3.76	3.22	H7	4.11	3.74	3.11
H1	3.57	3.01	H1	3.98	3.90	3.40	H1	4.12	3.93	3.11
H8	3.57	3.16	H8	3.99	3.61	3.04	H8	4.11	3.60	3.30
H6	3.07	2.42	H6	3.96	4.16	3.90	H6	4.09	4.12	2.56
H11a	5.93	5.28	H11a	4.63	5.98	5.26	H11a	4.66	5.84	2.91
H11b	3.57	2.91	H11b	4.63	3.24	2.51	H11b	4.66	3.60	5.12
H18	7.25	6.82	H18	7.38	7.49	7.07	H18	7.65	7.61	7.18
H17	7.37	6.89	H17	7.44	7.65	7.14	H16	7.12	8.04	7.50
H16	7.87	7.26	H15	7.32	6.82	6.12	H15	7.14	7.61	7.02
H19	6.64	5.98	H19	7.51	7.41	6.90	H19	7.11	6.77	6.12

a : 6-311++ G(d, p) , b: LanL2DZ

Table 13. Calculated ^{1}H NMR chemical shift (ppm) of A1, A2 and A3 $CDCl_3$ by experimental method and DFT/B3LYP methods

A1				A2				A3			
Atom	Exp.	a	b	Atom	Exp.	a	b	Atom	Exp.	a	b
H6	4.03	3.18	2.61	H6	4.04	3.02	2.66	H6	4.18	4.51	4.14
H9	4.03	4.21	3.80	H9	4.04	3.79	3.32	H7	4.10	3.76	3.24
H7	4.03	3.59	3.06	H8	4.04	3.79	3.21	H8	4.10	3.51	2.99
H8	4.03	3.70	3.27	H7	4.04	3.65	2.90	H9	4.18	3.13	2.61
H5	4.08	3.59	3.13	H5	4.08	3.65	3.21	H1	4.16	3.64	3.09
H1	4.08	3.59	3.06	H1	4.08	3.56	2.99	H2	4.16	3.76	3.24
H2	4.08	3.88	3.50	H2	4.08	3.79	3.32	H3	4.16	3.76	3.24
H3	4.08	3.59	3.36	H3	4.08	3.79	3.32	H4	4.16	3.76	3.38
H4	4.08	3.59	3.27	H4	4.08	3.65	3.32	H5	4.16	3.93	3.48
H11	4.61	5.86	2.66	H11	4.61	4.79	4.19	H11	4.89	4.14	5.07
H19	6.97	6.61	5.94	H20	6.97	7.44	6.74	H17	6.93	6.25	5.56
H18	7.28	7.28	6.88	H19	7.30	7.68	7.011	H16	7.13	7.02	6.56
H17	7.30	7.55	7.12	H18	7.33	7.44	6.85	H15	7.16	7.31	6.92
H16	7.32	7.70	7.23	H17	7.35	7.44	6.74	H14	7.18	7.71	7.20
H15	7.26	7.39	6.88	H16	7.29	6.76	6.46	H13	6.91	7.71	7.11
H13	1.75	1.62	0.85	H13	1.96	2.50	1.96	H20	7.26	7.97	7.50
-	-	-	-	H14	1.02	1.10	0.84	H21	7.13	7.51	7.01
-	-	-	-	-	-	-	-	H22	7.16	7.51	7.01
-	-	-	-	-	-	-	-	H23	7.18	7.02	6.56
-	-	-	-	-	-	-	-	H24	7.29	7.02	6.27

a : 6-311++ G(d, p) , b: LanL2DZ

Table 14. Calculated ^{1}H NMR chemical shift (ppm) of A21 and A22 $CDCl_3$ by experimental method and DFT/B3LYP methods

A21				A22			
Atom	Exp.	a	b	Atom	Exp.	a	B
H13a	1.88	1.42	0.47	H9	4.14	3.76	2.71
H13b	1.98	0.99	0.31	H5	4.18	3.49	2.93
H13c	1.88	0.99	0.01	H2	4.18	3.76	3.25
H9	4.10	3.83	3.27	H4	4.18	3.76	3.25
H5	4.21	3.51	3.00	H3	4.18	3.76	3.25
H2	4.21	3.83	3.37	H7	4.21	3.49	3.07
H4	4.21	3.83	3.27	H1	4.18	3.76	3.25
H3	4.21	3.83	3.37	H8	4.21	3.76	2.71
H7	4.15	3.83	3.37	H6	4.14	3.35	3.25
H1	4.21	3.83	3.37	H11a	4.60	3.76	3.62
H8	4.15	3.83	3.37	H11b	4.60	4.80	3.75
H6	4.10	3.96	3.49	H12	7.62	5.42	5.88
H11a	2.01	3.83	3.27	H13	6.97	7.53	6.66
H11b	2.01	4.61	3.94	H14	6.86	7.63	7.00
H12	7.36	5.56	5.27	H15	7.27	7.14	6.66
H15	7.00	7.02	6.14	H16	6.86	7.63	7.00
H18	6.97	7.53	6.97	H17	6.97	7.32	6.37
H17	6.94	7.02	6.57	H20	7.73	7.53	7.69
H16	6.97	7.47	6.84	H21	7.54	7.32	7.00
H19	7.00	7.31	6.43	H22	7.44	7.42	6.88
				H23	7.54	7.14	6.29
				H24	7.73	7.32	6.66

a : 6-311++ G(d, p) , b: LanL2DZ

6.2. ^{13}C NMR

Table 15. Calculated ^{13}C NMR chemical shift (ppm) of A16, A17 and A18 $CDCl_3$ by experimental method and DFT/B3LYP methods

A16			A17			A18			
Atom	**a**	**b**	**Atom**	**a**	**b**	**Atom**	**Exp.**	**a**	**b**
C6	74.08	64.22	C6	72.45	62.07	C1	84.37	87.50	79.90
C10	72.21	60.30	C10	72.36	60.35	C2	68.43	73.46	63.57
C9	71.58	59.10	C9	72.36	59.89	C3	68.31	73.09	61.74
C8	72.21	59.10	C8	72.24	59.89	C4	68.31	73.09	61.74
C7	72.29	60.09	C7	74.14	60.35	C5	68.31	73.46	63.57
C5	69.95	58.51	C5	74.97	63.88	C6	68.70	73.46	61.64
C4	70.87	60.30	C4	73.26	60.99	C8	68.70	72.79	59.98
C3	72.21	60.22	C3	72.55	60.99	C7	68.70	72.94	60.39
C2	69.94	59.48	C2	73.86	63.88	C9	68.70	72.79	59.98
C1	90.72	81.89	C1	88.49	80.54	C10	68.70	72.94	60.39
C11	42.55	36.96	C11	45.98	40.13	C11	42.23	45.32	40.64
C12	152.06	137.95	C12	154.07	136.68	C12	152.92	156.47	141.05
C13	138.56	134.32	C13	106.36	96.53	C13	111.04	109.42	100.02
C14	133.62	121.20	C14	157.40	147.83	C14	126.56	132.19	120.97
C15	117.92	109.19	C15	116.13	104.22	C15	152.92	144.65	138.69
C16	141.17	130.52	C16	133.25	122.95	C16	126.56	131.23	119.10
C17	115.77	105.66	C17	123.27	111.57	C17	111.04	115.16	103.86

a : 6-311++ G(d, p) , b: LanL2DZ

Table 16. Calculated ^{13}C NMR chemical shift (ppm) of A23, A24 and A25 $CDCl_3$ by experimental method and DFT/B3LYP methods

A23			A24			A25			
Atom	a	b	Atom	a	b	Atom	Exp.	a	b
C6	74.77	65.79	C6	73.21	63.45	C1	69.60	71.31	59.69
C10	89.66	80.45	C10	92.49	79.30	C2	69.60	71.73	59.92
C9	73.74	64.42	C9	71.47	62.44	C3	69.60	71.84	60.09
C8	70.33	59.01	C8	71.23	60.49	C4	69.60	72.29	60.36
C7	73.08	62.30	C7	71.98	62.04	C5	69.60	73.45	61.44
C5	73.74	61.90	C5	75.74	61.03	C6	76.74	72.72	63.15
C4	72.23	59.70	C4	72.21	59.82	C8	76.74	72.77	62.03
C3	71.68	59.45	C3	71.74	59.96	C7	76.74	70.91	59.28
C2	71.68	59.77	C2	71.98	60.23	C9	68.63	74.91	66.01
C1	72.32	60.65	C1	72.21	60.05	C10	77.54	88.71	80.0
C11	50.62	45..21	C11	54.94	49.12	C11	48.73	52.67	47.02
C12	173.48	167.45	C12	174.07	167.29	C12	148.44	172.12	167.16
C13	23.79	16.89	C13	24.22	17.64	C13	22.93	23.26	16.34
C14	145.31	132.91	C14	152.23	136.39	C14	143.87	157.63	144.36
C15	140.34	129.66	C15	131.92	126.49	C15	132.54	132.36	120.68
C16	137.13	127.88	C16	131.92	122.56	C16	131.90	130.31	118.42
C17	132.77	122.21	C17	123.83	113.80	C17	133.89	153.18	143.76
C18	130.86	121.49	C18	155.30	145.45	C18	82.26	129.30	117.85
C19	157.39	147.97	C19	128.47	123.16	C19	129.45	135.70	124.92

a : 6-311++ G(d, p) , b: LanL2DZ

Table 17. Calculated ^{13}C NMR chemical shift (ppm) of 13, 14 and 15 $CDCl_3$ by experimental method and DFT/B3LYP methods

A13				A14				A15			
Atom	Exp.	a	b	Atom	Exp	a	b	Atom	Exp.	a	b
C6	68.17	74.39	63.86	C6	68.26	72.25	61.82	C6	68.3	69.29	58.97
C10	84.98	87.42	80.12	C10	85.34	90.81	83.09	C10	84.86	92.13	83.80
C9	67.69	74.39	63.86	C9	68.26	73.05	62.13	C9	68.26	73.14	62.43
C8	68.17	73.84	61.58	C8	68.26	72.93	61.82	C8	68.3	71.14	59.30
C7	68.17	73.84	61.58	C7	68.26	71.74	59.75	C7	68.3	72.23	60.25
C5	76.76	74.87	61.93	C5	68.73	73.18	61.04	C5	77.76	72.23	60.67
C4	76.76	72.69	60.30	C4	68.73	71.98	60.33	C4	77.76	72.95	60.89
C3	76.76	72.14	60.0	C3	68.73	71.74	59.95	C3	77.76	73.29	60.56
C2	76.76	72.14	60.0	C2	68.73	71.36	59.75	C2	77.76	72.23	59.80
C1	76.76	72.69	60.30	C1	68.73	71.88	59.75	C1	77.76	72.23	59.80
C11	42.29	44.87	40.30	C11	43.10	44.04	37.31	C11	42.66	43.03	38.12
C12	118.12	155.27	140.77	C12	119.85	153.59	137.10	C12	151.04	157.28	141.29
C13	134.32	98.02	85.30	C13	117.29	117.27	106.65	C13	112.18	112.49	102.23
C14	132.76	139.65	129.26	C14	112.87	118.89	105.22	C14	134.12	141.10	130.42
C15	116.61	118.79	109.18	C15	120.7	125.66	115.46	C15	77.6	101.24	88.65
C16	116.61	138.49	126.39	C16	130.07	134.16	123.45	C16	133.76	139.40	129.03
C17	95.62	110.69	101.69	C17	114.84	122.10	108.94	C17	111.08	118.65	106.84
C18	150.06	121.25	113.20	C18	148.34	123.18	115.25	C18	151.04	126.52	117.38

a : 6-311++ G(d, p) , b: LanL2DZ

Table 18. Calculated ^{13}C NMR chemical shift (ppm) of A26, A27 and A28 by experimental method and DFT/B3LYP methods.

A26			A27			A28			
Atom	a	b	Atom	a	b	Atom	Exp.	a	b
C6	73.34	64.4	C6	74.85	65.37	C6	68.61	74.45	64.12
C10	89.24	80.56	C10	88.96	79.82	C10	82.31	89.33	80.72
C9	74.92	65.70	C9	73.12	64.22	C9	68.61	74.38	65.18
C8	73.09	62.05	C8	71.17	59.57	C8	76.65	70.77	61.85
C7	69.81	58.82	C7	72.44	62.17	C7	76.65	72.47	59.27
C5	74.01	62.05	C5	73.19	61.46	C5	68.42	73.38	62.02
C4	71.88	60.67	C4	71.43	59.57	C4	68.42	71.42	60.41
C3	71.36	59.82	C3	71.76	59.57	C3	68.42	71.77	60.13
C2	71.24	59.38	C2	72.92	59.87	C2	68.42	71.77	59.91
C1	71.74	59.92	C1	73.00	60.59	C1	68.42	72.12	59.72
C11	50.31	45.13	C11	51.45	45.80	C11	48.73	51.72	46.29
C12	172.69	166.52	C12	172.61	165.98	C12	168.95	174.29	168.66
C13	23.65	17.03	C13	23.48	16.98	C13	22.88	23.75	17.24
C14	154.99	142.14	C14	152.01	138.97	C14	132.54	155.49	141.22
C15	119.36	105.81	C15	142.01	132.38	C15	129.5	134.91	123.18
C16	140.35	130.40	C16	119.87	105.44	C16	129.5	140.22	130.20
C17	132.36	121.49	C17	138.04	127.40	C17	132.54	114.95	102.20
C18	136.76	125.62	C18	134.49	124.24	C18	133.41	139.62	129.33
C19	137.81	128.0	C19	137.59	126.25	C19	118.06	136.92	126.29
C20	120.00	110.84	C20	121.91	112.69	C20	146.74	123.84	116.42

a : 6-311++ G(d, p) , b: LanL2DZ

Table 19. Calculated ^{13}C NMR chemical shift (ppm) of A1, A2 and A3 $CDCl_3$ by experimental method and DFT/B3LYP methods

A1				A2				A3			
Atom	Exp.	a	b	Atom	Exp	a	b	Atom	Exp.	a	b
C1	68.55	73.32	62.11	C1	68.57	73.35	61.88	C1	68.69	72.07	59.85
C2	68.55	72.20	60.52	C2	68.57	72.754	59.93	C2	68.69	71.58	59.93
C3	68.55	70.99	59.91	C3	68.57	72.52	60.16	C3	68.69	71.15	60.11
C4	68.55	70.53	59.81	C4	68.57	72.44	60.16	C4	68.69	70.97	60.28
C5	68.55	71.20	59.98	C5	68.57	71.81	60.22	C5	68.69	73.01	61.88
C6	77.16	75.04	64.86	C6	77.16	72.52	63.49	C6	77.16	73.89	63.95
C7	77.58	72.55	61.61	C7	77.59	71.46	61.24	C8	77.58	71.85	60.87
C8	77.58	70.03	58.70	C8	77.59	72.75	59.70	C7	77.58	70.89	58.95
C9	77.16	73.65	64.81	C9	77.16	74.46	65.14	C9	77.16	75.96	66.47
C10	83.33	90.35	81.98	C10	83.33	89.186	80.44	C10	83.26	88.78	80.58
C11	48.40	51.05	45.11	C11	48.65	53.10	46.51	C11	49.88	53.59	48.29
C12	169.74	172.95	169.16	C12	173.17	178.18	174.41	C12	126.75	151.68	138.54
C13	22.83	23.40	17.35	C13	27.87	31.16	26.39	C13	128.33	132.72	121.56
C14	127.87	151.99	137.31	C14	9.67	8.78	7.18	C14	128.97	134.03	122.95
C15	128.55	133.80	122.75	C15	127.86	149.43	134.40	C15	128.97	130.91	119.89
C16	129.43	134.32	123.74	C16	128.77	137.83	125.85	C16	128.97	133.38	122.52
C17	129.43	132.41	121.89	C17	129.46	132.61	121.01	C17	128.33	136.16	124.71
C18	129.43	133.30	122.75	C18	129.46	131.29	119.39	C18	170.07	176.85	171.19
C19	128.55	136.98	125.62	C19	129.46	134.26	121.56	C19	127.68	142.50	127.18
-	-	-	-	C20	128.77	132.22	118.82	C20	128.77	136.49	126.50
-	-	-	-	-	-	-	-	C21	129.50	132.78	121.65
-	-	-	-	-	-	-	-	C22	129.50	134.99	123.33
-	-	-	-	-	-	-	-	C23	129.50	131.02	119.89
-	-	-	-	-	-	-	-	C24	128.77	134.53	123.71

a : 6-311++ G(d, p) , b: LanL2DZ

Table 20. Calculated ^{13}C NMR chemical shift (ppm) of A21 and A22. $CDCl_3$ by experimental method and DFT/B3LYP methods

