



Benzoin-Based Bis-Pyrrole Carbonitriles as Targeted EGFR Inhibitors: Insights from DFT Calculations, Molecular Docking, and Pharmacokinetic Profiling

Abdalhamid Alarafi ^{1*}, Mohammed Alajilani ², Haneen Wali ³, Naowara Alarafi ⁴

^{1,2} Department of Chemistry, Faculty of Arts and Sciences, University of Benghazi, Al-Abyar, Libya

³ Libyan Center for Environmental Sciences and Technology Studies, Sabratha, Libya

⁴ Department of Chemistry, Faculty of Sciences, University of Benghazi, Benghazi, Libya

كربونيتريلات ثنائي البيروول المشتقة من البنزوين كمثبطات مستهدفة لمستقبل عامل النمو المحفز (EGFR) من حسابات نظرية دالة الكثافة، الالتحام الجزيئي، وتوصيف الحركية الدوائية

عبدالحميد العرفي ^{1*}، محمد الجيلاني ²، حنين والي ³، نوارا العرفي ⁴

^{1,2} قسم الكيمياء، كلية العلوم، جامعة بنغازي، الأبيار، ليبيا

³ المركز الليبي لدراسات وبحوث علوم وتكنولوجيا البيئة، صبراتة، ليبيا

⁴ قسم الكيمياء، كلية العلوم، جامعة بنغازي، بنغازي، ليبيا

*Corresponding author: abdahamid.abdallah@uob.edu.ly

Received: August 03, 2026

Accepted: September 30, 2026

Published: October 06, 2026

Abstract:

This study examined four previously synthesized and structurally characterized bispyrrole derivatives to evaluate their potential interactions with the epidermal growth factor receptor (EGFR) and to explore the correlation between molecular structure, electronic properties, and predicted drug-likeness. An integrated computational workflow was applied, combining molecular docking, density functional theory (DFT), and in silico ADMET analysis. Docking simulations were performed against the EGFR tyrosine kinase domain (PDB ID: 1M17), utilizing erlotinib as a reference inhibitor. Redocking of the co-crystallized ligand yielded a root-mean-square deviation (RMSD) of 0.8627 Å, confirming the reliability of the protocol in reproducing the experimental binding pose. Among the evaluated derivatives, compound 2 achieved the most favorable docking score (-9.521 kcal/mol), surpassing erlotinib (-8.24 kcal/mol), with predicted interactions involving LYS721, HIS781, and LEU694. DFT calculations executed at the B3LYP/6-31G level revealed that compound 2 possessed the narrowest HOMO-LUMO gap (3.3423 eV), indicating enhanced electronic reactivity within the

series. Conversely, compound 1 displayed the widest energy gap (4.2945 eV) along with a more favorable predicted drug-likeness profile and a bioavailability score of 0.55.

ADMET predictions identified notable limitations concerning molecular weight, lipophilicity, potential cytochrome P450 inhibition, and projected toxicity. These findings demonstrate that high docking affinity does not strictly parallel favorable pharmacokinetic properties. In conclusion, compound 2 emerges as a promising candidate for experimental follow-up, though its actual EGFR inhibitory activity and therapeutic potential require empirical validation.

Keywords: Bis-pyrrole derivatives; Epidermal growth factor receptor; EGFR; Molecular docking; Density functional theory; Frontier molecular orbitals; ADMET; Drug discovery.

المخلص

تناولت هذه الدراسة أربع مشتقات من ثنائي البيروول تم تخليقها وتوصيفها مسبقاً، وذلك لتقييم تفاعلاتها المحتملة مع مستقبل عامل النمو المتبادل (EGFR) واستكشاف العلاقة بين التركيب الجزيئي والخصائص الإلكترونية وتقدير التشابه الدوائي. تم استخدام منهجية حاسوبية متكاملة تجمع بين الالتحام الجزيئي، ونظرية دالة الكثافة (DFT)، وتحليلات التوزيع الحيوي والامتصاص والتمثيل الغذائي والإخراج والسمية. (ADMET) أجريت عمليات الالتحام ضد نطاق كيناز التيروسين لمستقبل (EGFR) بالرمز (PDB ID: 1M17)، مع استخدام مركب الإبرلوتينيب كمثبط مرجعي. أدى إعادة الالتحام للجاند المتبلور مشاركةً إلى قيمة انحراف معياري جذري متوسط (RMSD) بلغت 0.8627 Å، مما أكد قدرة البرتوكول على إعادة إنتاج وضع الارتباط التجريبي بدقة. ومن بين المشتقات الأربعة، حقق المركب 2 أفضل درجة التهام) -125.9.

كيلو سرعة حرارية/مول)، متفوقاً على مثبط الإبرلوتينيب (-8.42 كيلو سرعة حرارية/مول)، مع وجود تفاعلات متوقعة مع الأحماض الأمينية LYS721 و HIS781 و LEU694. أظهرت حسابات نظرية دالة الكثافة عند مستوى B3LYP/6-31G أن المركب 2 يمتلك أصغر فجوة طاقة بين أعلى مدار مولكولي مشغول وأدنى مدار غير مشغول (3.3243 إلكترون فولت)، مما يشير إلى زيادة تفاعليته الإلكترونية ضمن السلسلة. في المقابل، سجل المركب 1 أوسع فجوة طاقة (4.5492 إلكترون فولت) إلى جانب خصائص تشابه دوائي مقدرة أكثر ملاءمة مع درجات إتاحة حيوية بلغت 0.55. حددت تحليلات (ADMET) قصوراً ملحوظاً مرتبطاً بالوزن الجزيئي، ومُعامل الذوبان الدهني، والتنشيط المحتمل لأنزيم السيتوكروم P450، والسمية المتوقعة. توضح هذه النتائج أن الكفاءة العالية في الالتحام الجزيئي لا تترجم بالضرورة إلى خصائص حركية دوائية مواتية. وخلاصة القول، يعد المركب 2 مرشحاً واعداً للمتابعة التجريبية، على الرغم من أن نشاطه التنشيطي تجاه المستقبل وإمكاناته العلاجية لا تزال بحاجة إلى مؤازرة وتأكيد مخبري.

الكلمات المفتاحية: مشتقات ثنائي البيروول؛ مستقبل عامل النمو المتبادل؛ الالتحام الجزيئي؛ نظرية دالة الكثافة؛ المدارات الجزيئية الحدودية؛ التوزيع الحيوي والسمية؛ اكتشاف الأدوية.

Introduction

Cancer is still among the main causes of death worldwide and places a heavy burden on healthcare systems. GLOBOCAN estimates indicate that about 20 million new cases and 9.7 million cancer deaths were recorded globally in 2022 [1]. Progress in molecular oncology has produced targeted therapies, yet drug resistance, side effects, and wide variation in patient response continue to limit their benefit. For this reason, research

keeps focusing on molecular targets involved in tumor development and progression, especially those linked to abnormal cell signaling .

The epidermal growth factor receptor (EGFR), a member of the ErbB family of receptor tyrosine kinases, takes part in several cellular processes, including proliferation, differentiation, migration, and survival. Ligand binding triggers conformational changes in EGFR that lead to receptor dimerization and activation of the intracellular kinase domain. This activation starts downstream signaling through pathways such as RAS–RAF–MEK–ERK and PI3K–AKT, which control cell growth and survival. Overexpression of the receptor or activating mutations disrupt this signaling and have been associated with the development and progression of various malignancies [2]. EGFR has therefore become a key target in anticancer drug discovery .

EGFR tyrosine kinase inhibitors (TKIs) have markedly improved outcomes for selected patients with EGFR-driven cancers, notably non-small-cell lung cancer (NSCLC). Erlotinib, gefitinib, afatinib, and osimertinib belong to different generations of inhibitors designed to block EGFR-mediated signaling, and clinical studies have confirmed their effectiveness in appropriately selected patients [3]. However, acquired resistance restricts their long-term benefit. Secondary mutations, activation of bypass signaling pathways, and other molecular adaptations can lower treatment efficacy, which underlines the need to explore alternative chemical scaffolds and therapeutic strategies [4].

Nitrogen-containing heterocycles hold an important place in medicinal chemistry, as their structural diversity enables the design of molecules with varied physicochemical and biological properties. Within this group, pyrrole derivatives have drawn considerable interest because they occur in biologically active natural products and in synthetic drugs. Pyrrole-based structures have been studied for a wide range of pharmacological activities, including antimicrobial, anti-inflammatory, and anticancer effects [5]. The pyrrole ring is also versatile, allowing different functional groups and structural modifications to be introduced, which makes it possible to study how molecular architecture affects biological interactions.