A21				A22			
Atom	Exp.	A	b	Atom	Exp.	a	b
C6	69.82	73.63	62.98	C1	68.92	72.14	60.10
C10	81.42	88.30	82.08	C2	68.92	72.60	60.10
C9	69.82	75.66	65.51	C3	68.92	73.01	60.10
C8	70.01	72.02	60.31	C4	68.92	73.01	59.94
C7	70.01	72.96	62.29	C5	68.92	73.84	61.56
C5	69.90	73.63	61.74	C6	68.92	74.02	62.52
C4	69.90	72.39	60.31	C7	70.05	73.31	59.69
C3	69.90	72.08	60.31	C8	70.05	71.77	60.41
C2	69.90	71.55	59.93	C9	68.92	76.74	65.09
C1	69.90	72.02	60.03	C10	81.41	86.53	80.78
C11	54.32	59.38	50.27	C11	51.71	61.41	53.48
C12	175.99	180.49	178.62	C12	148.43	157.52	141.48
C13	21.13	20.80	14.34	C13	113.62	118.58	104.88
C14	148.74	157.19	139.41	C14	129.34	134.83	123.00
C15	114.54	116.84	103.60	C15	120.03	125.04	112.58
C16	129.43	134.19	123.09	C16	129.34	134.11	122.72
C17	121.0	124.45	111.82	C17	113.62	120.83	103.86
C18	129.43	134.19	122.95	C18	166.57	179.92	173.60
C19	114.54	119.82	10.99	C19	133.08	138.13	123.83
-	-	-	-	C20	127.26	135.46	126.34
-	-	-	-	C21	128.30	132.64	121.39
-	-	-	-	C22	132.16	136.13	125.35

-	-	-	-	C23	128.30	131.37	120.72
-	-	-	-	C24	127.26	134.11	122.46

a : 6-311++ G(d, p) , b: LanL2DZ

7. Vibrational analysis.

Table 21. The observed FT-IR of studied compounds and its calculated wavenumbers (in cm^{-1}) with their probable assignments

compound	Experimental frequencies (cm-1)	Calculated frequencies		**Vibrational Assignments**
	FT-IR	**a**	**b**	
A16	1610	1675	1657	$\upsilon_s(C_{13}C_{14}\ C_{16}\ C_{17})sy+\ \upsilon_s(NC_{12})sy+\ \upsilon_s(O_1N_2O_2)Asy$
		1612	1547	$\upsilon_s(N_2O_2)sy+\ \pi_s\ (NH)sy$
		1573	1446	$\upsilon_m(O_1N_2O_2)Asy+\ \pi_m\ (NH)sy+\ \upsilon_m(C_{15}\ C_{16})sy$
	1506.4	1546	1466	$\upsilon_s(O_1N_2O_2)Asy+\ \pi_s\ (N_1C_{12})sy+\upsilon_s(C_{13}C_{14}\ C_{16}\ C_{17})sy$
		1380	1380	$\upsilon_m(C_{17}\ N_2)sy$
		1274	1274	$\upsilon_s(O_1N_2O_2)sy$
		874	874	$\pi_w\ (O_1N_2O_2)sy$
A17	1535.73	1660	1668	$\upsilon_s(C_{13}C_{14}\ C_{16}\ C_{17})sy+\ \upsilon_s(NC_{12})sy+\ \upsilon_s(O_1N_2O_2)Asy$
	1258.3	1581	1470	$\upsilon_s(O_1N_2O_2)Asy+\ \pi_s\ (NH)sy+\ \upsilon_s(C_{14}\ C_{15}\ C_{16})asy+\ \upsilon_s(C_{12}\ C_{13})sy$
		-	1428	$\upsilon_m(O_1N_2O_2)Asy+\ \upsilon_m(N_1C_{12})sy$
		1364	1271	$\upsilon_s(C_{16}\ N_2)sy+\ \upsilon_s(N_1C_{12})sy+\ \alpha_s\ (C_{11}H_{11a}H_{11b})asy$
		845	799	$\pi_w\ (O_1N_2O_2)sy$
A18	1611.9	1643	1660	$\upsilon_s(C_{13}C_{14}\ C_{16}\ C_{17})sy$
		1626	1623	$\upsilon_m(O_1N_2O_2)Asy+\ \upsilon_m(C_{17}C_{12}\ C_{13}+C_{14}\ C_{15}C_{16})asy+\ \pi_m\ (NH)sy$
		1559	1423	$\upsilon_s(O_1N_2O_2)Asy+\pi_s\ (NH)sy$
		1543	1388	$\upsilon_s(O_1N_2O_2)Asy+\pi_s\ (NH)sy+\ \upsilon_s(N_1C_{12})sy$
	1286	1345	1280	$\upsilon_s(C_{15}\ N_2)sy+\ \upsilon_s(O_1N_2O_2)sy$
		1126	1214	$\upsilon_s(C_{15}\ N_2)sy$
		879	802	$\pi_w\ (O_1N_2O_2)sy$
A23		1731	1637	$\upsilon_s(C_{12}O_1)sy$
		1586	1519	$\upsilon_s(O_1N_2O_2)Asy+\ \upsilon_s(C_{15}\ C_{16}\ C_{17})asy$
		1387	1423	$\alpha_m\ (C_{13}H_{13a}H_{13b}H_{13c})asy+\ \upsilon_m(N_2C_{19})sy$
		1296	1300	$\upsilon_s(C_{11}N_1C_{12})asy$
		733	812	$\pi_w\ (O_1N_2O_2)sy$
A24	1656	1740	1641	$\upsilon_s(C_{12}O_1)sy$
	1531	1586	1523	$\upsilon_s(O_1N_2O_2)Asy+\ \upsilon_s(C_{14}\ C_{15}\ C_{16})asy$
		1406	1413	$\upsilon_m(N_1C_{12}C_{13})Asy+\ \alpha_m\ (C_{13}H_{13a}H_{13b}H_{13c})asy$
		1371	1278	$\upsilon_s(C_{18}N_2)sy+\ \upsilon_s(N_1C_{12})sy$
		855	821	$\pi_w\ (O_1N_2O_2)sy$
A25	1659	1722	1624	$\upsilon_s(C_{12}O_1)sy$
	1593	1576	1633	$\upsilon_s(O_2N_2O_3)Asy+\ \upsilon_s(C_{19}C_{14}\ C_{15}+C_{18}\ C_{17}C_{16})asy$
		1364	1274	$\upsilon_m(N_2C_{17})sy+\ \upsilon_s(C_{11}N_1C_{12})asy$
		1284	1211	$\upsilon_s(N_1C_{12}C_{13})Asy$
		868	815	$\pi_w\ (O_1N_2O_2)sy$
A13	2208.3	2307	2237	$\upsilon_m\ (C_{13}\text{-}C_{18}N_2)sy$
	1610	1642	1656	$\upsilon_s(C_{12}C_{13}C_{14}\ C_{15}\ C_{16}\ C_{17})sy$
		1205	1215	$\upsilon_w(C_{13}C_{18})sy$
A14	2223.8	2329	2259	$\upsilon_m\ (C_{14}\text{-}C_{18}N_2)sy$
		1639	1655	$\upsilon_s(C_{12}C_{13}C_{14}\ C_{15}\ C_{16}\ C_{17})sy$

	1507.9	1182	1200	$\upsilon_w(C_{14}C_{18})sy$
A15	2210.6	2318	2248	$\upsilon_m(C_{15}\text{-}C_{18}N_2)sy$
	1603.7	1651	1667	$\upsilon_s(C_{12}C_{13}C_{14}\ C_{15}\ C_{16}\ C_{17})sy$
		1198	1219	$\upsilon_w(C_{15}C_{18})sy$
A26	2210.6	2332	2265	$\upsilon_w(C_{15}\text{-}C_{20}N_2)sy$
	1602.3	1728	1646	$\upsilon_s(C_{12}O_1)sy$
		1632	1637	$\upsilon_s(C_{12}C_{13}C_{14}\ C_{15}\ C_{16}\ C_{17})sy$
		1283	1300	$\upsilon_s(N_1C_{12}C_{13})Asy$
		1221	1235	$\upsilon_w(C_{15}C_{20})sy$
A27	2214.2	2334	2266	$\upsilon_w(C_{16}\text{-}C_{20}N_2)sy$
	1636.1	1722	1636	$\upsilon_s(C_{12}O_1)sy$
		1634	1660	$\upsilon_s(C_{12}C_{13}C_{14}\ C_{15}\ C_{16}\ C_{17})sy$
		1292	1306	$\upsilon_s(N_1C_{12}C_{13})Asy$
		1223	1238	$\upsilon_w(C_{16}C_{20})sy$
A28	2214.2	2332	2264	$\upsilon_w(C_{17}\text{-}C_{20}N_2)sy$
	1636.1	1721	1630	$\upsilon_s(C_{12}O_1)sy$
		1642	1541	$\upsilon_s(C_{14}\ C_{15}\ C_{16}\ C_{17}\ C_{18}C_{19})sy$
		1287	1301	$\upsilon_s(N_1C_{12}C_{13})Asy$
		1219	1227	$\upsilon_w(C_{17}C_{20})sy$
		441	405	$\upsilon_w(C_{12}Cl)sy$
A21	2920	3051	3068	$\upsilon_w(C_{13}H_2)\ sy$
	1666.4	1752	1680	$\upsilon_s(C_{12}O_1)\ sy$
	1593.1	1639	1653	$\upsilon_s(C_{15}\ C_{16}\ C_{17}\ C_{18}\ C_{19}C_{20})sy$
		1340	1387	$\upsilon_s(N_1N_2C_{12})\ sy$
		1331	1356	$\upsilon_s\ (N_2C_{12}C_{13})\ asy + \pi_s(N_2H)\ sy$
		1199	1240	$\upsilon_m(C_{11}N_1N_2)asy$
A22	1651.0	1723	1647	$\upsilon_s(C_{12}O_1)\ sy$
	1595.0	1641	1653	$\upsilon_w(C_{19}C_{20}\ C_{21}\ C_{22}\ C_{23}\ C_{24})sy$
	1579.6	1693	1654	$\upsilon_s(C_{12}\ C_{13}C_{14}C_{15}\ C_{16}\ C_{17})sy$
		1350	1384	$\upsilon_s\ (N_2C_{18}C_{19})\ asy + \pi_m(N_2H)\ sy$
		1158	1244	$\upsilon_m(C_{11}N_1N_2)\ asy$
A1	2360	3243	3302	$\upsilon_w(C_1H_1C_2H_2C_3H_3C_4H_4C_5H_5)sy$
		3240	3296	$\upsilon_m(C_6H_6C_7H_7C_8H_8C_9H_9)sy$
		3232	3283	$\upsilon_w(C_1H_1C_2H_2C_3H_3C_4H_4C_5H_5C_6H_6C_7H_7)asy$
		3194	3233	$\upsilon_w(C_{15}H_{15}C_{16}H_{16}C_{17}H_{17}C_{18}H_{18}C_{19}H_{19})sy$
		3180	3212	$\upsilon_w(C_{15}H_{15}C_{16}H_{16}C_{17}H_{17}C_{18}H_{18}C_{19}H_{19})asy$
		3147	3147	$\upsilon_w(C_{13}H_{13a}H_{13b}H_{13c})asy$
		3128	3156	$\upsilon_w(C_{11}H_{11a}H_{11b})asy$
		3062	3088	$\upsilon_w(C_{11}H_{11a}H_{11b})sy$
	1651	1713	1626	$\upsilon_s(C_{12}O)sy$
	1593	1635	1525	$\upsilon_s(C_{15}C_{16}\ C_{18}C_{19})sy$
		1497	1510	$\upsilon_m(C_6C_{10}\ C_9)sy$
	1494	1486	1509	$\pi_m(C_{13}H_{13a}H_{13b}H_{13c})sy$
		1473	1504	$\pi_m(C_{11}H_{11a}H_{11b})sy$
	1387	1371	1416	$\upsilon_m(C_6C_{10}\ C_9\ C_1C_5\ C_4)asy + \upsilon_m(C_{12}N)\ sy$
	1284	1294	1306	$\upsilon_s\ (C_{12}C_{11}N)\ asy + \alpha_s(C_{11}H_{11a})asy$
		1211	1225	$\alpha_m(C_{11}H_{11a}H_{11b})asy + \upsilon_m(C_6C_{10}C_9)asy$
	1023	1133	1122	$\upsilon_w(C_1C_2\ C_3\ C_4C_5)sy$
		1043	1048	$\upsilon_m(C_6C_7\ C_8\ C_4C_9)asy$
	806	841	815	$\alpha_m(C_1H_1C_2H_2C_3H_3C_4H_4C_5H_5C_6H_6C_7H_7C_8H_8)asy$
	701	731	726	$\alpha_m(C_1H_1C_2H_2C_3H_3C_4H_4C_5H_5C_6H_6C_7H_7C_8H_8)asy$
	561	506	494	$\upsilon_m(C_6FeC_9FeC_4FeC_5FeC_{10}FeC_1Fe)sy$
	544	479	459	$\upsilon_m(C_6FeC_7FeC_8FeC_9FeC_2FeC_3FeC_4FeC_5Fe)sy$
		461	440	$\upsilon_m(C_6FeC_7FeC_8FeC_9FeC_1FeC_2FeC_3FeC_4FeC_5FeC_{10}Fe)asy$
A2	2159	3232	3282	$\upsilon_w(C_1H_1C_2H_2C_3H_3C_4H_4C_5H_5)asy$

		3230	3281	υ_w ($C_6H_6C_7H_7$ $C_8H_8C_9H_9$)asy
		3178	3226	υ_m($C_{16}H_{16}C_{17}H_{17}$ $C_{18}H_{18}C_{19}H_{19}C_{20}H_{20}$)asy
		3127	3181	υ_w($C_{11}H_{11a}H_{11b}$)asy
		3104	3153	υ_m($C_{13}H_{13a}H_{13b}C_{14}H_{14a}H_{14b}H_{14c}$)asy
		3063	3062	υ_w ($C_{11}H_{11a}H_{11b}$)sy
		3040	3049	υ_m($C_{14}H_{14a}H_{14b}H_{14c}$)sy
		3022	3078	υ_m($C_{13}H_{13a}H_{13b}$)sy
	1654	1722	1618	υ_s($C_{18}O$)sy
	1592	1633	1653	υ_w($C_{16}C_{17}$ $C_{18}C_{19}$)sy
		1525	1636	υ_m(C_{15} C_{16} $C_{17}C_{18}$ $C_{19}C_{20}$)asy
	1493	1501	1539	π_w($C_{14}H_{14a}H_{14b}H_{14c}$)sy+ π_w($C_{13}H_{13a}H_{13b}$)sy
	1469	1438	1402	υ_m(C_6 C_7 C_8C_9)asy
	1432	1418	1415	υ_m ($C_{12}NC_{15}$) asy+ α_m($C_{13}H_{13a}H_{13b}$)asy+ α_m($C_{11}H_{11a}H_{11b}$)sy
	1402	1393	1388	υ_m(C_1C_5 C_4)asy+ υ_m(C_6C_{10} C_9)asy
	1374	1341	1377	α_m($C_{11}H_{11a}H_{11b}$)asy+ α_m($C_{13}H_{13a}H_{13b}$)sy+ υ_m ($C_{12}NC_{15}$) asy
	1262	1271	1281	α_m($C_{11}H_{11a}H_{11b}$)asy+ π_m ($C_6H_6C_7H_7C_8H_8C_9H_9$)sy
		1256	1270	υ_m(C_6C_7 C_8 C_4C_9 C_{10}) sy
		1235	1245	υ_m ($C_{12}NC_{15}$) asy++ π_m ($C_{16}H_{16}C_{17}H_{17}C_{19}H_{19}C_{20}H_{20}$)asy
	1177	1194	1214	π_m ($C_{16}H_{16}C_{17}H_{17}C_{19}H_{19}C_{20}H_{20}$)asy
	1106	1133	1122	υ_m(C_1C_2 C_3 C_4C_5)sy
	1071	1125	1134	υ_m ($C_{11}N$) sy
	1023	1047	1056	π_m ($C_6H_6C_7H_7C_8H_8C_9H_9$)asy
	925	847	810	α_m($C_1H_1C_2H_2C_3H_3C_4H_4C_5H_5C_6H_6C_7H_7C_8H_8C_9H_9C_{16}H_{16}C_{17}H_{17}$ $C_{18}H_{18}C_{19}H_{19}C_{20}H_{20}$)asy
	806	836	791	α_m($C_1H_1C_2H_2C_3H_3C_4H_4C_5H_5$)asy
	701	707	720	α_m($C_{16}H_{16}C_{17}H_{17}$ $C_{18}H_{18}C_{19}H_{19}C_{20}H_{20}$)asy
	613	488	474	υ_m ($C_{10}FeC_6FeC_9FeC_1FeC_4Fe$)sy
	564	479	455	υ_m($C_1FeC_2FeC_3FeC_4FeC_5FeC_6Fe$ $C_7FeC_8FeC_9Fe$)sy
		460	440	υ_m ($C_1FeC_2FeC_3FeC_4FeC_5FeC_6Fe$ $C_7FeC_8FeC_9Fe$)sy
A3	2159	3233	3285	υ_w ($C_6H_6C_7H_7$ $C_8H_8C_9H_9$)asy
		3230	3280	υ_w($C_1H_1C_2H_2C_3H_3C_4H_4C_5H_5$)asy
		3222	3271	υ_w ($C_6H_6C_7H_7C_8H_8C_9H_9$)asy
		3200	3241	υ_w ($C_{20}H_{20}C_{21}H_{21}C_{22}H_{22}C_{23}H_{23}C_{24}H_{24}$)sy
		3197	3235	υ_w ($C_{13}H_{13}C_{14}H_{14}C_{15}H_{15}C_{16}H_{16}C_{17}H_{17}$)sy
		3189	3228	υ_m ($C_{13}H_{13}C_{14}H_{14}C_{15}H_{15}C_{16}H_{16}C_{17}H_{17}$)asy
		3184	3220	υ_m($C_{20}H_{20}C_{21}H_{21}C_{22}H_{22}C_{23}H_{23}C_{24}H_{24}$)asy
		3073	3101	υ_w ($C_{11}H_{11a}H_{11b}$)sy
	1639	1690	1584	υ_s($C_{18}O$)sy
	1594	1639	1651	υ_w($C_{20}C_{21}$ $C_{23}C_{24}$)sy
		1617	1633	υ_m($C_{19}C_{20}C_{21}$ C_{22} $C_{23}C_{24}$)asy
		1525	1525	υ_m($C_{12}C_{13}C_{14}$ C_{15} $C_{16}C_{17}$)asy+ υ_m ($C_{12}NC_{18}$) asy
	1490	1469	1501	π_m ($C_{11}H_{11a}H_{11b}$)sy
	1443	1406	1437	υ_m(C_6C_7 C_8 C_4C_9 C_{10})asy+ α_m ($C_{11}H_{11a}H_{11b}$)sy
	1428	1385	1412	υ_m(C_1C_5 C_4)asy+ α_m ($C_{11}H_{11a}H_{11b}$)sy+ υ_m ($C_{18}NC_{19}$) asy
	1379	1371	1391	υ_m ($C_{12}C_{18}N$) asy+ υ_m(C_6C_{10} C_9)asy
	1369	1321	1373	υ_m($C_{12}C_{13}C_{14}$ C_{15} $C_{16}C_{17}$)asy
	1295	1296	1306	υ_s($C_{19}C_{18}N$) asy+ α_s($C_{11}H_{11a}$)asy
	1281	1270	1284	π_m ($C_6H_6C_7H_7C_8H_8C_9H_9$)sy+ α_m($C_{11}H_{11a}H_{11b}$)sy
	1178	1201	1228	α_w($C_{20}H_{20}C_{21}H_{21}C_{22}H_{22}C_{23}H_{23}C_{24}H_{24}C_{25}H_{25}C_7H_7$ C_8H_8)sy
	1142	1153	1159	υ_m($C_{18}C_{19}$)sy+ υ_m($C_{11}NC_{18}$)asy
	1103	1132	1122	υ_w(C_1C_2 C_3 C_4C_5)sy
	1021	1043	1053	π_m ($C_6H_6C_7H_7$ C_8H_8 C_9H_9)sy
	963	973	982	υ_m ($C_{10}C_{11}N$)sy
	815	832	808	α_m ($C_1H_1C_2H_2C_3H_3C_4H_4C_5H_5$)asy
	769	712	730	α_m ($C_{13}H_{13}C_{14}H_{14}C_{15}H_{15}C_{16}H_{16}C_{17}H_{17}$)asy
	697	501	485	υ_m ($C_{10}FeC_6FeC_9FeC_1FeC_4Fe$)sy
	613	478	455	υ_m ($C_1FeC_2FeC_3FeC_4FeC_5FeC_6Fe$ $C_7FeC_8FeC_9Fe$)sy
	579	469	440	υ_m ($C_1FeC_2FeC_3FeC_4FeC_5FeC_6Fe$ $C_7FeC_8FeC_9Fe$)sy

asy – asymmetric, Sy – symmetric, υ – stretching, π- in plane bending, α- out of plane bending, s - strong, m - medium.

8. Frontier molecular orbitals

Table 22. Frontier molecular orbitals by DFT/B3LYP methods with 6-311++G(d, p) : a; and LanL2DZ basis sets: b.

	6-311++G (d, p)			LanL2DZ		
Molecules	**E_{HOMO} (eV)**	**E_{LUMO} (eV)**	**ΔE**	**E_{HOMO} (eV)**	**E_{LUMO} (eV)**	**ΔE**
A0	-5.442	-0.490	4.952	-5.113	-0.742	4.371
A1	-5.469	-0.925	4.544	-5.475	-0.781	4.693
A2	-5.687	-0.626	5.061	-5.714	-0.762	4.952
A3	-5.442	-1.252	4.191	-5.442	-1.225	4.218
A13	-5.769	-1.415	4.354	-5.761	-1.273	4.487
A14	-5.714	-1.524	4.191	-5.698	-1.367	4.331
A15	-5.905	-1.143	4.762	-5.832	-1.024	4.809
A16	-5.606	-2.640	2.966	-5.640	-2.950	2.690
A17	-5.823	-2.612	3.211	-5.880	-2.916	2.963
A18	-5.987	-2.259	3.728	-6.055	-2.544	3.511
A21	-5.742	-0.653	5.089	-5.717	-0.842	4.876
A22	-5.660	-1.252	4.408	-5.683	-1.318	4.365
A23	-5.578	-2.912	2.667	-5.628	-3.374	2.254
A24	-5.959	-2.721	3.238	-5.973	-3.093	2.880
A25	-5.687	-3.075	2.612	-5.733	-3.398	2.334
A26	-5.551	-2.095	3.456	-5.551	-2.014	3.537
A27	-5.606	-2.095	3.510	-5.630	-2.026	3.603
A28	-5.608	-1.983	3.625	-5.687	-1.926	3.755
A4	-5.448	-2.765	2.683	-	-	-
A5	-5.496	-0.462	5.034	-	-	-
A6	-5.388	-0.489	4.898	-	-	-
A7	-5.496	-0.462	5.034	-	-	-
A8	-5.586	-0.582	5.004	-	-	-
A9	-5.632	-0.634	4.998	-	-	-
A10	-5.551	-0.653	4.898	-	-	-
A11	-5.633	-0.653	4.979	-	-	-
A12	-5.551	-0.68	4.871	-	-	-
A19	-5.632	-1.768	3.864	-	-	-
A20	-5.741	-1.279	4.462	-	-	-

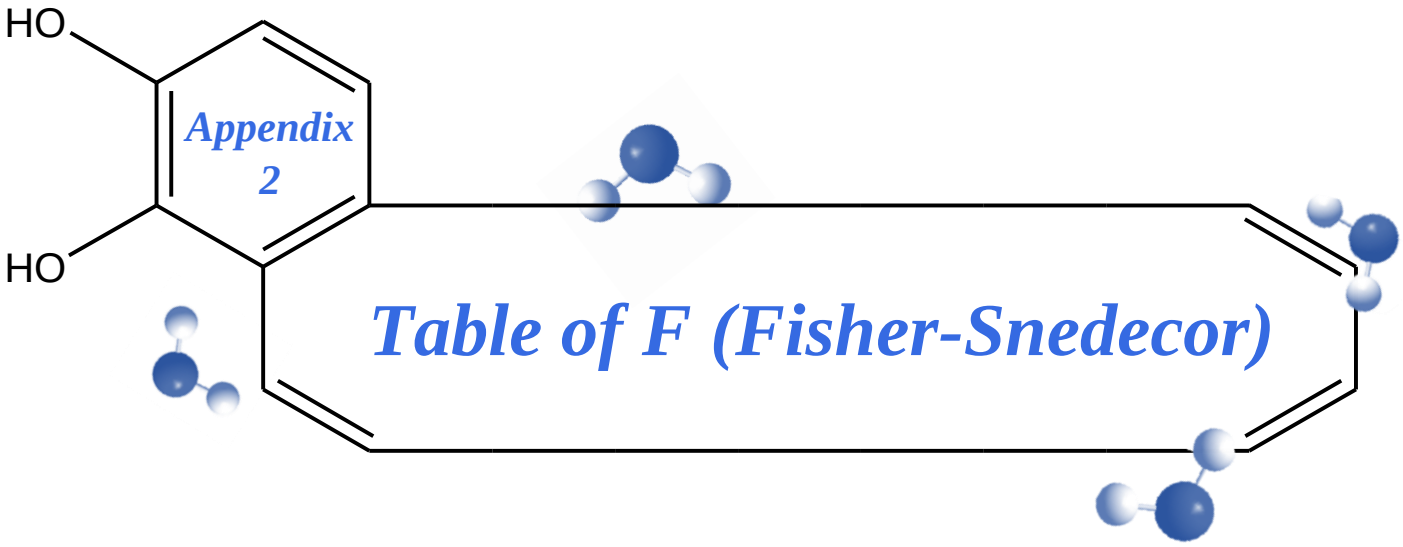


Table of F (Fisher-Snedecor)

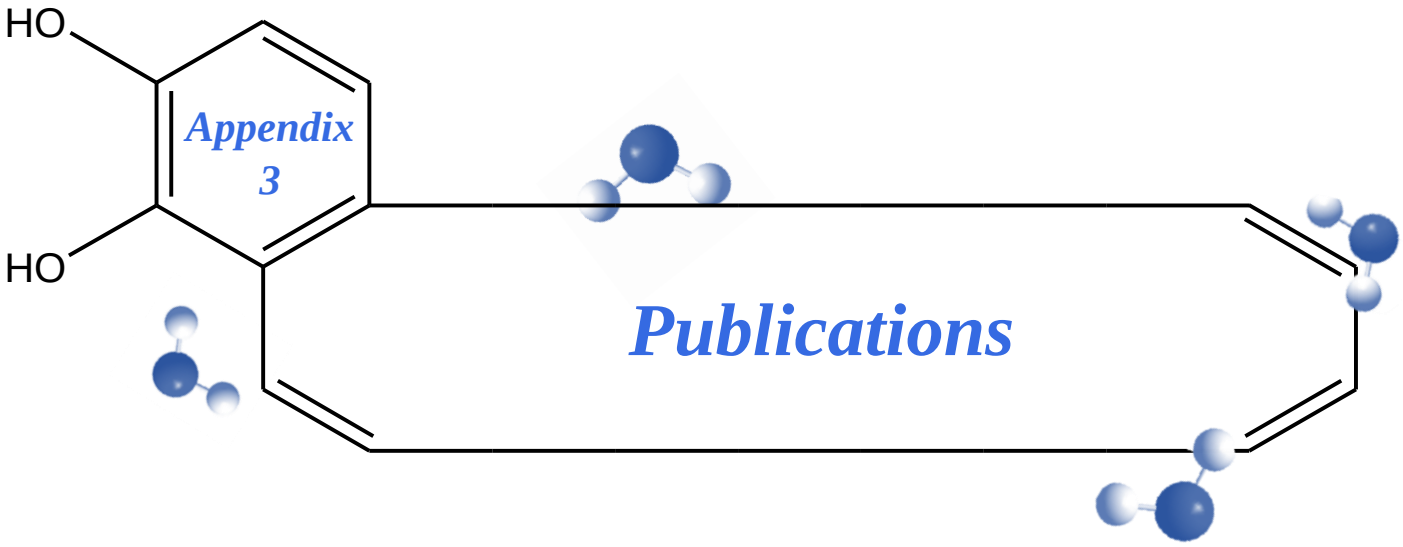
ν_1= degrees of freedom for numerator

ν_1= degrees of freedom for denominator

$F_{95}(\nu_1, \nu_2)$ α=0.05

ν_2/ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.44	199.50	215.70	224.58	230.16	233.98	236.76	238.88	240.54	241.88	243.90	245.94	248.01	249.05	250.09	251.14	252.19	253.25	254.31
2	18.512	19.000	19.164	19.246	19.296	19.329	19.353	19.371	19.384	19.395	19.412	19.429	19.445	19.454	19.462	19.470	19.479	19.487	19.495
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123	8.7855	8.7446	8.7029	8.6602	8.6385	8.6166	8.5944	8.5720	8.5494	8.5264
4	7.7086	6.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	5.9988	5.9644	5.9117	5.8578	5.8025	5.7744	5.7459	5.7170	5.6877	5.6581	5.6281
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725	4.7351	4.6777	4.6188	4.5581	4.5272	4.4957	4.4638	4.4314	4.3985	4.3650
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990	4.0600	3.9999	3.9381	3.8742	3.8415	3.8082	3.7743	3.7398	3.7047	3.6689
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767	3.6365	3.5747	3.5107	3.4445	3.4105	3.3758	3.3404	3.3043	3.2674	3.2298
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881	3.3472	3.2839	3.2184	3.1503	3.1152	3.0794	3.0428	3.0053	2.9669	2.9276
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789	3.1373	3.0729	3.0061	2.9365	2.9005	2.8637	2.8259	2.7872	2.7475	2.7067
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204	2.9782	2.9130	2.8450	2.7740	2.7372	2.6996	2.6609	2.6211	2.5801	2.5379
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962	2.8536	2.7876	2.7186	2.6464	2.6090	2.5705	2.5309	2.4901	2.4480	2.4045
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964	2.7534	2.6866	2.6169	2.5436	2.5055	2.4663	2.4259	2.3842	2.3410	2.2962
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144	2.6710	2.6037	2.5331	2.4589	2.4202	2.3803	2.3392	2.2966	2.2524	2.2064
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458	2.6022	2.5342	2.4630	2.3879	2.3487	2.3082	2.2664	2.2229	2.1778	2.1307
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876	2.5437	2.4753	2.4034	2.3275	2.2878	2.2468	2.2043	2.1601	2.1141	2.0658
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377	2.4935	2.4247	2.3522	2.2756	2.2354	2.1938	2.1507	2.1058	2.0589	2.0096
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943	2.4499	2.3807	2.3077	2.2304	2.1898	2.1477	2.1040	2.0584	2.0107	1.9604
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563	2.4117	2.3421	2.2686	2.1906	2.1497	2.1071	2.0629	2.0166	1.9681	1.9168
19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227	2.3779	2.3080	2.2341	2.1555	2.1141	2.0712	2.0264	1.9795	1.9302	1.8780
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928	2.3479	2.2776	2.2033	2.1242	2.0825	2.0391	1.9938	1.9464	1.8963	1.8432