The synthetic accessibility of substituted pyrroles has been widely investigated, including through multicomponent reactions that give structurally diverse heterocyclic compounds [6]. Amino, cyano, and aromatic substituents, in particular, can alter molecular polarity, electron distribution, and potential intermolecular interactions. Nevertheless, the mere presence of these groups does not establish biological activity. Their role in molecular recognition depends on their spatial arrangement, their electronic environment, and their compatibility with the target binding site.

In an earlier experimental study, Alarafi et al. (2025) reported the synthesis and structural characterization of several heterocyclic derivatives prepared from benzoin, malononitrile, and various diamines .]7[

That study described bis-pyrrole derivatives with aliphatic, phenylene, and biphenyl linkers, as well as other heterocyclic products formed through alternative reaction pathways. The compounds were characterized by spectroscopic techniques, namely infrared spectroscopy, proton nuclear magnetic resonance (^1H NMR), and carbon-13 nuclear magnetic resonance (^{13}C NMR). These experimental results established the synthetic origin and proposed structures of the compounds examined in the present work.

The earlier study concentrated on synthesis, product formation, and structural characterization. It did not examine how these bis-pyrrole derivatives interact with EGFR, nor did it systematically assess their electronic properties and predicted pharmacokinetic behavior. This gap justifies extending the experimental work with

computational methods. In particular, the different linkers and aromatic substitution patterns make it possible to explore how structural variation affects electronic distribution, molecular flexibility, and potential interactions within the EGFR kinase binding site.

Molecular docking is a structure-based computational approach for studying possible interactions between small molecules and biological macromolecules. By predicting how a ligand is oriented within a defined binding pocket, docking can reveal potential hydrogen bonds, hydrophobic contacts, and other noncovalent interactions involved in molecular recognition. The crystal structure of the EGFR tyrosine kinase domain complexed with erlotinib (PDB ID: 1M17) offers an experimentally determined basis for such studies [8]. Even so, docking scores are approximations produced by scoring functions and should not be read as direct measures of binding affinity or inhibitory potency. Docking is therefore most informative when interpreted together with other molecular properties and experimental evidence.]9[

Density functional theory (DFT) is a complementary method for studying the electronic characteristics of organic compounds. The energies and spatial distributions of the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) indicate electron-donating and electron-accepting tendencies. Related descriptors, such as chemical hardness, softness, and electrophilicity, further help describe electronic behavior. Molecular electrostatic potential (MEP) analysis adds information on how electron-rich and electron-deficient regions are distributed over the molecular surface [10]. These descriptors can help interpret structural differences among related compounds, although they do not necessarily relate directly to biological binding.

Besides molecular recognition and electronic properties, early evaluation of pharmacokinetics and toxicity is essential when selecting compounds for further pharmaceutical study. Poor solubility, excessive lipophilicity, unfavorable absorption, metabolic interactions, and potential toxicity can hinder development even when interaction with the biological target appears favorable. Computational platforms such as SwissADME, ADMETlab 3.0, and ProTox 3.0 offer predictive tools for assessing physicochemical properties, drug-likeness, absorption, distribution, metabolism, excretion, and toxicity [11–13]. Although these methods can reveal possible development-related limitations, their predictions need experimental confirmation. Against this background, the present study provides an integrated computational evaluation of four previously synthesized bis-pyrrole derivatives reported by Alarafi et al. [2025].

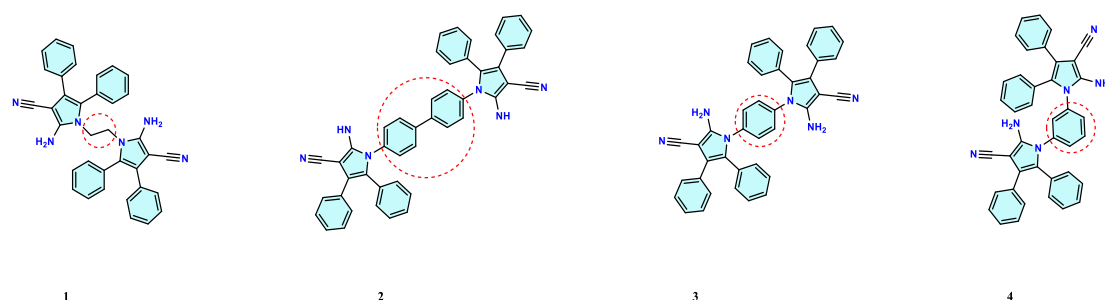
The compounds were studied by molecular docking against the EGFR kinase domain (PDB ID: 1M17), with erlotinib as the reference inhibitor. DFT calculations at the B3LYP/6-31G level were carried out to examine electronic structure, frontier molecular orbital energies, global reactivity descriptors, and molecular electrostatic potential distributions. In silico ADMET predictions were also used to evaluate physicochemical, pharmacokinetic, and toxicological characteristics. By combining these complementary analyses, the study aims to clarify how structural differences influence predicted EGFR interactions and electronic behavior, and to identify limitations that should be addressed before experimental biological evaluation.

2. Computational Experimental

2.1. Chemical Structures

The four bis-pyrrole derivatives examined here were synthesized and structurally characterized earlier by Alarafi et al. (2025) [7]. According to that report, the

compounds were prepared from benzoic acid, malononitrile, and various diamines, and their structures were confirmed by spectroscopic techniques. For the present work, compounds I, III, IV, and VI of the original study were renumbered as compounds 1, 2, 3, and 4, respectively. Their two-dimensional structures were redrawn in ChemDraw 2023 and then converted into three-dimensional conformations for the computational analyses. The structures and their substitution patterns are shown in Scheme 1.



Scheme 1. General chemical structures and substitution pattern of the investigated compounds 1–4

2.2 Ligand preparation

The initial 3D conformations of ligands 1–4 were optimized with the Energy Minimize module in MOE 2024 [51]. After partial charges were assigned, energy minimization was run with the Amber:EHT force field and an implicit reaction-field solvation model (R-Field 1:80). Minimization continued until the RMS gradient fell below 0.1 kcal/mol/Å².

2.2. Receptor Preparation

The 3D crystal structure of the catalytic kinase domain of human epidermal growth factor receptor (EGFR; PDB ID: 1M17, chain A), co-crystallized with its selective inhibitor erlotinib (AQ4), was downloaded from the RCSB Protein Data Bank (Figure 1). Before docking, all crystallographic water molecules were removed, while the cocrystallized ligand was kept to define the active site. The protein was then prepared with the QuickPrep module in MOE 2024. This protocol kept the chain sequence, neutralized electrostatic charges, assigned protonation states in 3D with Protonate3D (including side-chain flips for ASN, GLN, and HIS residues), and added missing hydrogen atoms. A tethered energy minimization followed (strength: 10; buffer: 0.25 Å; atoms farther than 8.0 Å fixed) until the RMS gradient dropped below 0.1 kcal/mol/Å².

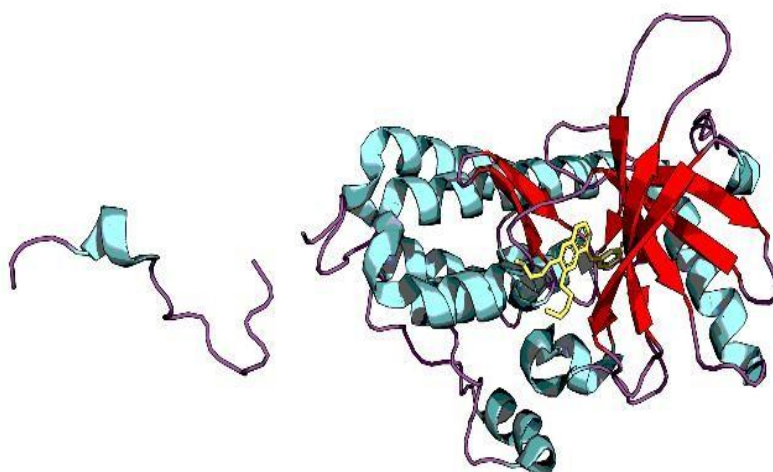


Figure 1. 3D representation of the single-chain human EGFR kinase domain (PDB ID: 1M17, Chain A) co-crystallized with its selective inhibitor, erlotinib (AQ4).