21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660	2.3210	2.2504	2.1757	2.0960	2.0540	2.0102	1.9645	1.9165	1.8657	1.8117
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419	2.2967	2.2258	2.1508	2.0707	2.0283	1.9842	1.9380	1.8894	1.8380	1.7831
23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201	2.2747	2.2036	2.1282	2.0476	2.0050	1.9605	1.9139	1.8648	1.8128	1.7570
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002	2.2547	2.1834	2.1077	2.0267	1.9838	1.9390	1.8920	1.8424	1.7896	1.7330
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821	2.2365	2.1649	2.0889	2.0075	1.9643	1.9192	1.8718	1.8217	1.7684	1.7110
26	4.2252	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655	2.2197	2.1479	2.0716	1.9898	1.9464	1.9010	1.8533	1.8027	1.7488	1.6906
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501	2.2043	2.1323	2.0558	1.9736	1.9299	1.8842	1.8361	1.7851	1.7306	1.6717
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360	2.1900	2.1179	2.0411	1.9586	1.9147	1.8687	1.8203	1.7689	1.7138	1.6541
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229	2.1768	2.1045	2.0275	1.9446	1.9005	1.8543	1.8055	1.7537	1.6981	1.6376
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107	2.1646	2.0921	2.0148	1.9317	1.8874	1.8409	1.7918	1.7396	1.6835	1.6223
40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240	2.0772	2.0035	1.9245	1.8389	1.7929	1.7444	1.6928	1.6373	1.5766	1.5089
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401	1.9926	1.9174	1.8364	1.7480	1.7001	1.6491	1.5943	1.5343	1.4673	1.3893
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588	1.9105	1.8337	1.7505	1.6587	1.6084	1.5543	1.4952	1.4290	1.3519	1.2539
∞	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799	1.8307	1.7522	1.6664	1.5705	1.5173	1.4591	1.3940	1.3180	1.2214	1.0000



Publications

Send Orders for Reprints to reprints@benthamscience.net

Anti-Cancer Agents in Medicinal Chemistry, 2021, *21*, 1-12

RESEARCH ARTICLE

In Vitro and *In Silico* Determination of some N-ferrocenylmethylaniline Derivatives as Anti-Proliferative Agents against MCF-7 Human Breast Cancer Cell Lines

Nadjiba Zegheb[1], Cherifa Boubekri[2], Touhami Lanez[1,*], Elhafnaoui Lanez[1], Tuba Tüylü Küçükkılınç[3], Esin Öz[4], Ali Khennoufa[1], Saida Khamouli[2] and Salah Belaidi[2]

[1]*VTRS Laboratory, University of El Oued B.P.789, 39000, El Oued, Algeria;* [2]*Department of Material Sciences, University of Biskra B. P. 145 RP, Biskra 07000, Algeria;* [3]*Department of Biochemistry, Faculty of Pharmacy, Hacettepe University, Ankara, Turkey; 4Department of Biochemistry, Faculty of Medicine, Hacettepe University, Ankara, Turkey*

Abstract: ***Background*:** Since the binding of estradiol to its receptor promotes breast cancer cell proliferation (in the ER+ tumours), many molecules targeting this protein have been synthesized to counteract the estradiol action. Ferrocene derivatives have proved their efficiency against hormone-dependent breast cancer cells (MCF-7).

***Objective*:** In this study, we aimed to find new ferrocene derivatives having pharmacochemistry properties as potential drugs against human breast cancer cells.

***Methods*:** A series of 29 N-ferrocenylmethylaniline derivatives A0-A28 were synthesised, and their anti-proliferative activity against both hormone-dependent (MCF-7) and independent (MDA-MB 231) human breast cancer cell lines were performed using the MTT test. Molecular docking and drug-likeness prediction were also performed for the five most active derivatives towards MCF-7. A QSAR model was also developed for the perdition of the anti-proliferative activity against MCF-7 cell lines using molecular descriptors and MLR analysis.

***Results*:** All studied derivatives demonstrated better cytotoxicity against MCF-7 compared to the MDA-MB-231 cell lines, and compounds A2, A9, A14, A17 and A27 were the most potent ones but still less active than the standard anti-cancer drug, crizotinib. The QSAR study revealed good predictive ability, as shown by $R^2_{cv} = 0.848$.

***Conclusion*:** *In vitro* and *in silico* results indicated that derivatives A2, A9, A14, A17, and A27 possess the highest anti-proliferative activity; these results can be used to design more potent N-ferrocenylmethylaniline derivatives as anti-proliferative agents.

ARTICLE HISTORY

Received: November 16, 2020
Revised: April 18, 2021
Accepted: April 26, 2021

DOI:
10.2174/1871520621666210624141712

Keywords: N-ferrocenylmethylaniline, breast cancer, MCF-7, MDA-MB-231, estrogen receptor, progesterone receptor, QSAR, antiproliferative activity.

1. INTRODUCTION

Breast cancer is the most common cancer in women, followed by colorectal cancer and lung cancer, with over 2 million new cases worldwide in 2020. According to the World Health Organization, breast cancer is the leading global cause of cancer-related death among women, with almost 684 996 breast cancer-related deaths in 2020 [1].

Although breast cancer incidence and mortality have seen a decline in recent years, thanks to early detection and increasingly appropriate treatments, breast cancer, unfortunately, remains the leading cause of cancer death in women [2]. Most women with breast cancer will undergo some type of surgical excision, often combined with other treatments, such as radiation therapy, chemotherapy, hormone therapy, and, very recently, biology therapy [3]. Surgery is meant to remove the tumor from the breast and lymph nodes. It is almost always followed by radiotherapy. Biology therapy refers to the treatment with Herceptin®, a monoclonal antibody, for patients overexpressing a specific "protein" (HER2) [4-6].

In the case of ER-breast tumours, chemotherapy is often prescribed to kill cancer cells that could not be removed by other means or that may have already invaded other parts of the body. However, the molecules used in chemotherapy are highly cytotoxic and can reduce mortality by up to 20% [7-9]. The finding of the hormone responsiveness of ER+ breast cancer cases triggered the development of antiestrogens, and molecules, which could compete with estradiol at its receptor and counterbalance its proliferative action.

Thus, for the patients with ER+ tumours, adjuvant treatment with antiestrogens [10, 11], or more recently with aromatase inhibitors, are currently available for their treatment [12]. This hormone therapy appears as a simple and safer alternative to ablative surgery and undiscriminate radio- and chemotherapy for patients having hormone-dependent breast cancer [10, 13, 14], but it causes side effects like hot flashes, vaginal dryness or discharge, tumor flare with bone pain, high calcium level in the blood, and other rare but serious side effects like uterine cancer, blood clots and strokes [15].

The *in vivo* and *in vitro* antiproliferative activity of ferrocene derivatives against Ehrlich ascites tumor (EAT) cells was observed as early as 1985, though their effectiveness was limited by their poor water solubility [16, 17]. Furthermore, ferrocene derivatives have proved their efficiency against hormone-dependent breast cancer cells (MCF-7) that express the estrogen receptor alpha (ERα). Moreover, studies on the potential anti-proliferative effect of compounds containing ferrocene indicated that the presence of the ferrocene moiety is necessary to generate the anti-proliferative effect of these compounds [18, 19].

*Address correspondence to this author at the Department of Chemistry, Faculty of sciences, University of El Oued, P.O. Box: 789, 39000, El Oued, Algeria; Tel/Fax: +213661655550;
E-mail: touhami-lanez@univ-eloued.dz

Herein, we investigate the *in vitro* and *in silico* anti-proliferative activity of a series of 29 N-ferrocenylmethylaniline derivatives as novel cytotoxic agents against MCF-7 and MDA-MB-231 cell lines. We also developed a QSAR model using molecular descriptors and a multiple linear regression method for predicting the activity of N-ferrocenylmethylaniline derivatives as potent anti-human breast cancer.

2. EXPERIMENTS

2.1. Synthesis

The synthesis of N-ferrocenylmethylaniline derivatives (A0-A28) used in this work as potent inhibitory on MCF-7 cell proliferation was synthetized following our previously reported procedure [20-24]. Their molecular structures are shown in Fig. (**1**) and listed in Table **1**.

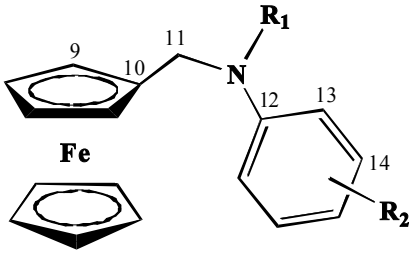


Fig. (1). Structure of N-ferrocenylmethylaniline derivatives.

2.2. MTT assay against MDA-MB-231 and MCF-7

Cell culture and proliferation assays were carried out at the department of biochemistry at the faculty of pharmacy of the Hacettepe University, Turkey. The cancer cell lines MDA-MB-231 (ATCC® HTB-26) and MCF-7 (ATCC® HTB-22) were cultured in Dulbecco's Modified Eagle's Medium (DMEM) including 1% L-glutamine, 1% pen-strep solution and 10% of fetal bovine serum FBS, at 37°C in a humidified atmosphere containing 5% CO2 for 48 hours. The cells were then seeded in 96-well plates at 4000 cells per well in 100 ml of DMEM. After 6 hours, the medium was changed, and DMEM with 1% L-glutamine, 1% pen-strep solution and 2% of FBS was added, and it was left to stand for 8 hours. The tested compounds were then added at different concentrations and incubated for 48 hours at 37°C in the presence of 5% CO2. Crizotinib (Sigma, PZ0191, USA), an approved anti-cancer drug, was also added as a positive control.

After incubation, 10 µL of MTT reagent (5 mg/ml) was added to each well, and the plates were incubated for 2 hours at 37°C under the string. The MTT formazan crystals were dissolved by adding 100 µL DMSO.

Relative cell viability was estimated by comparing absorbance values at 570 nm (for absorbance of MTT formazan) and 690 nm (for the reference wavelength) of treated group and untreated control group using a FLUOstar Omega microplate reader. IC50 values (concentration at which cell proliferation was inhibited by 50%) were calculated using GraphPad Prism software.

2.3. Structure Optimization

Chemical structures of compounds A0-A28 were fully optimized using density functional theory (DFT) and the B3LYP level of theory [25, 26] with 6-311G++(d, p) basis set using Gaussian 09W software [27], and the optimized structures were saved as PDBQT files using ADT.

2.4. Molecular Docking Study

The molecular docking study was performed for the five best active compounds towards MCF-7 using AutoDock 4.2 software with the AutoDock Tools (ADT) interface [28]. Estrogen alpha (ERα) and progesterone receptors (PR) (PDB ID: 2IOG [29] 17 and 1A28 [30], respectively) were chosen as targets for this study (Table **2**). Both receptors were first prepared using Discovery Studio Software [31]. The preparation involved the elimination of all water molecules and cofactors to obtain the raw protein's structure, after which the active site was defined. The polar hydrogen atoms were added along with the charges using ADT.

The docking procedure consisted of two main steps; the first step was the generation of interaction maps of the receptor using the AutoGrid tool, with the grid boxes sized 50 x 50 x 50 Å, spacing 0.5 Å with coordinates X = 24.685, Y = 5.51 and Z = 23.906 for ERα, and 50 x 50 x 50 Å, spacing 0.5 Å with coordinates X= 26.573, Y= 8.374, Z= 64.687 for PR. The second step consisted of the docking of the ligands into the receptor with AutoDock4. Lamarckian genetic algorithm was utilized for docking search with a number of evaluations: 2,500,000, population size: 150. The number of runs was 20 for each ligand. The graphical interface of ADT was used to dock and view the obtained results. Finally, different values of binding energy were obtained, and the best poses with the lower docking energy [32-34] were selected and used in the docking analysis using Protein-Ligand Interaction Profiler [35].

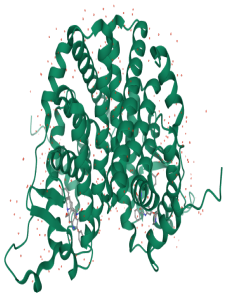
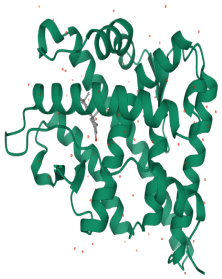
2.5. Drug-likeness Screening

The drug-like nature was predicted for the five potent compounds using the free web-server SwissADME [36, 37] by entering the SMILES of the compounds. The Lipinski rule of five, Ghose, Veber, Egan and Muegge rules were used for this analysis as well as the bioavailability score.

Table 1. Chemical structures of N-ferrocenylmethylaniline derivatives used for anti-proliferative activity study and QASR analysis.

Code	R^1	R^2	Code	R^1	R2	Code	R^1	R^2
A0	H	H	A10	H	4-Cl	A20	H	4-COMe
A1	COMe	4-CN	A11	H	3-Br	A21	NHCOMe	H
A2	COEt	H	A12	H	4-Br	A22	NHCOPh	H
A3	COPh	H	A13	H	2-CN	A23	COMe	2-NO_2
A4	CO-Ph-2-NO_2	H	A14	H	3-CN	A24	COMe	3-NO_2
A5	H	2-Me	A15	H	4-CN	A25	COMe	4-NO_2
A6	H	3-Me	A16	H	2-NO_2	A26	COMe	2-CN
A7	H	2-Et	A17	H	3-NO_2	A27	COMe	3-CN
A8	H	2-Cl	A18	H	4-NO_2	A28	COMe	H
A9	H	3-Cl	A19	H	2-COMe			

Table 2. Target receptors' information chosen for docking studies.

	Estrogen receptors alpha (ERα)	
	PDB ID	2IOG
	Resolution (Å)	1.60
	R-Value Free	0.328
	R-Value	0.212
	Space Group	P 6_5 2 2
	Chains	A
	Organism	Homo sapiens
	Progesterone receptors (PR)	
	PDB ID	1A28
	Resolution (Å)	1.80
	R-Value Free	0.227
	R-Value	0.191
	Space Group	P 1 2_1 1
	Chains	A, B
	Organism	Homo sapiens

2.6. Quantitative Structure–activity Relationship (QSAR)

A data set containing 29 N-ferrocenylmethylaniline derivatives as potent anti-proliferative agents were selected for the QSAR study. The data set was prepared by converting the obtained IC50 values (μM) to the logarithmic scale [pIC50 = -log IC50] as well as the molecular descriptors, which were calculated using Gaussian 09W [27] and HyperChem (8.03) [38] software and the web-server SwissADME [36, 37].

The QSAR properties module from HyperChem (8.03) [38] were used to calculate molar polarizability (Pol), molar volume (MV) and molar weight (MW). The semi-empirical PM3 optimized geometry was used to calculate the heat of formation (HF).

The Quantum Chemical descriptors, dipole moment (DM), energy of frontier orbital's HOMO, LUMO, energetic gap (GapE), atomic net charges (qFe1, qN1, qC9, qC10, qC11, qC12, qC13, qC14 and qC15) and total energy (E), were computed through Gaussian 09W software [27] by using DFT/B3LYP with 6-311G++(d,p) basis set.

Moreover, partition coefficient (logP), the molar refractivity (MR), topological polar surface area (TPSA) and water solubility (logS) were obtained using the free web-server SwissADME [36, 37].

The final step in QSAR model development was model validation. In order to test the validity of the predictive power of the selected MLR model, as opposed to traditional regression methods, the cross-validation method estimated a model's trustworthiness by predicting data. This approach uses the following cross-validated parameters: PRESS (predicted residual sum squares), SSY (sum of the squares of response value), R^2cv (overall predictive ability), S_{PRESS} (uncertainty of prediction), and PSE (predictive square error), equations 1, 2, 3 and 4 [39-43].

$$PSE = \sqrt{\frac{PRESS}{n}} \quad (1)$$

$$S_{PRESS} = \sqrt{\frac{PRESS}{n-p-1}} \quad (2)$$

$$R^2_{CV} = 1 - \frac{PRESS}{SSY} = Q^2 \quad (3)$$

$$S = \sqrt{\frac{SSY}{n-1}} \quad (4)$$

3. RESULTS AND DISCUSSION

3.1. Anti-proliferative STUDY

The anti-proliferative assay was performed to evaluate the inhibitory effect of the studied compounds against both hormone-dependent (MCF-7) and hormone-independent (MDA-MB-231) breast cancer cells.

As demonstrated in Fig. (**2**), all compounds showed promising cytotoxicity against MCF-7, with the most potent ones being A2, A9, A14, A17 and A27, with IC_{50} values equal to 17.9, 14.14, 14,.06, 19.34 and 15.9 μM, respectively. The only inactive compounds identified were A16 and A23, with IC_{50} equal to 201.6 and 227.2 μM, respectively. In contrast, most of the tested compounds were inactive against MDA-MB-231, with IC_{50} ranges from 101.7 to 318.4 μM, with the exception of compounds A10, A14, A25 and A1, which showed moderate cytotoxicity, with IC_{50} values equal to 83.24, 60.93, 31.4 and 79.56 μM, respectively. A similar result has been presented previously, where a ferrocenic derivative (1,2-bis(4′-hydroxyphenyl)-2-ferrocenyl-but-1-ene) showed very weak cytotoxicity against MDA-MB-231. However, 1,1-bis(4′-hydroxyphenyl)-2-ferrocenyl-but-1-ene indicated strong effects against those cells [18]. In addition, A25 indicated cytotoxicity against MDA-MB-231, as well as MCF-7, which may be explained by the existence of ERβ in both MDA-MB-231 and MCF-7 that could mediate the cytotoxicity [18, 44].

Results from (Fig. **2**) further highlight the fact that compound A14 was the most potent compound against both cancer cell lines. In addition, the high activity of A14, as compared with its two analogues, A13 and A15, indicated that the cyano group in the meta position might increase its activity against cancer cell lines. Although the activity of compounds A13, A14 and A15 against cell

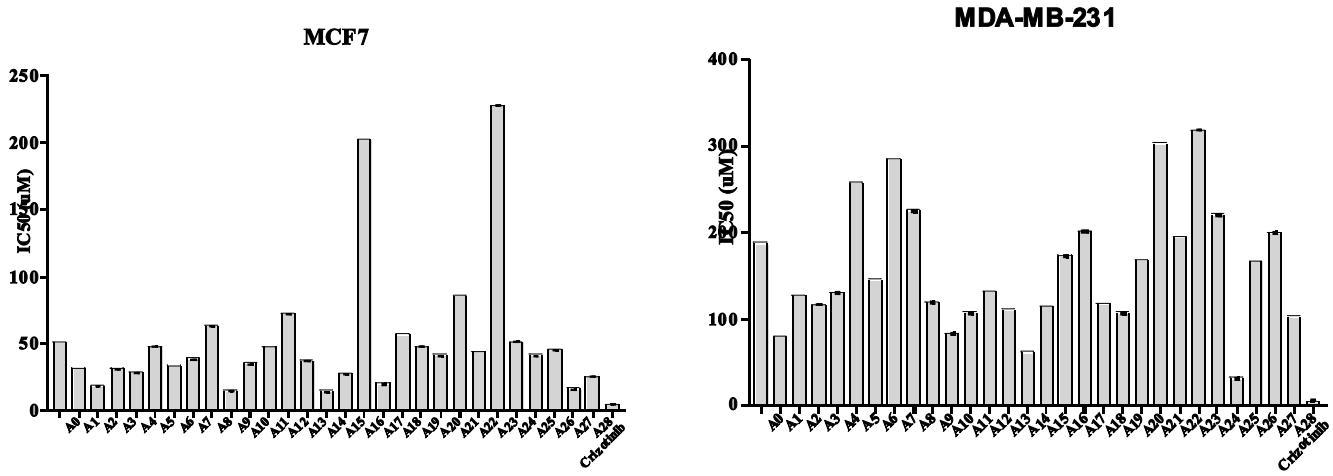


Fig. (2). IC_{50} values of compounds A0-A28 and Crizotinib of cell proliferation assay against MCF-7 and MDA-MB-231.

lines was weakened by the acylation of their azote atom, the acylated form of the meta cyano derivative A27 still possessed the highest activity compared to A26 and A1. Overall, the results obtained indicate that the presence of electron-withdrawing groups and electron-donating groups in the aromatic ring attached to the ferrocene moiety could perhaps be the key behind the different efficacy of this studied series of N-ferrocenylmethylaniline derivatives, which operates *via* the same mechanism as that of any anti-proliferative drugs. In addition, the amine group of the studied series is usually activated when an electron-withdrawing group is in the meta position of the aromatic ring, which is the reason behind the high cytotoxicity of compounds A2, A9, A14, A17 and A27 against MCF-7.

Although most of the studied compounds showed remarkable cytotoxicity, Crizotinib, which is used as a standard anti-cancer drug in this study, showed stronger cytotoxicity against both cell lines with IC50±SEM values equal to 3.5±1.5 µM for MDA-MB-231 and 3.5±1.5 µM for MCF-7.

3.2. Molecular Docking

Molecular docking was performed to determine the preferred binding site and binding mode between the five most active compounds against MCF-7. Since MDA-MB-231 lacks the estrogen and progesterone receptors, molecular docking was performed on the crystal structure of estrogen (ERα) and progesterone (PR) receptors.

The overall results observed from the docking study indicated high binding affinity between the five chosen molecules (A2, A9, A14, A17 and A27), and both targets, ERα and PR, with binding free energy in the range from -8.01 to -9.83 kcal.mol^{-1}, Table **3**.

Among the studied compounds, A27 showed the highest binding affinity towards both receptors, with the lowest binding free energy values of -9.05 and 9.83 kcal.mol^{-1}, respectively.

Molecular docking results further pointed out that hydrogen bonding, hydrophobic forces, and π-π stacking interactions are involved in the binding process between nearby residues of the receptors ERα and PR and the five studied compounds. Tables **4** and **5** summarize the interacting residues with their corresponding distances. As seen in Table **4**, all compounds contained two hydrogen bonds, except A2, which interacted beside A9 *via* π-π stacking interaction with the amino acids Phe 404 and Trp 393, respectively (Fig. **3**). The types of amino acids involved in the interactions confirm that all studied compounds were located at the estrogen binding cavity, and this observation is further supported by previously published work [45].

Our results stand in good agreement with previous studies that confirm that progesterone binds to its receptor *via* hydrogen and Van Der Waals interactions using the same amino acids, as shown in our study (Fig. **4**) [46, 47]. The molecular docking study also revealed that the identified compounds have higher binding affinity with ERα than with PR. Finally, compound A27 showed high binding affinity with both receptors, and figure 5 illustrates the secondary structure view of the adducts, A27-ERα (Fig. **5a**) and A27-PR (Fig.**5b**).

3.3. Drug-likeness Prediction

Drug-likeness studies qualitatively measured the chance of a molecule becoming an oral drug by assessing its bioavailability. Five different rules-based filters were used to calculate the drug-likeness for the five chosen compounds. The results highlighted that all the compounds show good drug-likeness scores with zero violation of understudy drug-likeness rules, except for A9, which had 1 violation (MLogP > 4.15) with Lipinski rule of five, and 2 violations (XLOGP3 > 5 and Heteroatoms <2) with Muegge filter. The Abbot bioavailability score predicts the chance of compound to have at least 10% oral bioavailability in rat or quantifiable Caco-2 cell line permeability experiments and defined by a probability score of 11%, 17%, 56%, and 85% [48]. The candidate compounds showed a score of 55%, indicating good bioavailability. Table **6** shows the obtained results of drug-likeness rules and the bioavailability.

3.4. Quantitative structure-activity relationship (QSAR)

Here, we tried to develop the best 2D-QSAR model to explain the correlations between the physicochemical properties and the anti-proliferative activity against MCF-7 cell lines. Table **7** shows the values of the calculated descriptors used in order to establish the 2D-QSAR model. MLR analysis was then performed and the best model was selected from several QSAR equations based on different statistical parameters, such as correlation coefficient (R > 0.8), squared correlation coefficient (R2 > 0.6), standard error of estimate (SEE < 0.3) and F-statistic values at 96% significance level.

$$\begin{aligned} pIC50 = & -13.85 - 2.179\,HOMO + 0.0003\,E \\ & +0.107\,MR + 0.022\,MV + 0.005\,HF \\ & -0.617\,Pol - 0.029\,TPSA \\ & -0.836\,qC1 - 0.829\,qN1 \end{aligned} \quad (5)$$

After multiple regression analysis using SPSS statistics 22 software [49], the best statistically significant generated correlation was given in the following model, (relation (5)).

Table 3. Binding free energy (ΔG) and binding constant (K) for the 5 studied molecules with ERα and PR.

Compounds	ERα		PR	
-	-ΔG (Kcal/mol)	K (M^{-1})	-ΔG (Kcal/mol)	K (M^{-1})
A2	8.12	724.44 x10^3	8.93	2783.31 x10^3
A9	8.01	603.42 x10^3	8.97	2974.60 x10^3
A14	8.23	869.73 x10^3	8.75	2063.76 x10^3
A17	7.48	250.11 x10^3	9.82	12213.86 x10^3
A27	9.05	3397.52 x10^3	9.83	12418.51 x10^3

Table 4. Distances of formed bonds between the ligands and ERα residues.

Bonds type	A2		A9		A14		A17		A27	
	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist
Hydrophobic interactions	LEU346	3.36	PRO324	3.87	LEU346	3.94	LEU346	3.35	THR347	3.52
	LEU384	3.27	ILE326	3.49	THR347	3.51	ALA350	3.58	LEU349	3.10
	LEU387	3.75	GLU353	3.35	LEU349	3.34	LEU384	3.62	ALA350	3.13
	LEU391	3.59	MET357	3.37	ALA350	3.18	LEU387	3.78	TRP383	3.12
	ILE424	3.41	LEU387	3.89	ALA350	3.20	LEU391	3.62	LEU387	3.62
	ILE424	3.73	ARG394	3.31	TRP383	3.17	-	-	PHE404	3.79
	PHE425	3.35	PHE445	3.14	LEU387	3.40	-	-	LEU525	3.20
	HIS524	3.54	PHE445	3.59	PHE404	3.47	-	-	-	-
	LEU525	3.85	-	-	LEU525	3.78	-	-	-	-
	-	-	-	-	LEU525	3.32	-	-	-	-
H-bonds	-	-	GLU353	2.01	ARG394	1.90	ARG394	1.67	ARG394	1.98
	-	-	ARG394	2.89	ARG394	2.98	ARG394	2.97	ARG394	3.11
π-Stacking	PHE404	5.05	TRP393	4.46	-	-	-	-	-	-

Table 5. Distances of formed bonds between the ligands and PR residues.