2.3. Molecular Docking and Validation Protocol

Docking calculations and scoring were carried out in the Molecular Operating Environment (MOE 2024). The crystal structure of the EGFR kinase domain (PDB ID: 1M17) was solved by X-ray diffraction at a resolution of 2.60 Å. Resolutions between 1.5 and 3.0 Å are generally regarded as suitable for docking, as they provide reliable structural detail and an accurate binding pocket.

The docking protocol was validated by removing the co-crystallized native inhibitor (erlotinib/AQ4) from the binding pocket and re-docking it into the same site with identical docking parameters. The protocol was considered reliable because the root mean square deviation (RMSD) was below the accepted threshold of 2.0 Å and the binding free energy score was favorable.

2.4. Computational Details and DFT Calculations

Density functional theory (DFT) calculations were performed with the Gaussian package and GaussView to evaluate the structural, electronic, and global reactivity properties of the target compounds, as well as their molecular electrostatic potential (MEP/ESP) maps. Full geometry optimizations were carried out in the gas phase using the hybrid B3LYP functional with the 6-31G basis set. Convergence was reached under the standard optimization criteria, and the very low RMS force gradients confirmed that the optimized geometries were stable stationary points on the potential energy surface (PES).

Within conceptual DFT (c-DFT), the global reactivity descriptors were derived from the frontier molecular orbital energies (E_{HOMO} and E_{LUMO}) using the following expressions:

$$\text{Energy Gap: } \Delta E = E_{\text{LUMO}} - E_{\text{HOMO}}$$

$$\text{Chemical Potential: } \mu = \frac{E_{\text{HOMO}} + E_{\text{LUMO}}}{2}$$

$$\text{Global Hardness: } \eta = \frac{E_{\text{LUMO}} - E_{\text{HOMO}}}{2}$$

$$\text{Global Softness: } S = \frac{1}{2\eta}$$

$$\text{Electrophilicity Index: } \omega = \frac{\mu^2}{2\eta}$$

$$\text{Nucleophilicity Index: } N = E_{\text{HOMO}(\text{Compound})} - E_{\text{HOMO}(\text{TCE})}$$

Moreover, 3D molecular electrostatic potential (MEP/ESP) surfaces were generated at the optimized B3LYP/6-31G level to examine charge distribution and to locate the sites likely to take part in molecular interactions: electron-rich regions (red), which are prone to electrophilic attack, and electron-deficient regions (blue), which are prone to nucleophilic attack.

2.5 In silico ADME and toxicity profiling

Pharmacokinetic, physicochemical, and toxicity characteristics were assessed to establish drug-likeness. Lipinski's rule compliance and various physicochemical

characteristics were estimated using SwissADME. Comprehensive ADME endpoints (absorption, clearance, metabolic stability) were assessed using ADMETLab 3.0, while organ toxicity, cytotoxicity, and LD₅₀ values were derived using ProTox 3.0.

3.Results and discussion

3.1 Docking Protocol Validation

To assess the accuracy and reproducibility of the docking protocol before docking the test compounds, a self-docking (re-docking) experiment was performed. The cocrystallized inhibitor erlotinib (AQ4) was removed from the binding pocket of the human EGFR kinase domain (PDB ID: 1M17) and re-docked under identical settings. As shown in Figure 2, the top-ranked re-docked pose (yellow) superimposed closely on the native crystallographic conformation (pink), with an RMSD of 0.8627 Å. This value is well below the commonly accepted threshold of < 2.0 Å, indicating that the protocol can reproduce the experimentally observed binding geometry. To rigorously evaluate the accuracy, predictive capacity, and reproducibility of the applied molecular docking protocol prior to virtual screening, a self-docking (re-docking) study was performed. The co-crystallized selective inhibitor, erlotinib (AQ4), was extracted from the active binding pocket of the human EGFR kinase domain (PDB ID: 1M17) and subsequently re-docked under identical computational settings.

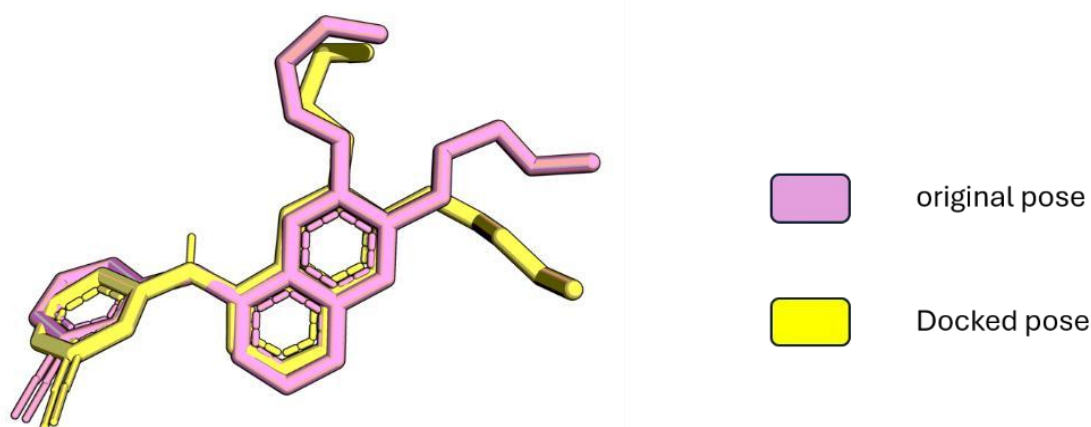


Figure 2. Validation of the docking protocol by re-docking erlotinib (AQ4) into the catalytic active site of human EGFR (PDB ID: 1M17)

3.2. Molecular Docking Analysis and Comparative Binding Modes

To characterize the binding modes and the structural features that govern the interaction of derivatives 1-4 with EGFR, docking simulations were run in the ATP-binding site of the human EGFR kinase domain (PDB ID: 1M17). As listed in Table 1 and shown in the 2D interaction diagrams (Figure 3.2) and 3D binding poses (Figure 3.3), all four compounds obtained favorable docking scores, ranging from -7.37 to -9.521 kcal/mol, with ($\text{RMSD} \leq 1.655$ Å) for the refined poses. Compound 2 had the best score, -9.521 kcal/mol, ($\text{RMSD} = 1.171$ Å) exceeding that of the reference inhibitor erlotinib ($S = -8.24$ kcal/mol, $\text{RMSD} = 0.8627$ Å).

The binding mode of compound 2 combines two hydrogen bonds with a hydrophobic π -H interaction. N27 bonds with the catalytic residue LYS 721 (3.21 Å, $E = -1.3$ kcal/mol), and N59 bonds with HIS 781 (3.58 Å, $E = -1.1$ kcal/mol). In addition, its 6-membered aromatic ring contacts LEU 694 (4.30 Å, $E = -0.5$ kcal/mol), which appears to hold the molecule deep in the pocket (Figures 3.2B and 3.3B). Compound 3 ($S = -8.32$ kcal/mol, $\text{RMSD} = 1.554$ Å), also scored better than the reference drug. It

formed an H-acceptor bond with LYS 721 (N17, 3.18 Å, $E = -1.5$ kcal/mol) and a π -H contact with VAL 702 (3.59 Å, $E = -0.7$ kcal/mol) (Figures 3.2C and 3.3C). Compound 4 ($S = -7.985$ kcal/mol, RMSD = 1.574 Å) was the only derivative to reproduce the hinge-region interaction typical of erlotinib, forming a hydrogen bond with MET 769 (N40, 3.22 Å, $E = -1.2$ kcal/mol). It was further stabilized by contacts with PRO 770 (3.57 Å), LYS 721 (3.19 Å), LYS 721 (N17, 3.18 Å, $E = -1.5$ kcal/mol), and VAL 702 (3.59 Å, $E = -0.7$ kcal/mol) (Figures 3.2D and 3.3D). Compound 1 ($S = -7.37$ kcal/mol, RMSD = 1.655 Å) adopted a stable pose held by an H-bond with LYS 721 (N13, 3.21 Å, $E = -1.3$ kcal/mol) and by two π -H contacts, with VAL 702 (3.56 Å) and CYS 773 (3.92 Å) (Figures 3.2A and 3.3A). Taken together, these electrostatic and hydrophobic interactions indicate a good steric and chemical fit of the derivatives, especially compound 2, within the EGFR active site. This supports their selection for experimental testing of EGFR inhibition.

Table 1. The results obtained from docking of Ligands with 1M17 (EGFR).

Ligand	S score (kcal/mol)	RMSD (Å)	Atom of compound	Atom of receptor	Involved receptor residues	Type of interaction bond	Distance (Å)	E(kcal/mol)
1	-7.37	1.655	N 13 5-ring 5-ring	NZ CG2 CB	LYS 721 VAL 702 CYS 773	H-acceptor pi-H pi-H	3.21 3.56 3.92	-1.3 -0.5 -0.6
2	-9.521	1.171	N 27 N 59 6-ring	NZ NE2 CD1	LYS 721 HIS 781 LEU 694	H-acceptor H-acceptor pi-H	3.21 3.58 4.30	-1.3 -1.1 -0.5
3	-8.32	1.554	N 17 5-ring	NZ CG2	LYS 721 VAL 702	H-acceptor pi-H	3.18 3.59	-1.5 -0.7
4	-7.985	1.574	N 40 N 17 N 49 5-ring	O CA NZ CG2	MET 769 PRO 770 LYS 721 VAL 702	H-donor H-acceptor H-acceptor pi-H	3.22 3.57 3.19 3.58	-1.2 -0.5 -1.2 -0.5
Reference Ligand	-8.24	0.8627	N2 44	N	MET 769	H-acceptor	3.00	-4.2

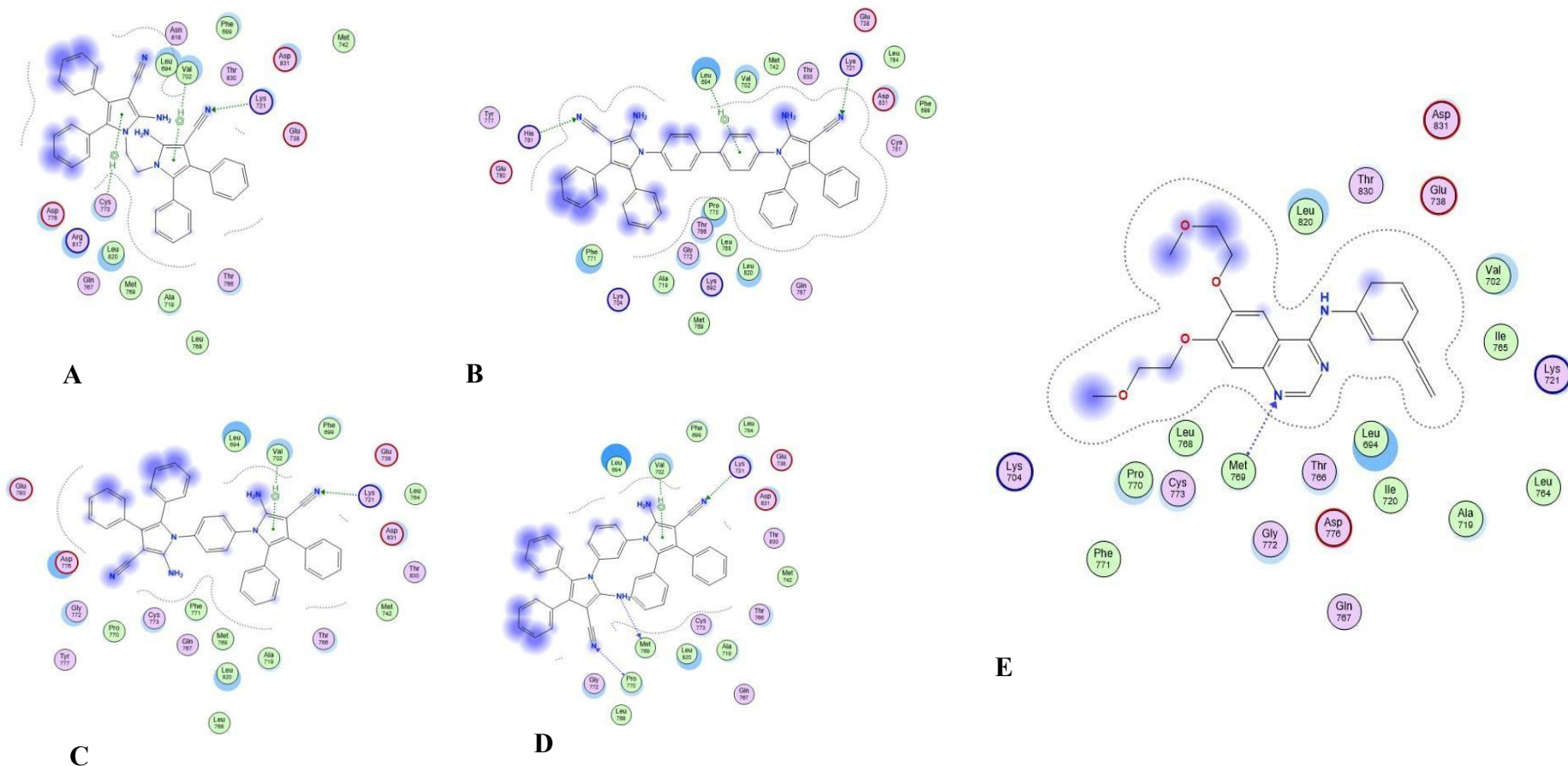
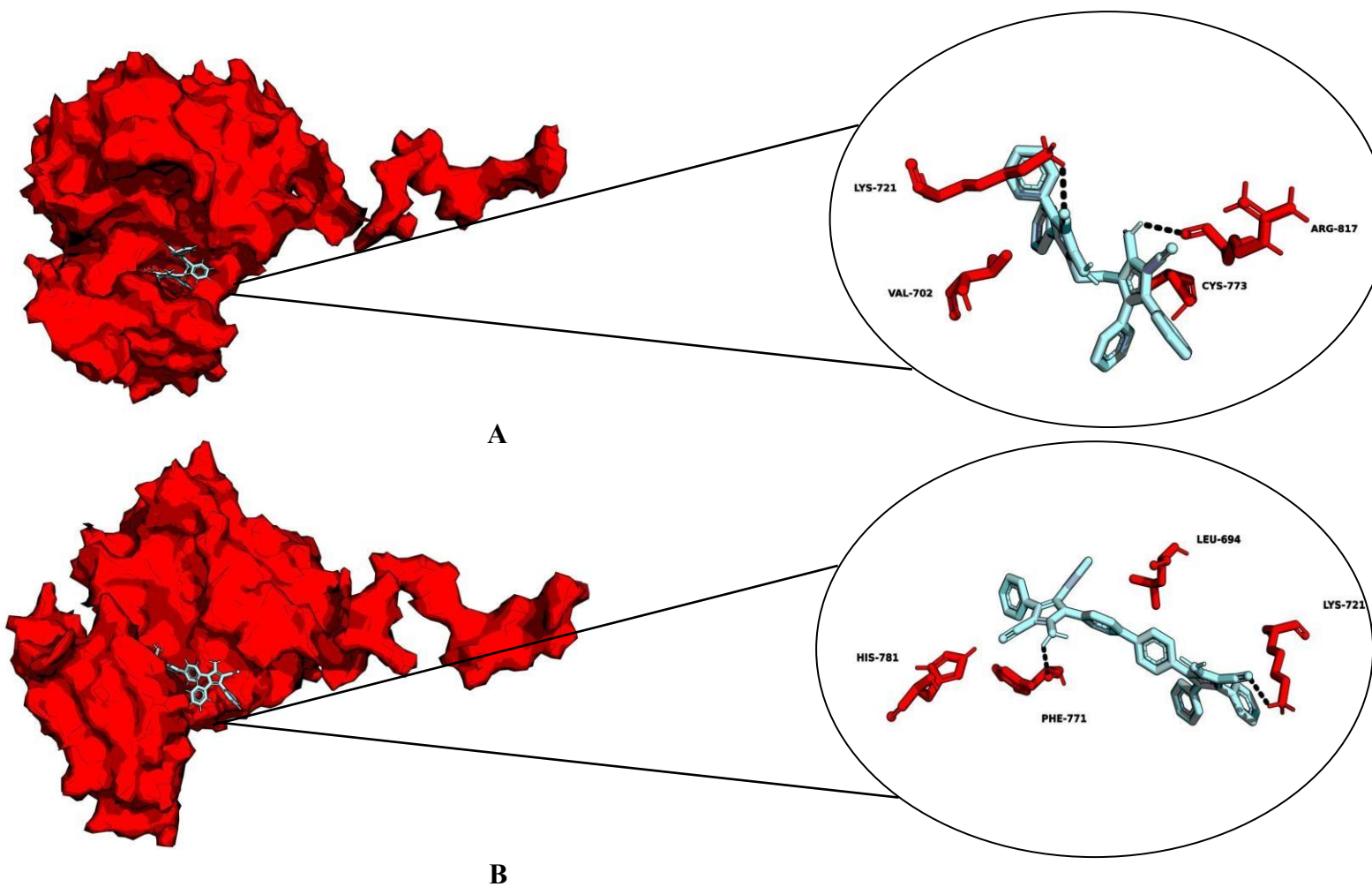
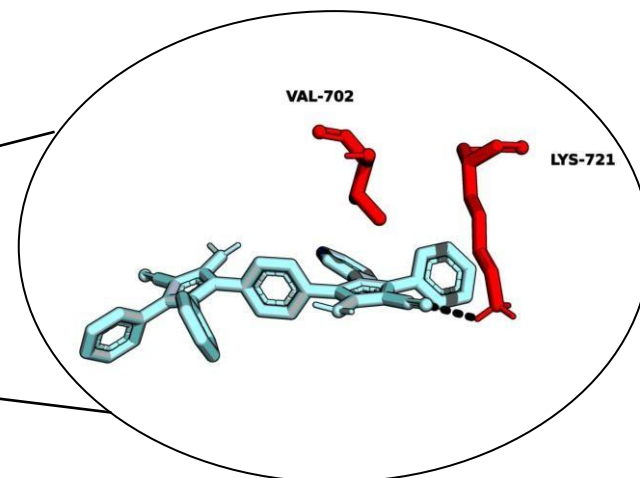
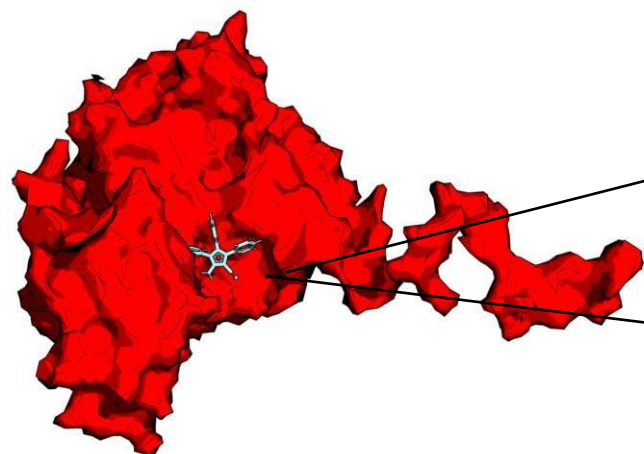
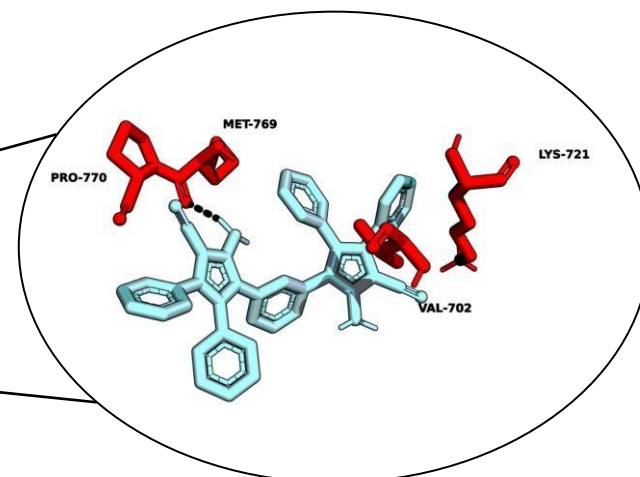
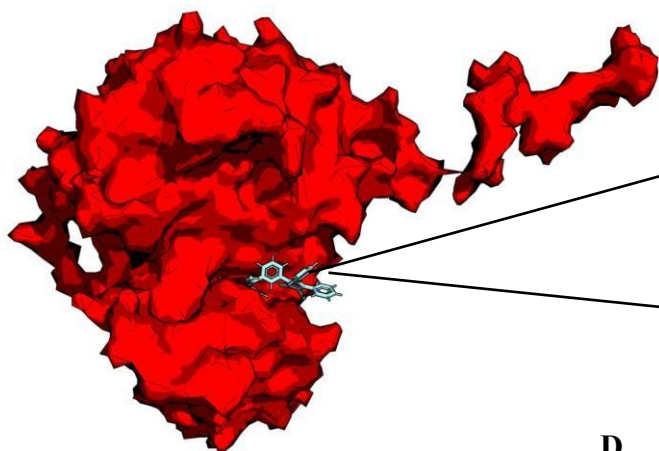