Bonds type	A2		A9		A14		A17		A27	
	AA	Dist	AA	Dist	AA	Dist	AA	Dist	AA	Dist
Hydrophobic interactions	LEU715	3.40	LEU718	3.76	LEU715	3.38	LEU715	3.37	LEU715	3.37
	LEU715	3.93	LEU721	3.42	VAL760	3.85	LEU718	3.18	LEU718	3.18
	LEU718	3.36	LEU721	3.97	LEU763	3.88	LEU718	3.48	LEU718	3.48
	LEU718	3.51	TRP755	3.78	PHE778	3.67	PHE794	3.90	PHE794	3.90
	MET759	3.47	PHE778	3.35	LEU797	3.35	TYR890	3.79	TYR890	3.79
	LEU763	3.80	TYR890	3.28	TYR890	3.60	-	-	-	-
	LEU797	3.46	THR894	3.81	PHE905	3.87	-	-	-	-
	LEU887	3.27	-	-	PHE905	3.83	-	-	-	-
	LEU887	3.54	-	-	-	-	-	-	-	-
	THR894	3.60	-	-	-	-	-	-	-	-
	PHE905	3.14	-	-	-	-	-	-	-	-
H-bonds	-	-	ASN719	1.92	GLN725	2.18	LEU718	2.18	LEU718	2.18
	-	-	-	-	ARG766	2.08	GLN725	1.88	GLN725	1.88
	-	-	-	-	-	-	LEU763	3.70	LEU763	3.70
	-	-	-	-	-	-	ARG766	1.75	ARG766	1.75
π-Stacking	-	-	-	-	-	-	PHE778	5.19	PHE778	5.19

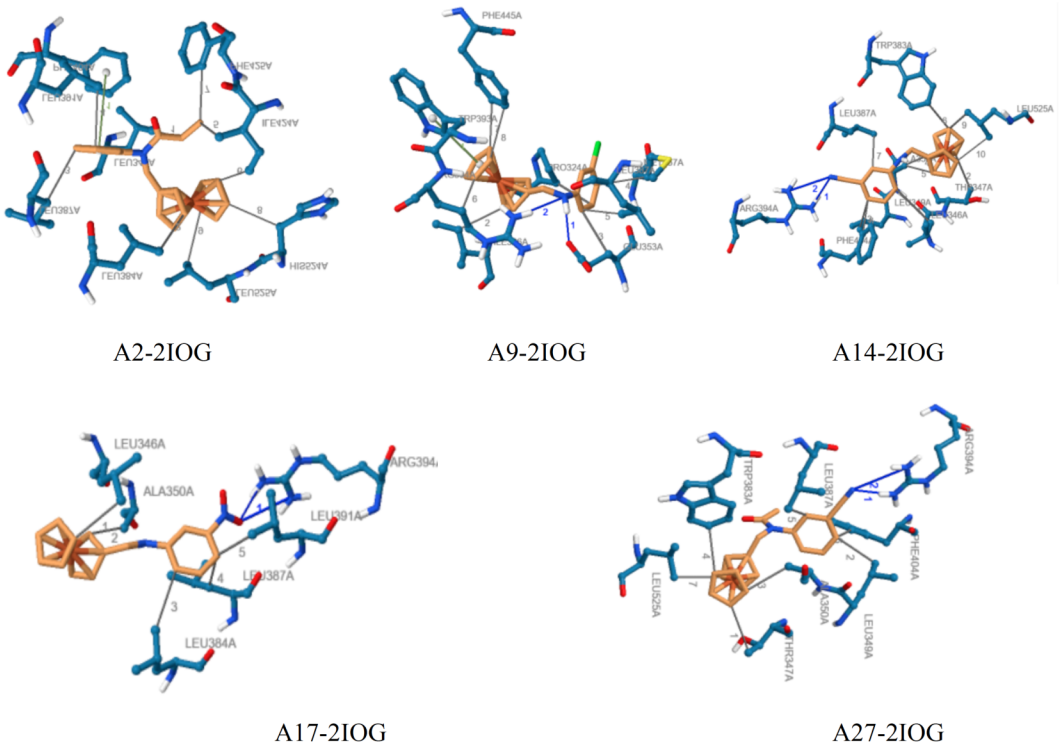


Fig. (3). 3D interaction between studied ligands and the target ERα (PDB ID: 2IOG), where the grey lines show hydrophobic bonds and the blue lines show H-bonds. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

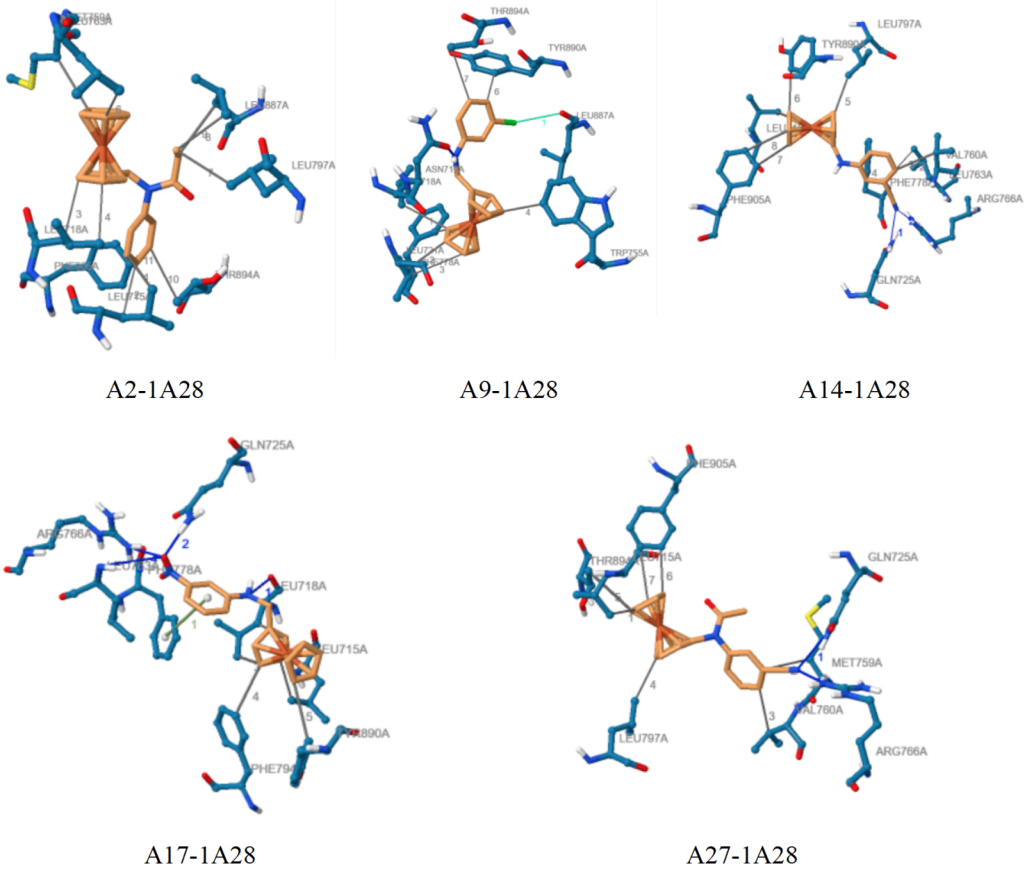
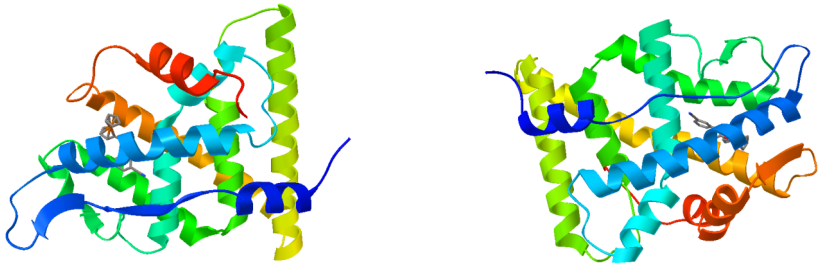


Fig. (4). 3D interaction between studied ligands and the target PR (PDB ID: 1A28), where the grey lines show hydrophobic bonds and the blue lines show H-bonds. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).



(a) **(b)**

Fig. (5). Docking poses of A27 with 2IOG (**a**) and 1A28 (**b**), illustrating the secondary structural view of the interactions between the proteins and the examined ligand. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

Table 6. Details of different drug-likeness rules and the bioavailability.

Compounds	Lipinski	Ghose	Veber	Egan	Muegge	Bioavailability Score
A2	Yes; 0 violation	Yes	Yes	Yes	Yes	0.55
A9	Yes; 1 violation: MLOGP > 4.15	Yes	Yes	Yes	No; 2 violations: XLOGP3 > 5, Heteroatoms < 2	0.55
A14	Yes; 0 violation	Yes	Yes	Yes	Yes	0.55
A17	Yes; 0 violation	Yes	Yes	Yes	Yes	0.55
A27	Yes; 0 violation	Yes	Yes	Yes	Yes	0.55

Table 7. Values of molecular descriptors used in the multiple regression analysis.

Comp	pIC50	HOMO (eV)	LUMO (eV)	ΔE (eV)	Total E (a.u)	DM (Debay)	MR ($Å^3$)	log P	log S	MV ($Å^3$)	HF (kcal/mol)	TPSA ($Å^2$)	MW (amu)	Pol ($Å^3$)
A0	4.298	-5.442	-0.49	4.952	-1976.69	1.796	78.79	3.03	-4.76	798.72	-148.52	12.03	291.18	30.25
A1	4.504	-5.608	-1.983	3.624	-2221.68	4.966	93.42	2.63	-4.49	949.61	-152.842	44.1	358.22	35.86
A2	4.747	-5.687	-0.626	5.061	-2168.71	3.810	93.51	3.15	-4.77	941.97	-191.884	20.31	347.24	35.84
A3	4.514	-5.442	-1.252	4.191	-2321.17	3.456	108.6	3.9	-6.01	1047.36	-152.411	20.31	395.28	41.83
A4	4.555	-5.448	-2.765	2.683	-2525.73	6.184	117.43	3.19	-6.79	1100.54	-156.861	66.13	440.2	43.55
A5	4.328	-5.496	-0.462	5.034	-2016.02	1.229	83.76	3.3	-5.14	841.92	-155.7	12.03	305.2	32.09
A6	4.478	-5.388	-0.489	4.898	-2016.02	1.606	83.76	3.3	-5.14	848.88	-157.865	12.03	305.2	32.09
A7	4.418	-5.496	-0.462	5.034	-2055.34	1.222	88.57	3.56	-5.58	890.31	-158.701	12.03	319.23	33.92
A8	4.201	-5.586	-0.582	5.004	-2436.32	2.138	83.8	3.5	-5.41	837.18	-155.037	12.03	325.62	32.18
A9	4.850	-5.632	-0.634	4.998	-2436.32	3.418	83.8	3.5	-5.41	842.77	-155.491	12.03	325.62	32.18
A10	4.465	-5.551	-0.653	4.898	-2436.32	3.874	83.8	3.5	-5.41	842.09	-155.44	12.03	325.62	32.18
A11	4.327	-5.633	-0.653	4.979	-4550.24	3.48	86.49	3.57	-5.48	860.72	-141.27	12.03	370.07	32.88
A12	4.144	-5.551	-0.68	4.871	-4550.24	3.96	86.49	3.57	-5.48	861.34	-141.25	12.03	370.07	32.88
A13	4.431	-5.769	-1.415	4.354	-2068.97	4.486	83.51	2.89	-5.54	857.83	-111.68	35.82	316.19	32.1
A14	4.852	-5.714	-1.524	4.191	-2068.97	6.092	83.51	2.78	-4.97	855.64	-113.73	35.82	316.19	32.1
A15	4.569	-5.905	-1.143	4.762	-2068.97	8.250	83.51	2.78	-4.97	851.21	-118.746	35.82	316.19	32.1
A16	3.696	-5.606	-2.64	2.966	-2181.27	5.147	87.62	2.56	-6.12	837.99	-159.441	57.85	336.17	31.97
A17	4.714	-5.823	-2.612	3.211	-2181.27	6.752	87.62	2.45	-5.55	866.96	-156.37	57.85	336.17	31.97
A18	4.253	-5.987	-2.259	3.728	-2181.27	9.448	87.62	2.45	-5.55	865.39	-158.208	57.85	336.17	31.97
A20	4.327	-5.741	-1.279	4.462	-2129.38	6.326	88.99	2.96	-4.79	910.78	-189.498	29.1	333.21	34.01
A19	4.389	-5.632	-1.768	3.864	-2129.38	5.052	88.99	3	-5	900.48	-187.028	29.1	333.21	34.01
A21	4.069	-5.742	-0.653	5.089	-2184.74	3.953	91.5	2.64	-4.82	926.6	-164.587	32.34	348.23	35.36
A22	4.365	-5.66	-1.252	4.408	-2376.52	3.424	96.31	2.93	-5.32	1072.71	-128.642	32.34	410.3	43.18
A23	3.644	-5.578	-2.912	2.667	-2333.96	6.31	97.52	2.15	-5.07	944.33	-194.722	66.13	378.21	35.72
A24	4.292	-5.959	-2.721	3.238	-2333.96	8.893	97.52	2.15	-5.07	930.05	-201.07	66.13	378.21	35.72
A25	4.386	-5.687	-3.075	2.612	-2333.96	2.193	97.52	2.15	-5.07	953.21	-196.525	66.13	378.21	35.72
A26	4.351	-5.551	-2.095	3.456	-2221.67	6.395	93.42	2.63	-4.49	943.26	-151.403	44.1	358.22	35.86
A27	4.799	-5.606	-2.095	3.51	-2221.67	4.035	93.42	2.63	-4.49	949.31	-152.688	44.1	358.22	35.86
A28	4.603	-5.469	-0.925	4.544	-2129.40	3.87	88.7	2.85	-4.28	894.27	-188.402	20.31	333.21	34.01

Table 7 Contd….

Comp	qFe (C)	qC9 (C)	qC10 (C)	qC11 (C)	qC12 (C)	qC13 (C)	qC14 (C)	qC15 (C)	qN1 (C)
A0	0.561	-0.095	0.662	-1.14	-0.324	-0.124	-0.272	-0.373	0.206
A1	0.651	-0.419	1.003	-1.355	-0.821	0.134	-0.277	1.749	0.665
A2	0.548	-0.4	-0.316	-0.539	-0.745	-0.465	-0.129	-0.29	0.574
A3	0.454	-0.167	0.749	-0.8	-0.281	-0.381	-0.397	0.018	0.546
A4	0.457	-0.392	0.842	-0.895	-0.171	0.073	-0.402	-0.274	0.521
A5	0.639	-0.18	0.345	-0.929	-0.313	0.355	-0.081	-0.409	0.184
A6	0.59	-0.099	0.736	-1.189	-0.46	0.396	0.603	-0.749	0.219
A7	0.629	-0.264	0.563	-0.975	-0.062	0.572	-0.216	-0.509	0.185
A8	0.555	-0.199	0.692	-1.085	-1.39	1.111	-0.198	-0.524	0.303
A9	0.544	-0.137	0.668	-1.087	-0.661	-0.339	0.527	-0.717	0.205
A10	0.551	-0.155	0.682	-1.121	-0.646	-0.24	-0.649	0.554	0.218
A11	0.559	-0.122	0.682	-1.126	-0.248	-0.339	0.091	-0.218	0.199
A12	0.556	-0.115	0.645	-1.115	-0.464	-0.224	-0.253	0.081	0.211
A13	0.868	-0.324	-0.22	-0.451	-0.579	1.479	0.151	-0.367	0.108
A14	0.593	-0.368	-0.832	-1.267	-0.407	0.849	1.339	-0.703	0.192
A15	0.527	-0.2	-0.026	-0.304	-0.058	-0.069	-0.259	1.84	0.078
A16	0.429	-0.176	0.731	-0.998	-0.257	-0.032	0.066	-0.395	0.093
A17	0.71	-0.355	-0.185	-0.707	-0.455	0.338	0.203	0.286	0.121
A18	0.821	-0.373	-0.152	-0.468	-0.066	0.466	-0.002	-0.451	0.066
A19	0.778	-0.405	-0.029	-0.535	-0.405	-0.766	-1.217	1.596	0.102
A20	0.842	-0.417	-0.116	-0.553	-0.404	0.703	-0.005	-0.498	0.235
A21	0.802	-0.385	0.361	-0.905	-0.584	0.007	-0.086	-0.301	0.901
A22	0.539	-0.423	0.768	-0.936	-0.112	-0.433	-0.226	0.074	-0.573
A23	0.696	-0.31	-0.772	-0.986	-0.199	-0.199	0.071	-0.429	0.855
A24	0.784	-0.302	0.893	-0.85	-0.035	-0.543	-0.217	0.104	0.536
A25	-0.146	-0.36	0.468	-0.541	0.465	-0.305	-0.347	0.367	-0.244
A26	0.704	-0.198	1.048	-1.302	-0.954	1.198	-0.078	-0.374	0.703
A27	0.79	-0.47	-0.913	-1.26	-0.788	-0.383	1.445	-0.316	0.591
A28	0.736	-0.455	0.913	-1.132	-0.66	0.152	-0.278	-0.382	0.55

The model shows R = 0.92 and R^2= 0.84, which allowed us to firmly confirm the correlation between the descriptors (HOMO, E, MR, MV, HF, POL, TPSA, qC11 and qN1) and the anti-proliferative activity against the MCF-7 cell line. The calculated F value of the obtained model was equal to 9.33, which was greater than the one tabulated by a wide margin, required for a meaningful regression. The positive value of Q equal to 6.75 for QSAR's model suggests its high predictive power. Moreover, the accuracy in the analysis was indicated by the low standard error of estimate, which was equal to 0.14. All the t-statistic values were found to be significant, confirming the significance of each descriptor.