Figure 3. 2D molecular interaction diagrams of target derivatives and the co-crystallized reference inhibitor within the catalytic active site of human EGFR (PDB ID: 1M17): (A) Compound 1, (B) Compound 2, (C) Compound 3, (D) Compound 4, and (E) Reference ligand (Erlotinib / AQ4).





C



D

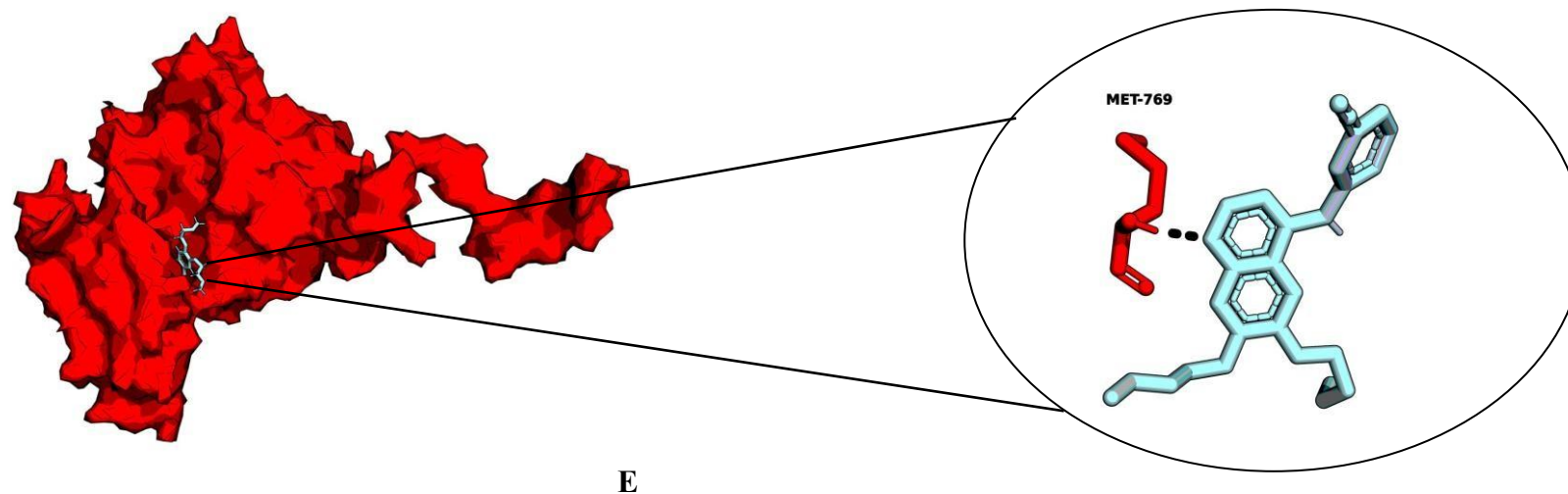


Figure 4. 3D molecular interaction diagrams of target derivatives and the co-crystallized reference inhibitor within the catalytic active site of human EGFR (PDB ID: 1M17): (A) Compound 1, (B) Compound 2, (C) Compound 3, (D) Compound 4, and (E) Reference ligand (**Erlotinib** / **AQ4**).

3.3 Quantum Chemical Calculations and Geometric Optimization

Full geometry optimizations for the investigated organic derivatives (Compounds 1–4) were carried out at the DFT/B3LYP level using the 6-31G basis set in the gas phase. The absence of imaginary frequencies along with the stringent convergence criteria achieved for the root-mean-square (RMS) force gradients (0.000002 –

0.000004 Hartree/Bohr) conclusively validates that all optimized conformers reside at local energy minima on the potential energy surface (PES)

The ground-state electronic total energies and dipole moments (μ) are summarized in Table 1. As expected, Compound 2 exhibits the lowest total electronic energy (–2100.157390 Hartree), demonstrating superior thermodynamic stability relative to the remaining members of the series. Interestingly, Compound 3 and Compound 4 display nearly identical electronic energies (–1869.148613 Hartree and –1869.149949 Hartree, respectively), which strongly indicates structural isomerism with subtle differences in spatial molecular arrangement.

Molecular polarity plays a crucial role in predicting molecular solubility, solvent-solute interactions, and binding affinities toward bio-macromolecular receptors or surfaces.

The calculated dipole moments reveal a wide polarity span following the order:

Compound 1 (10.0751 D) > Compound 4 (7.2858 D) > Compound 3 (4.0346 D)
> Compound 2 (1.1107 D)

The exceptionally high dipole moment observed for Compound 1 (10.0751 D) points to a highly asymmetric electron density distribution, inducing pronounced intramolecular charge transfer (ICT) from electron-donating moieties to electron-withdrawing centers. Conversely, the lower dipole moment of Compound 2 (1.1107 D) reflects a more symmetric charge distribution across its framework.

Table 2. Calculated Total Energies, Dipole Moments, and RMS Gradient Norms of the Studied Compounds.

Compound	Total Energy (Hartree)	Total Energy (eV)	Dipole Moment (μ , Debye)	RMS Gradient Norm (Hartree/Bohr)
1	-1716.754813	-46715.30	10.0751	0.000003
2	-2100.157390	-57148.23	1.1107	0.000002
3	-1869.148613	-50862.15	4.0346	0.000003
4	-1869.149949	-50862.18	7.2858	0.000004

3.4 Frontier Molecular Orbital (FMO) Analysis

Frontier molecular orbitals (FMOs), namely the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO), determine the chemical reactivity, kinetic stability, and optoelectronic behavior of organic molecules. The HOMO energy (E_{HOMO}) correlates with the electron-donating ability, whereas the LUMO energy (E_{LUMO}) reflects the electron-accepting capacity. The computed FMO energy levels and the corresponding energy gaps ($\Delta E = E_{\text{LUMO}} - E_{\text{HOMO}}$) are detailed in Table 2, while the spatial distribution maps are illustrated in Figure 1.

The energy gap (ΔE) serves as an established index for chemical reactivity and kinetic stability; a larger energy gap signifies high kinetic stability and low chemical reactivity (a "hard" molecule), whereas a smaller gap denotes higher polarizability and reactivity (a "soft" molecule). Among the evaluated series, the energy gap increases in the following order:

Compound 2 (3.3423 eV) < Compound 3 (3.7367 eV) < Compound 4 (3.7458 eV)
< Compound 1 (4.2945 eV)

Compound 2 possesses the narrowest energy gap ($\Delta E = 3.3423$ eV), rendering it the most kinetically reactive and most easily polarizable molecule in the series. This reduced bandgap is primarily attributed to a lower E_{LUMO} (-1.7821 eV), which facilitates charge transfer processes upon electronic excitation. In contrast, Compound 1 exhibits the largest energy gap ($\Delta E = 4.2945$ eV) alongside the highest E_{HOMO} (-5.0893 eV), indicating superior intrinsic kinetic stability and enhanced electron donor potential compared to the other derivatives.

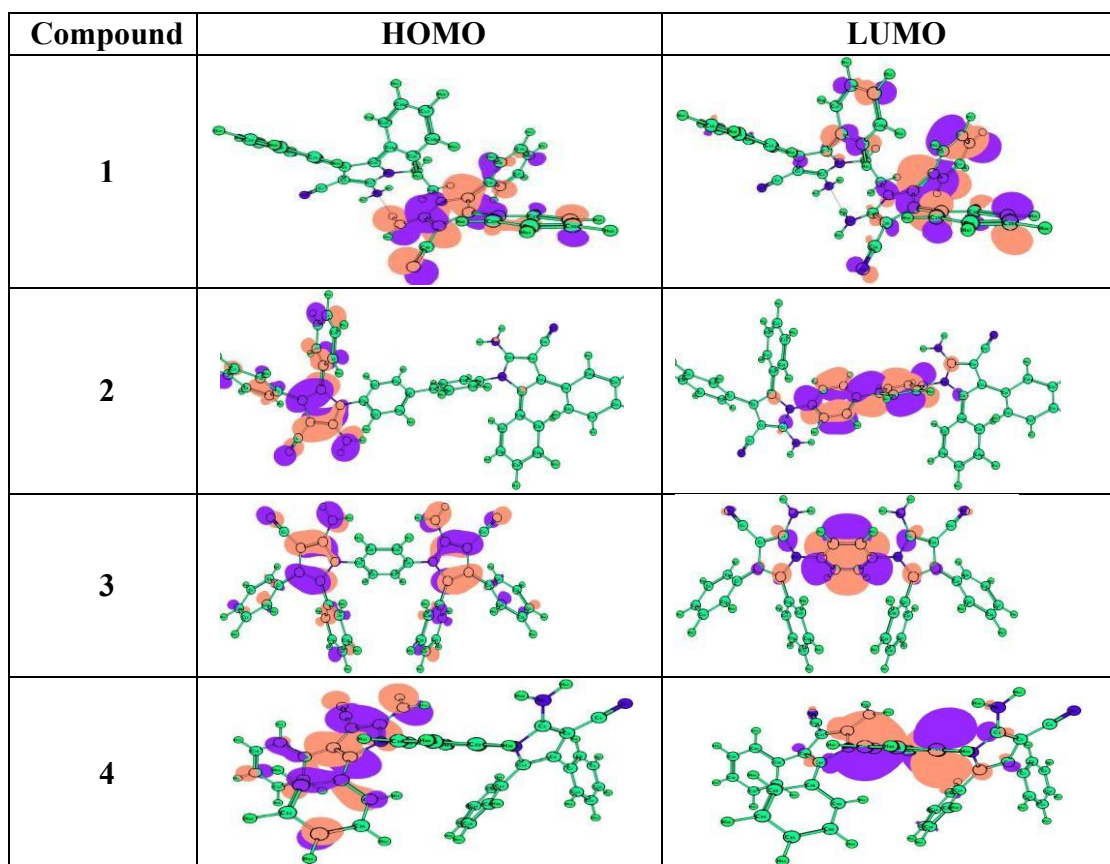


Figure 5. Spatial distribution and 3D iso-surface energy level representations of the frontier molecular orbitals (HOMO and LUMO) for Compounds 1–4 computed at the DFT/B3LYP/6-31G level of theory.

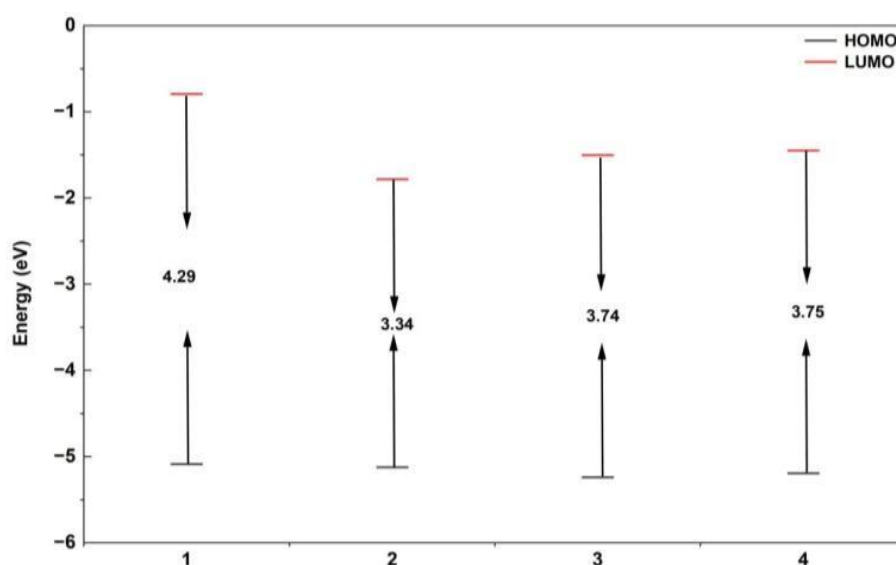


Figure 6. HOMO and LUMO energy levels and corresponding energy gaps of the four investigated compound(1 -4).