The equation showed that the negative coefficients of HOMO, Pol, TPSA, qC11 and qN1 explain that any increase in HOMO energy, polarizability, topological polar surface area and the charges of C11 and N1 of the compounds causes a decrease in the biological activity. Conversely, the positive coefficients of E, MR, MV and HF indicated that any increase in these parameters causes an increase in biological activity.

The correlation matrix between the descriptors obtained by the MLR analysis and the biological activity is reported in Table **8**, where the descriptors obtained afterwards are correlated with each other and the biological activity. Table **8** further shows that the heat of formation and TPSA are important parameters in the correlation between the selected descriptor and the anti-proliferative activity against MCF-7 cell lines.

The generated model was validated for predictive ability inside the model using the ''Leave One Out'' method (LOO). From the results depicted in Table **9**, the obtained model was statistically significant. The PRESS value was lower than SSY, which indicates that the model predicts better than chance and can be considered statically significant. Furthermore, the PRESS/SSY ratio was equal to 0.151, which should be lower than 0.4 [50]. The high R^2_{cv} and low S_{PRESS} (equal to 0.848 and 0.148, respectively) reflect its good predictive potential. Also, the high value of R^2_{adj} is an important criterion for our QSAR model to have the best qualification.

The predictive error of the coefficient correlation PE is another parameter used to evaluate the predictive power of the obtained model. For the developed model, the condition of R > 6 PE was satisfied, hence, the model has good predictive power.

Fig. (**6**) shows the plot of linear regression of predicted versus experimental values of the biological activity. The plot for the model shows a good correlation with experimentally reported data having R^2 = 0.848. The plot of residuals pIC_{50} against the experimental values of the biological activity does not show any systematic error (Fig. **7** and Table **10**). The propagation of the residuals on both sides of zero indicates that no systematic error exists, as suggested by Jalali-Heravi and Kyani [51].

Table 8. Correlation Matrix for the obtained model.

-	pIC50	HOMO	E	MR	MV	HF	pol	TPSA	qC11	qN1
pIC50	1.000	-	-	-	-	-	-	-	-	-
HOMO	-0.125-	1.000	-	-	-	-	-	-	-	-
E	0.168	-.086-	1.000							
MR	-0.018	0.135	0.01	1.000						
MV	0.076	0.102	0.053	0.926	1.000					
HF	0.283	0.008	-0.117	-0.351	-0.191	1.000				
pol	0.051	0.155	0.019	0.906	0.988	-0.14	1.000			
TPSA	-0.204	-0.393	0.264	0.55	0.442	-0.344	0.35	1.000		
qC11	0.053	-0.502	0.214	0.113	0.095	-0.146	0.066	0.206	1.000	
qN1	-0.134	0.108	0.068	0.303	0.145	-0.325	0.098	0.151	-0.285	1.000

Table 9. Cross-validation parameters.

R^2_{adj}	R^2_{cv}	PE	6PE	SSY	PRESS	S_{PRESS}	PRESS/SSY
0.758	0.848	0.021	0.129	2.191	0.332	0.148	0.151

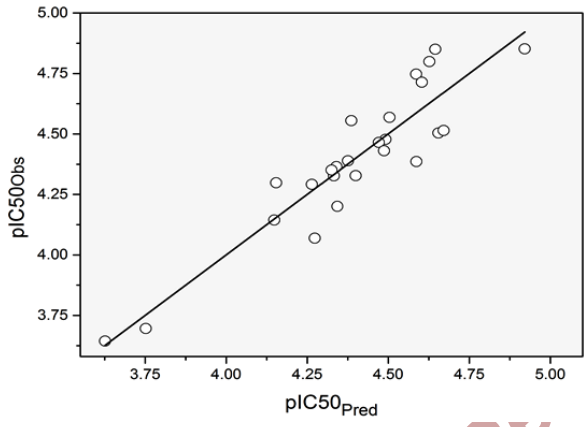


Fig. (6). Scatter plot between the Observed and Predicted Activities of Model.

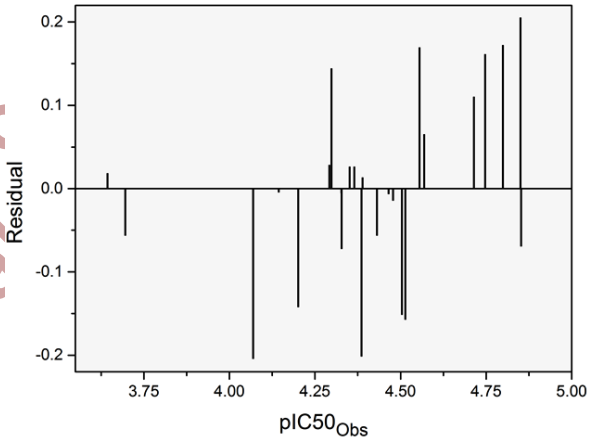


Fig. (7). Plots of the residual values against the experimentally observed values.

Table 10. Observed and calculated activity of N-ferrocenylmethylaniline derivatives.

Compounds	$pIC50_{Obs}$	$pIC50_{Pred}$	Residual
A0	4.298	4.154	0.144
A1	4.504	4.655	-0.151
A2	4.747	4.586	0.161
A3	4.514	4.671	-0.157
A4	4.555	4.386	0.169
A5	4.328	4.400	-0.072
A6	4.478	4.492	-0.014
A8	4.201	4.343	-0.142
A9	4.850	4.645	0.205
A10	4.465	4.471	-0.006
A11	4.327	4.332	-0.005
A12	4.144	4.148	-0.004
A13	4.431	4.487	-0.056

Table 10 Contd…

Compounds	$pIC50_{Obs}$	$pIC50_{Pred}$	Residual
A14	4.852	4.921	-0.069
A15	4.569	4.504	0.065
A16	3.696	3.752	-0.056
A17	4.714	4.604	0.110
A19	4.389	4.376	0.013
A21	4.069	4.273	-0.204
A22	4.365	4.340	0.026
A23	3.644	3.626	0.018
A24	4.292	4.264	0.028
A25	4.386	4.587	-0.201
A26	4.351	4.325	0.026
A27	4.799	4.627	0.172

Thus, the current QSAR research shows that in the generation of these molecules, this model can be successfully applied to predict anti-proliferative activity against the MCF-7 cell line.

CONCLUSION

In the present work, the anti-proliferative effect of 29 N-ferrocenylmethylaniline derivatives against MCF-7 cell lines was investigated and the results indicated that among the derivatives studied, five were very active; the binding affinity of these five active derivatives towards estrogen and progesterone receptors was further studied by molecular docking. The results of anti-proliferative assays and molecular docking were in good agreement. Furthermore, the application of drug-likeness rules shows that most of these compounds, theoretically, can be orally bioavailable. Finally, MLR was used to develop a linear QSAR model for the prediction of the anti-proliferative activity of N-ferrocenylmethylaniline derivatives. The built model displayed a good correlation between the structure and anti-proliferative activity of the studied derivatives. The validation of the model by ''Leave One Out'' method (LOO) method suggests that the model can be used to design new N-ferrocenylmethylaniline derivatives with improved anti-proliferative activity. The knowledge gained by this work can be used to design more potent N-ferrocenylmethylaniline derivatives as anti-proliferative agents.

ETHICAL APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No Animals/Humans were used in the studies that are the base of this research.

CONSENT FOR PUBLICATION

Not applicable

AVAILABILITY OF DATA AND MATERIALS

Not applicable

FUNDING

Financial support was received from the Directorate-General of Scientific Research and Technological Development (DGRSDT) (project number B00L01UN390120150001), Algeria.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to the Directorate-General of scientific research and technological development (DGRSDT) for financial support (project number B00L01UN390120150001). The authors extend thanks to the VTRS staff for providing facilities.

REFERENCES

[1] World Health Organization. *Cancer fact sheets,* Available from: http://gco.iarc.fr/today/fact-sheets-cancers [Accessed November 10, 2020]

[2] Jéhannin-Ligier, K.; Dantony, E.; Bossard, N.; Molinié, F.; Defossez, G.; Daubisse-Marliac, L.; Delafosse, P.; Remontet, L.; Uhry, Z. Projection de l'incidence et de la mortalité par cancer en France métropolitaine en 2017.*Rapport technique*; Saint-Maurice: Santé publique France, **2017**, pp. 1-80.

[3] American Cancer Society. *Breast cancer,* Available from: https://www.cancer.org/cancer/breast-cancer.html [Accessed November 10, 2020]

[4] Hudis, C.A. Trastuzumab--mechanism of action and use in clinical practice. *N. Engl. J. Med.*, **2007**, *357*(1), 39-51. http://dx.doi.org/10.1056/NEJMra043186 PMID: 17611206

[5] Ismael, G.; Rosa, D.D.; de Azambuja, E.; Braga, S.; Piccart-Gebhart, M. Trastuzumab (herceptin) for early-stage breast cancer. *Hematol. Oncol. Clin. North Am.*, **2007**, *21*(2), 239-256. http://dx.doi.org/10.1016/j.hoc.2007.03.003 PMID: 17512447

[6] Wang, J.; Xu, B. Targeted therapeutic options and future perspectives for HER2-positive breast cancer. *Signal Transduct. Target. Ther.*, **2019**, *4*(1), 34. http://dx.doi.org/10.1038/s41392-019-0069-2 PMID: 31637013

[7] Hortobagyi, G.N.; Buzdar, A.U. Current status of adjuvant systemic therapy for primary breast cancer: Progress and controversy. *CA Cancer J. Clin.*, **1995**, *45*(4), 199-226. http://dx.doi.org/10.3322/canjclin.45.4.199 PMID: 7600278

[8] Group, E.B.C.T.C. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: An overview of the randomised trials. *Lancet,* **2005**, *365*(9472), 1687-1717. http://dx.doi.org/10.1016/S0140-6736(05)66544-0 PMID: 15894097

[9] Gong, Z.; Chen, M.; Ren, Q.; Yue, X.; Dai, Z. Fibronectin-targeted dual-acting micelles for combination therapy of metastatic breast cancer. *Signal Transduct. Target. Ther.*, **2020**, *5*(1), 12. http://dx.doi.org/10.1038/s41392-019-0104-3 PMID: 32296050

[10] Jordan, V.C. Antiestrogens and selective estrogen receptor modulators as multifunctional medicines. 2. Clinical considerations and new agents. *J. Med. Chem.*, **2003**, *46*(7), 1081-1111. http://dx.doi.org/10.1021/jm020450x PMID: 12646017

[11] Singh, M.N.; Stringfellow, H.F.; Paraskevaidis, E.; Martin-Hirsch, P.L.; Martin, F.L. Tamoxifen: Important considerations of a multifunctional compound with organ-specific properties. *Cancer Treat. Rev.*, **2007**, *33*(2), 91-100.

http://dx.doi.org/10.1016/j.ctrv.2006.09.008 PMID: 17178195

[12] Dixon, J.M. Exemestane and aromatase inhibitors in the management of advanced breast cancer. *Expert Opin. Pharmacother.,* **2004**, *5*(2), 307-316. http://dx.doi.org/10.1517/14656566.5.2.307 PMID: 14996627

[13] Meegan, M.J.; Lloyd, D.G. Advances in the science of estrogen receptor modulation. *Curr. Med. Chem.,* **2003**, *10*(3), 181-210. http://dx.doi.org/10.2174/0929867033368501 PMID: 12570707

[14] Magarian, R.; Overacre, L.; Singh, S.; Meyer, K. The medicinal chemistry of nonsteroidal antiestrogens: A review. *Curr. Med. Chem.,* **1994**, *1*, 61-104.

[15] American Cancer Society. *Treating breast cancer,* Available from: https://www.cancer.org/cancer/breast-cancer/treatment.html [Accessed November 10, 2020]

[16] Köpf-Maier, P.; Köpf, H.; Neuse, E.W. Ferricenium complexes: A new type of water-soluble antitumor agent. *J. Cancer Res. Clin. Oncol.,* **1984**, *108*(3), 336-340. http://dx.doi.org/10.1007/BF00390468 PMID: 6511806

[17] Köpf-Maier, P. Tumor inhibition by ferricenium complexes: Systemic effect *in vivo* and cell growth inhibition *in vitro*. *Z Naturforsch C Biosci,* **1985**, *40*(11-12), 843-846. http://dx.doi.org/10.1515/znc-1985-11-1215 PMID: 3832667

[18] Vessières, A.; Top, S.; Pigeon, P.; Hillard, E.; Boubeker, L.; Spera, D.; Jaouen, G. Modification of the estrogenic properties of diphenols by the incorporation of ferrocene. Generation of antiproliferative effects *in vitro*. *J. Med. Chem.,* **2005**, *48*(12), 3937-3940. http://dx.doi.org/10.1021/jm050251o PMID: 15943467

[19] Hillard, E.; Vessières, A.; Le Bideau, F.; Plazuk, D.; Spera, D.; Huché, M.; Jaouen, G. A series of unconjugated ferrocenyl phenols: Prospects as anticancer agents. *ChemMedChem,* **2006**, *1*(5), 551-559. http://dx.doi.org/10.1002/cmdc.200500035 PMID: 16892391

[20] Khand, I.U.; Lanez, T.; Pauson, P.L. Ferrocene derivatives. Part 24. Synthesis of dihydro-2-pyrindines and dihydro-3H-2-cyclopent[c]azepines by photolysis of their cyclopentadienyliron derivatives. *J. Chem. Soc., Perkin Trans. 1,* **1989**, 2075-2078. http://dx.doi.org/10.1039/p19890002075

[21] Boubekri, C.; Khelef, A.; Terki, B.; Lanez, T. Synthesis and electrochemical properties of n-(ferrocenylmethyl)aminobenzonitrile and n-(ferrocenylmethyl)nitroaniline derivatives. *ILCPA,* **2015**, *49*, 27-34. http://dx.doi.org/10.18052/www.scipress.com/ILCPA.49.27

[22] Lanez, E.; Bechki, L.; Lanez T. N6,9-bis(ferrocenylmethyl)adenine: Synthesis, cyclic voltammetric, spectroscopic characterization, and DFT calculations. *St. Cerc. St. CICBIA,* **2019**, *20*(4), 509-519.

[23] Khelef, A.; Terki, B.; Mahboub, M.S.; Lanez, T. N-Ferrocenymethyl-N-phenylpropionamide. *Acta Crystallogr.,* **2012**, *E68*, m647.

[24] Lanez, E.; Bechki, L.; Lanez, T. Ferrocenylmethylnucleobases: Synthesis, DFT calculations, electrochemical and spectroscopic characterisation. *Chem. Chem. Technol,* **2020**, *14*(2), 146-153. http://dx.doi.org/10.23939/chcht14.02.146

[25] Becke, A.D. Density□functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.,* **1993**, *98*(7), 5648-5652. http://dx.doi.org/10.1063/1.464913

[26] Miehlich, B. Results obtained with the correlation energy density functionals of Becke and Lee, Yang and Parr. *Chem. Phys. Lett.,* **1989**, *157*(3), 200-206. http://dx.doi.org/10.1016/0009-2614(89)87234-3

[27] Frisch, M. *Gaussian 09*; Gaussian, Inc.: Wallingford CT, **2009**.

[28] Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.,* **2009**, *30*(16), 2785-2791. http://dx.doi.org/10.1002/jcc.21256 PMID: 19399780

[29] Dykstra, K.D.; Guo, L.; Birzin, E.T.; Chan, W.; Yang, Y.T.; Hayes, E.C.; DaSilva, C.A.; Pai, L.Y.; Mosley, R.T.; Kraker, B.; Fitzgerald, P.M.; DiNinno, F.; Rohrer, S.P.; Schaeffer, J.M.; Hammond, M.L. Estrogen receptor ligands. Part 16: 2-Aryl indoles as highly subtype selective ligands for ERalpha. *Bioorg. Med. Chem. Lett.,* **2007**, *17*(8), 2322-2328. http://dx.doi.org/10.1016/j.bmcl.2007.01.054 PMID: 17289385

[30] Williams, S.P.; Sigler, P.B. Atomic structure of progesterone complexed with its receptor. *Nature,* **1998**, *393*(6683), 392-396. http://dx.doi.org/10.1038/30775 PMID: 9620806

[31] Systèmes, D. Biovia discovery studio modeling environment.*Dassault Systèmes Biovia*; San Diego, CA, USA, **2016**.

[32] Darwin, K.H.; Ehrt, S.; Gutierrez-Ramos, J.C.; Weich, N.; Nathan, C.F. The proteasome of mycobacterium tuberculosis is required for resistance to nitric oxide. *Science,* **2003**, *302*(5652), 1963-1966. http://dx.doi.org/10.1126/science.1091176 PMID: 14671303

[33] Bandyopadhyay, N.; Pradhan, A.B.; Das, S.; Lu, L.; Zhu, M.; Chowdhury, S.; Naskar, J.P. Synthesis, structure, DFT calculations, electrochemistry, fluorescence, DNA binding and molecular docking aspects of a novel oxime based ligand and its palladium(II) complex. *J. Photochem. Photobiol. B,* **2016**, *160*, 336-346. http://dx.doi.org/10.1016/j.jphotobiol.2016.04.026 PMID: 27179300

[34] Guhathakurta, B. Synthesis, characterization, structure, DNA binding aspects and molecular docking study of a novel Schiff base ligand and its bis (μ-chloro) bridged Cu (II) dimer. *Polyhedron,* **2017**, *126*, 195-204. http://dx.doi.org/10.1016/j.poly.2017.01.033

[35] Salentin, S.; Schreiber, S.; Haupt, V.J.; Adasme, M.F.; Schroeder, M. PLIP: Fully automated protein-ligand interaction profiler. *Nucleic Acids Res.,* **2015**, *43*(W1), W443-7. http://dx.doi.org/10.1093/nar/gkv315 PMID: 25873628

[36] Daina, A.; Michielin, O.; Zoete, V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.,* **2017**, *7*, 42717. http://dx.doi.org/10.1038/srep42717 PMID: 28256516

[37] Daina, A.; Michielin, O.; Zoete, V. iLOGP: A simple, robust, and efficient description of n-octanol/water partition coefficient for drug design using the GB/SA approach. *J. Chem. Inf. Model.,* **2014**, *54*(12), 3284-3301. http://dx.doi.org/10.1021/ci500467k PMID: 25382374

[38] Teppen, B.J. HyperChem, release 2: Molecular modeling for the personal computer. *J. Chem. Inf. Comput. Sci.,* **1992**, *32*(6), 757-759. http://dx.doi.org/10.1021/ci00010a025

[39] Purkayastha, S.K.; Jha, T.; Pal, D.K.; De, A.U. Possible antineoplastic agents: Part XIII. Synthesis, biological evaluation and QSAR studies of some 1-(substituted benzenesulphonyl)-5-oxopyrrolidine-2-carboxylic acid derivatives. *Anticancer Drug Des.,* **1993**, *8*(2), 95-100. PMID: 8494605

[40] Srikanth, K. As a part of a composite programme of rational drug design (RDD)', we had synthesized some substituted. *Indian J. Biochem. Biophys.,* **2001**, *38*, 120. PMID: 11563323

[41] Srikanth, K.; Debnath, B.; Jha, T. QSAR study on adenosine kinase inhibition of pyrrolo[2,3-d]pyrimidine nucleoside analogues using the hansch approach. *Bioorg. Med. Chem. Lett.,* **2002**, *12*(6), 899-902. http://dx.doi.org/10.1016/S0960-894X(02)00042-2 PMID: 11958989

[42] Jha, T. QSAR study on some substituted glutamine analogs as anticancer agents. Internet Electron. *J. Mol. Des,* **2003**, *2*, 539-545.

[43] Samanta, S.; Srikanth, K.; Banerjee, S.; Debnath, B.; Gayen, S.; Jha, T. 5-N-Substituted-2-(substituted benzenesulphonyl) glutamines as antitumor agents. Part II: Synthesis, biological activity and QSAR study. *Bioorg. Med. Chem.,* **2004**, *12*(6), 1413-1423. http://dx.doi.org/10.1016/j.bmc.2004.01.006 PMID: 15018914

[44] Vladusic, E.A.; Hornby, A.E.; Guerra-Vladusic, F.K.; Lupu, R. Expression of estrogen receptor β messenger RNA variant in breast cancer. *Cancer Res.,* **1998**, *58*(2), 210-214. PMID: 9443393

[45] Brzozowski, A.M.; Pike, A.C.; Dauter, Z.; Hubbard, R.E.; Bonn, T.; Engström, O.; Ohman, L.; Greene, G.L.; Gustafsson, J.A.; Carlquist, M. Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature,* **1997**, *389*(6652), 753-758. http://dx.doi.org/10.1038/39645 PMID: 9338790

[46] Mordasini, T.; Curioni, A.; Bursi, R.; Andreoni, W. The binding mode of progesterone to its receptor deduced from molecular dynamics simulations. *ChemBioChem,* **2003**, *4*(2-3), 155-161. http://dx.doi.org/10.1002/cbic.200390026 PMID: 12616628

[47] Tanenbaum, D.M.; Wang, Y.; Williams, S.P.; Sigler, P.B. Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains. *Proc. Natl. Acad. Sci. USA,* **1998**, *95*(11), 5998-6003. http://dx.doi.org/10.1073/pnas.95.11.5998 PMID: 9600906

[48] Alam, S.; Khan, F. 3D-QSAR, Docking, ADME/Tox studies on Flavone analogs reveal anticancer activity through Tankyrase inhibition. *Sci. Rep.,* **2019**, *9*(1), 5414. http://dx.doi.org/10.1038/s41598-019-41984-7 PMID: 30932078

[49] Jalali-Heravi, M.; Kyani, A. Use of computer-assisted methods for the modeling of the retention time of a variety of volatile organic compounds: A PCA-MLR-ANN approach. *J. Chem. Inf. Comput. Sci.,* **2004**, *44*(4), 1328-1335. http://dx.doi.org/10.1021/ci0342270 PMID: 15272841

[50] Podunavac-Kuzmanović, S.O.; Cvetković, D.D.; Barna, D.J. QSAR analysis of 2-amino or 2-methyl-1-substituted benzimidazoles against Pseudomonas aeruginosa. *Int. J. Mol. Sci.,* **2009**, *10*(4), 1670-1682. http://dx.doi.org/10.3390/ijms10041670 PMID: 19468332

[51] Srivastava, A.; Shukla, N. Quantitative structure activity relationship (QSAR) studies on a series of imidazole derivatives as novel ORL1 receptor antagonists. *J. Saudi Chem. Soc.,* **2013**, *17*(3), 321-328. http://dx.doi.org/10.1016/j.jscs.2011.04.014

Abstract

Breast cancer is the most common cancer in women in front of colorectal cancer and lung cancer. Ferrocene derivatives have proved their efficiency against breast cancer cells. In this thesis we aimed to find new ferrocene derivatives having pharmacochemistry properties as potential drugs against human breast cancer cells. A series of 29 N-ferrocenylmethylaniline derivatives (FMA) were optimized to determine the structural, crystal properties and molecular spectroscopy using DFT/ B3LYP method with 6-311 ++ G (d,p) and LanL2DZ basis sets. The Pharmacokinetics and ADME properties were also studied. The anti-proliferative activity of the studied derivatives against both hormone-dependent (MCF-7) and independent (MDA-MB 231) human breast cancer cell lines were performed using MTT test. Moreover, the binding affinity between FMA derivatives and MCF-7 was further studied by molecular docking using AutoDock 4.2 software, using both estrogen and progesterone receptors as targets. Next, another docking study was performed with aromatase and 17β-HSD1, the enzymes that catalyse the estrogen synthesis. Finally, the anti-proliferative activity of FMA derivatives against MCF-7 cell lines was investigated to predict a Quantitative Structure–Activity Relationship (QSAR) model using molecular descriptors and MLR analysis.

The FMA derivatives have shown an important cytotoxicity effect against MCF-7 cell lines, which was explained using molecular docking by the interaction of the compounds with the ERα and PR. Furthermore, the molecular docking study against aromatase shows that A4 and A25 were the best with the lowest binding free energies comparing to other compounds and ASD. However, moderate affinity was predicted against 17β-HSD1 for all compounds comparing to E2B. This indicate that these derivatives might be strong drug inhibitors of Aromatase. The best equation of regression contains HOMO, E, MR, MV, HF, POL, TPSA, qC11 and qN1 descriptors. The QSAR model indicates that these descriptors have significant relationships with the observed bioactivity. The model was validated for predictive ability inside the model using Leave One Out method (LOO), which proves that the model can be successfully applied to predict anti-proliferative activity against MCF-7 cell line.

Résumé

Le cancer du sein est le cancer le plus fréquent chez les femmes, devant le cancer colorectal et le cancer du poumon. Les dérivés du ferrocène ont prouvé leur efficacité contre les cellules cancéreuses du sein. Dans cette thèse, nous avons cherché à trouver de nouveaux dérivés du ferrocène ayant des propriétés pharmacochimiques comme médicaments potentiels contre les cellules humaines du cancer du sein. Une série de 29 dérivés de N-ferrocénylméthylaniline (FMA) ont été optimisés pour déterminer la structure, les propriétés cristallines et la spectroscopie moléculaire en utilisant la méthode DFT/ B3LYP avec les bases 6-311 ++ G (d,p) et LanL2DZ. Les propriétés de pharmacocinétique et ADME ont également été étudiées. L'activité antiprolifération des dérivés étudiés contre les lignées cellulaires de cancer du sein humain hormono-dépendant (MCF-7) et indépendant (MDA-MB 231) a été réalisée à l'aide du test MTT. En outre, l'affinité de liaison entre les dérivés de la FMA et le MCF-7 a été étudiée plus avant par amarrage moléculaire à l'aide du logiciel AutoDock 4.2, en utilisant les récepteurs d'œstrogène et de progestérone comme cibles. Ensuite, une autre étude d'amarrage moléculaire a été réalisée avec l'aromatase et 17β-HSD1, les enzymes qui catalysent la synthèse des œstrogènes. Enfin, l'activité antiproliférative des 29 dérivés de la FMA contre les lignées cellulaires MCF-7 a été étudiée pour prédire un modèle de relation quantitative structure-activité (QSAR) en utilisant des descripteurs moléculaires et l'analyse MLR.

Les dérivés du FMA ont montré un important effet de cytotoxicité contre les lignées cellulaires MCF-7, qui a été expliqué à l'aide d'un amarrage moléculaire par l'interaction des composés avec le ERα et le PR. En outre, l'étude de l'amarrage moléculaire contre l'aromatase montre que A4 et A25 étaient les meilleurs avec des énergies libres de liaison les plus faibles par rapport aux autres composés et à l'ASD. Cependant, une affinité modérée a été prédite contre 17β-HSD1 pour tous les composés en comparaison avec E2B. Cela indique que ces dérivés pourraient être de puissants inhibiteurs de l'aromatase. La meilleure équation de régression contient les descripteurs HOMO, E, MR, MV, HF, POL, TPSA, qC11 et qN1. Le modèle QSAR indique que ces descripteurs ont des relations significatives avec la bioactivité observée. La capacité prédictive du modèle a été validée à l'intérieur du modèle à l'aide de la méthode "Leave One Out" (LOO), ce qui prouve que le modèle peut être appliqué avec succès pour prédire l'activité antiproliférative contre la lignée cellulaire MCF-7.

الملخص

يعتبر سرطان الثدي هو أكثر أنواع السرطانات شيوعا عند النساء مقارنة بسرطان القولون والمستقيم وسرطان الرئة. أثبتت مشتقات الفيروسين فعاليتها ضد خلايا سرطان الثدي. هدفت هذه الأطروحة إلى إيجاد مشتقات فيروسينية جديدة لها خصائص كيميائية دوائية كأدوية محتملة ضد خلايا سرطان الثدي. تم تهيئة البنية الفراغية لسلسلة من 29 مشتق ل N- فيروسينيل ميثيل انيلين (FMA)لتحديد الخصائص الهيكلية والبلورية والتحليل الطيفي الجزيئي باستخدام طريقة DFT / B3LYP مع القواعد6-311G++(d,p) وLanL2DZ. كما تمت دراسة حركية الدواء وخصائصADME . تم إجراء النشاط المضاد للتكاثر للمشتقات المدروسة ضد كل من خلايا سرطان الثدي البشرية المعتمدة على الهرمونات (MCF-7) وغير المعتمدة على الهرمونات (MDA-MB 231) باستخدام اختبارMTT .علاوة على ذلك ، تمت دراسة العلاقة بين مشتقات FMA و MCF-7 بشكل أكبر من خلال الإرساء الجزيئي باستخدام برنامجAutoDock 4.2 ، باستخدام كل من مستقبلات هرمون الاستروجين والبروجسترون باعتبارها أهدافا. بعد ذلك ، تم إجراء دراسة أخرى بواسطة الإرساء الجزيئي باستخدام أروماتازو β-HSD117 ، وهي الإنزيمات التي تحفز تصنيع الإستروجين. أخيرًا ، تم التحقيق في النشاط المضاد للتكاثر لمشتقات FMA البالغ عددها 29 ضد خلايا MCF-7 للتنبؤ بنموذج علاقة التركيب والنشاط الكمي (QSAR) باستخدام الخصائص الجزيئية وتحليلMLR .

أظهرت مشتقات FMA تأثيرًا مهمًا في السمية الخلوية ضد خلايا MCF-7 ، وهذه الفعالية تم تفسيرها باستخدام الإرساء الجزيئي من خلال تفاعل المركبات مع ERα و PR. علاوة على ذلك ، أظهرت دراسة الإرساء الجزيئي ضد انزيم أروماتاز أن A4 و A25 كانا الأفضل مع أقل طاقة ربط حرة مقارنة بالمركبات الأخرى و ASD. ومع ذلك ، تم توقع تقارب متوسط ضد β-HSD117 لجميع المركبات مقارنة بـ E2B. يشير هذا إلى أن هذه المشتقات قد تكون مثبطات دوائية قوية للأروماتاز. تحتوي أفضل معادلة انحدار على الخصائص الجزيئية HOMO،E،MR،MV،HF،POL،TPSA،qC11 و qN1. يشير نموذج QSAR إلى أن هذه الخصائص لها علاقات مهمة مع النشاط الحيوي الملاحظ سابقا. تم التحقق من صحة النموذج للقدرة التنبؤية داخل النموذج باستخدام طريقةLOO ، الأمر الذي يثبت أنه يمكن تطبيق النموذج بنجاح للتنبؤ بالنشاط المضاد للتكاثر ضد خلاياMCF-7.