3.5 Global Reactivity Descriptors

To further quantitatively rationalize the chemical behavior of Compounds 1–4, conceptual DFT reactivity indices—including chemical hardness (η), softness (S), chemical potential (μ), global electrophilicity index (ω), and nucleophilicity index (N)—were evaluated using Koopmans' theorem approximations (Table 2).

3.5.1 Chemical Hardness (η) and Softness (S)

In line with the FMO bandgap trends, Compound 1 exhibits the highest chemical hardness ($\eta = 2.1473$ eV), confirming its strong resistance toward intramolecular charge redistribution under external chemical perturbations. Conversely, Compound 2 displays the highest chemical softness ($S = 0.5984$ eV⁻¹), highlighting its high reactivity toward biological targets, nucleophilic/electrophilic attack, or surface adsorption applications.

3.5.2 Global Electrophilicity (ω) and Nucleophilicity (N)

The global electrophilicity index (ω) measures the stabilization energy gained when a system acquires an additional electronic charge from the environment. The computed electrophilicity values follow the sequence:

Compound 2 (3.5680 eV) > Compound 3 (3.0450 eV) > Compound 4 (2.9458 eV)
> Compound 1 (2.0155 eV)

Compound 2 is characterized as a strong electrophile ($\omega = 3.5680$ eV), driven by its stabilized LUMO level. Regarding nucleophilicity (N), evaluated relative to tetracyanoethylene (TCE, $E_{\text{HOMO}} = -8.6559$ eV), all four compounds display strong nucleophilic character ($N > 3.4$ eV), with Compound 1 ($N = 3.5666$ eV) and Compound 2 ($N = 3.5315$ eV) ranking as the most potent nucleophiles.

Table 3: Frontier Molecular Orbital Energies and Conceptual DFT Global Reactivity Indices.

Compound	HOMO (eV)	LUMO (eV)	ΔE (eV)	η (eV)	S (eV ⁻¹)	μ (eV)	ω (eV)	N (eV)
1	-5.0893	-0.7948	4.2945	2.1473	0.4657	-2.9421	2.0155	3.5666
2	-5.1244	-1.7821	3.3423	1.6712	0.5984	-3.4533	3.5680	3.5315
3	-5.2415	-1.5048	3.7367	1.8684	0.5352	-3.3731	3.0450	3.4144
4	-5.1947	-1.4489	3.7458	1.8729	0.5339	-3.3218	2.9458	3.4612

3.6 Molecular Electrostatic Potential (MEP/ESP) Analysis

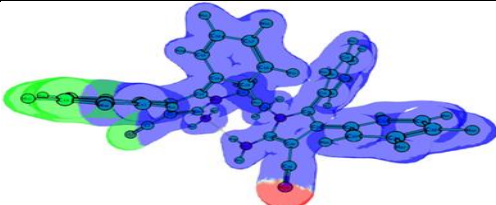
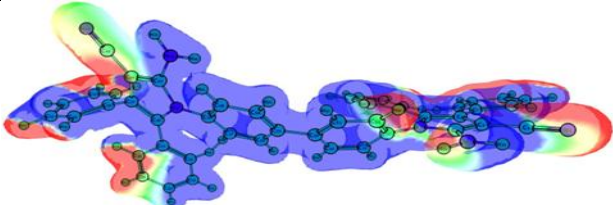
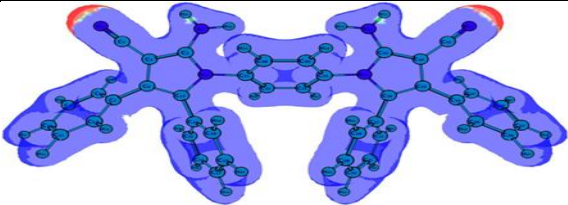
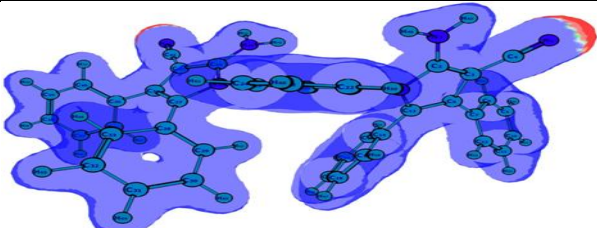
Molecular electrostatic potential (MEP) analysis describes the electrostatic character of a molecular surface and helps locate regions that may favor interaction with electrophilic or nucleophilic species. In the conventional MEP representation, regions of negative potential are shown in red and are favorable for attack by electrophiles, while regions of positive potential are shown in blue and may favor attack by nucleophiles. How these regions are read depends on the calculated potential and on the color scale used for display.

In the bis-pyrrole derivatives studied here, the cyano and amino groups are expected to cause variation in the surface potential. The electronegative nitrogen atoms of the cyano groups may contribute to negative regions, whereas the hydrogen atoms of the amino groups may contribute to positive regions. The exact spatial distribution and magnitude of these potentials should be confirmed from the calculated MEP surfaces .

The calculated dipole moments show clear differences in polarity among the compounds. Compound 1 has the highest dipole moment (10.0751 D), suggesting a relatively asymmetric charge distribution. Compound 2 has the lowest (1.1107 D), which may reflect partial cancellation of individual bond dipoles owing to its geometry. These results suggest that the arrangement of polar functional groups strongly influences overall molecular polarity.

Comparing compounds 3 and 4 illustrates how structural isomerism affects polarity. The two compounds share the same molecular formula, but their different substitution patterns give distinct dipole moments of 4.0346 D and 7.2858 D, respectively. The higher value for compound 4 may be explained by differences in the spatial orientation and vector sum of its polar bond contributions.

In summary, the calculated electrostatic features and dipole moments indicate that changes in molecular architecture alter the polarity of these bis-pyrrole derivatives, and such features may influence their electrostatic interactions with surrounding molecules and biological environments. However, MEP and dipole moment analyses alone cannot establish binding affinity, biological activity, or significant intramolecular charge transfer, so they should be interpreted together with the docking and frontier molecular orbital results.

Compound	MEP/ESP
1	
2	
3	
4	

3.7 ADME and toxicity profile analysis

3.7.1 Physicochemical Properties and Drug-Likeness

the physicochemical evaluation showed significant discrepancies in the drug-likeness characteristics of the studied compounds. Compound 1 displayed the best attributes, featuring the lowest molecular weight (544.65 g/mol), just one Lipinski violation, and the highest estimated bioavailability score (0.55). Conversely, compounds 2–4 exhibited greater molecular weights, numerous drug-likeness infractions, and diminished bioavailability scores (0.17). Even though all compounds met Veber's criteria, their differences from the Ghose, Egan, and Muegge

filters indicate possible constraints on their oral drug-like characteristics. In general, compound 1 exhibited a relatively advantageous physicochemical profile, although experimental confirmation is still required.

3.7.2 ADMET Properties

The expected ADMET profiles indicated variations in the pharmacokinetic and toxicological properties of the examined compounds. Permeability forecasts fluctuated moderately, with compound 1 exhibiting the highest numerical BBB score, indicating variations in possible tissue distribution. Metabolic forecasts revealed a strong possibility of CYP2C19, CYP2C9, and CYP2C8 inhibition for all compounds, emphasizing potential drug–drug interaction issues. Toxicity forecasts indicated that compound 1 had the lowest projected LD50 (500 mg/kg), while compounds 2–4 displayed greater values (2850–4000 mg/kg), implying relatively lower predicted acute toxicity. Nonetheless, the organ-specific toxicity ratings suggest possible safety issues that necessitate additional research. Together, these results underscore a balance between anticipated drug-likeness and toxicological characteristics, stressing the necessity of experimental confirmation prior to advancing drug development.

Table 4. Physicochemical properties and drug-likeness for Compounds (1-4)													
Ligands	Physiological properties							Drug likeness properties					
	MW (g/mol)	nHA	F. Csp3	nRB	nHBA	nHBD	TPSA (Å²)	Lipinski	Ghose	Veber	Egan	Muegge	Bio. score
1	544.65	42	0.06	7	2	2	109.48	Yes; 1 violation: MW>500	No; 3 violations: MW>480, WLOGP>5.6, MR>130	Yes	No; 1 violation: WLOGP>5.88	No; 1 violation: XLOGP3>5	0.55
2	666.77	52	0.00	7	4	2	81.50	No; 2 violations: MW>500, MLOGP>4.15	No; 4 violations: MW>480, WLOGP>5.6, MR>130, #atoms>70	Yes	No; 1 violation: WLOGP>5.88	No; 3 violations: MW>600, XLOGP3>5, #rings>7	0.17
3	592.69	46	0.00	6	2	2	109.48	No; 2 violations: MW>500, MLOGP>4.15	No; 4 violations: MW>480, WLOGP>5.6, MR>130, #atoms>70	Yes	No; 1 violation: WLOGP>5.88	No; 1 violation: XLOGP3>5	0.17
4	592.69	46	0.00	6	2	2	109.48	No; 2 violations: MW>500, MLOGP>4.15	No; 4 violations: MW>480, WLOGP>5.6, MR>130, #atoms>70	Yes	No; 1 violation: WLOGP>5.88	No; 1 violation: XLOGP3>5	0.17

Table 5. In silico ADMET profiling of Compounds 1–4 comprehensive pharmacokinetic and toxicity prediction

Entry	1	2	3	4
Absorption/distribution				
CaCo2	-5.02532	-4.88748	-4.98085	-4.99525
MDCK	2.32E-05	1.84E-05	2.06E-05	2.07E-05
pgb substrate	0.000155	1.02E-09	3.63E-08	3.15E-08
VDss	3.169657	4.955436	4.744457	4.814697
BBB	0.204446	1.32E-05	0.000184	0.000155
OATP1B1	0.000405	0.009859	0.000538	0.00055

OATP1B3	0.074949	0.851443	0.814393	0.865526
Metabolism				
CYP1A2-inh	0.999999642	0.00052026	0.81057781	0.981003642
CYP2C19-inh	1	0.999722302	1	1
CYP2C9-inh	1	0.999999762	1	1
CYP2D6-inh	9.93286E-05	0.00034692	3.51544E-08	9.85718E-08
CYP3A4-sub	0.141241595	0.998482883	0.219666988	1.08741E-05
CYP2B6-sub	0.460688859	0.00106927	3.78371E-06	1.88765E-10
CYP2C8-inh	0.998313665	0.999997139	1	1
Excretion				
cl-plasma	3.583039	1.554248	2.366557	2.546942
T1/2	0.578934	0.808281	0.751476	0.676237
Organ toxicity prediction				
Predicted LD50/	500mg/kg	2850mg/kg	4000mg/kg	4000mg/kg
Predicted Toxicity Class	4	5	5	5
Hepatotoxicity	0.81	0.54	0.54	0.54
Neurotoxicity-	0.83	0.81	0.81	0.81
Nephrotoxicity-	0.82	0.86	0.86	0.86
Respiratory toxicity	0.69	0.70	0.70	0.70
Cardiotoxicity	0.89	0.94	0.94	0.94

Conclusion

This research offered a comparative computational assessment of four structurally related organic derivatives via molecular docking, DFT calculations, and in silico ADMET evaluation. The results showcased clear disparities in their anticipated interactions with the EGFR kinase domain, electronic properties, and pharmacokinetic behaviors. Of the compounds studied, compound 2 demonstrated the best docking score (−9.521 kcal/mol), surpassing the reference inhibitor erlotinib (−8.24 kcal/mol) in the used scoring system. The anticipated interactions with LYS721, HIS781, and LEU694 indicate a beneficial binding orientation within the receptor's active site.

The DFT findings also differentiated compound 2 due to its comparatively narrow HOMO–LUMO energy gap (3.3423 eV) and elevated chemical softness, suggesting increased electronic sensitivity within the analyzed series. In comparison, compound 1 showed the largest energy gap (4.2945 eV), indicating relatively lower electronic reactivity within the chosen computational model. These results demonstrate how changes in structure can affect electronic characteristics and anticipated molecular interactions.

Significantly, the ADMET evaluation showed that good docking results did not automatically indicate enhanced drug-like properties. While compound 1 demonstrated a relatively advantageous physicochemical profile, compounds 2–4 presented elevated predicted LD50 values, indicating reduced estimated acute toxicity. Nonetheless, the detected drug-likeness infringements, possible cytochrome P450 inhibition, and predictions of organ-specific toxicity suggest that additional optimization and experimental assessment are required.

Collectively, the findings endorse the importance of combining molecular docking, quantum chemical computations, and ADMET forecasts to analyze structurally similar compounds and assist in their initial prioritization. Compound 2 could be regarded as a potential subject for additional study due to its docking performance and electronic properties; nonetheless, these computational results do not confirm EGFR inhibitory activity or therapeutic effectiveness. Subsequent research should encompass biochemical assays for EGFR inhibition, investigations into cellular activity and selectivity, as well as experimental evaluations of pharmacokinetics and toxicity to establish the biological significance of the predicted characteristics.

Compliance with ethical standards

Disclosure of conflict of interest

The authors declare that they have no conflict of interest.

References

- [1] Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R. L., Soerjomataram, I., & Jemal, A. (2024). Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 74(3), 229–263. DOI: 10.3322/caac.21834 .
- [2] Lemmon, M. A., Schlessinger, J., & Ferguson, K. M. (2014). The EGFR family: Not so prototypical receptor tyrosine kinases. *Cold Spring Harbor Perspectives in Biology*, 6(4), a020768. DOI: 10.1101/cshperspect.a020768 .
- [3] Soria, J. C., Ohe, Y., Vansteenkiste, J., et al. (2018). Osimertinib in untreated EGFRmutated advanced non-small-cell lung cancer. *The New England Journal of Medicine*, 378(2), 113–125. DOI: 10.1056/NEJMoa1713137 .
- [4] Gao, C., Wang, W., Liu, T., Li, X., Yu, Y., & Wu, J. (2025). Annual review of EGFR inhibitors in 2024. *European Journal of Medicinal Chemistry*, 292, 117677. DOI: 10.1016/j.ejmech.2025.117677 .
- [5] Bhardwaj, V., Gumber, D., Abbot, V., Dhiman, S., & Sharma, P. (2015). Pyrrole: A resourceful small molecule in key medicinal hetero-aromatics. *RSC Advances*, 5, 15233–15266. DOI: 10.1039/C4RA15710A .
- [6] Estévez, V., Villacampa, M., & Menéndez, J. C. (2014). Recent advances in the synthesis of pyrroles by multicomponent reactions. *Chemical Society Reviews*, 43, 4633–4657. DOI: 10.1039/C3CS60015G .
- [7] Alarafi, N., ALajilani, M., Aladiwish, W., Elfarsi, N., & Gheath, A. (2025). Synthesis and reactions of bis amino pyrroles based benzoin. *The North African Journal of Scientific Publishing*, 3(3), 213–218. DOI: 10.65414/najsp.v3i3.565 .
- [8] Stamos, J., Sliwowski, M. X., & Eigenbrot, C. (2002). Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. *Journal of Biological Chemistry*, 277, 46265–46272. DOI: 10.1074/jbc.M207135200 .
- [9] Pinzi, L., & Rastelli, G. (2019). Molecular docking: Shifting paradigms in drug discovery. *International Journal of Molecular Sciences*, 20(18), 4331. DOI: 10.3390/ijms20184331 .
- [10] Geerlings, P., De Proft, F., & Langenaeker, W. (2003). Conceptual density functional theory. *Chemical Reviews*, 103(5), 1793–1874. DOI: 10.1021/cr990029p .
- [11] Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports*, 7, 42717. DOI: 10.1038/srep42717 .
- [12] Fu, L., Shi, S., Yi, J., et al. (2024). ADMETlab 3.0: An updated comprehensive online ADMET prediction platform enhanced with broader coverage, improved

performance, API functionality and decision support. *Nucleic Acids Research*, 52(W1), W422–W431. DOI: 10.1093/nar/gkae236 .

- [13] Banerjee, P., Kemmler, E., Dunkel, M., & Preissner, R. (2024). ProTox 3.0: A webserver for the prediction of toxicity of chemicals. *Nucleic Acids Research*, 52(W1), W513–W520. DOI: 10.1093/nar/gkae303.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of **SJPHRT** and/or the editor(s). **SJPHRT** and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.