

Quantum Computational, Structural Characterization, Hirshfeld Surface, Electronic Properties and Molecular Docking Studies on N-(2,3,5,6-Tetra-Fluoropyridin-4-yl) Formamide

Bayrakdar Alpaslan^{*+}

Department of Medical Services and Techniques, Vocational School of Health Services, Iğdır University, Iğdır, TURKEY

ABSTRACT: *In this study, the structural, spectroscopic, and electronic properties of N-(2,3,5,6-Tetrafluoropyridin-4-yl) formamide compound were investigated theoretically using the DFT/B3LYP method and 6-311++G(d,p) basis set. The obtained results were compared with geometric structure and 1H-NMR data known in the literature and were seen to be compatible. Characters of intermolecular interactions were explained by Hirshfeld surface analysis and molecular electrostatic potential map analysis. Electronic properties of the N-(2,3,5,6-Tetra-fluoropyridin-4-yl) formamide compound were calculated considering frontier molecular orbital analysis. ADME study, conducted according to Lipinski's five criteria, revealed drug similarity properties of the N-(2,3,5,6-Tetrafluoropyridin-4-yl) formamide compound. In the last part of this study, the effect of nucleophilic substitution reactions on biochemical interactions between the perfluorinated compound and target protein was tested by molecular docking method using the Carbonic Anhydrase I (4WR7) enzyme.*

KEYWORDS: *Polyfluorinated pyridine; DFT; Hirshfeld surface analysis; Molecular docking; HOMO-LUMO; MEP.*

INTRODUCTION

N-(2,3,5,6-Tetrafluoropyridine-4-yl)formamide compound contains polysubstituted pyridine and formamide groups, which are commonly used in biochemical applications. Pyridines are cyclic structures formed by the replacement of a methyl group in the benzene ring with a nitrogen atom. Polysubstituted pyridines, by adding different groups to pyridines are important compounds used in many areas of chemistry.

Polyfluorinated pyridine groups are compounds used in pharmaceutical chemistry due to their unusual properties. It is also used in the synthesis of biologically active compounds because of the presence of strong electron-withdrawing fluorine atoms in the aromatic ring. The pyridine rings are assumed to interact with the nitrogen atom they traditionally have in the active sites of enzymes [1]. In this context, the polyfluoropyridines

* To whom correspondence should be addressed.

+ E-mail: alpaslan.bayrakdar@igdir.edu.tr

1021-9986/2023/8/2348-2450

13/\$/6.03

are an interesting field of study in areas such as supramolecular chemistry [2], dye chemistry [3], and electrochemical sensor synthesis [4], mainly due to aromatic nucleophilic displacement reactions of the fluorine atom [5]. Another functional group in the structure of the title compound is formamides. The Formamides are intermediates used in the synthesis of organic substances in many areas, such as herbicide chemicals [6], pesticides [7], and compounds with antifungal activity [8]. In addition, it is known that formamides are widely used in the synthesis of quinolones [9], which is a large synthetic antimicrobial group, and organic compounds [10] used in the treatment of bladder cancer.

In the literature, there are experimental studies in which halogen atoms are added as substituents on the benzene ring of amide groups to improve their inhibitory potential [11-13]. These studies showed that the biological activities of carbonic anhydrase inhibitors vary significantly with the amide groups into which different halogen atoms are incorporated. In other words, binding affinities of polyfluorinated derivatives have been reported to be greater than those of non-fluorinated ones [14]. In vitro and in silico studies are continuing so that newly synthesized compounds containing these bioactive molecules can be used in drug designs [14-17]. N-(2,3,5,6-Tetrafluoropyridin-4-yl)formamide compound containing perfluorinated pyridine and amide groups, which are biochemical groups, is the first molecule synthesized that contains a formamide functional group attached to a perfluorinated pyridine ring in its crystal structure [18]. Therefore, in this study, it was aimed to theoretically investigate the electronic and pharmacological properties of the title compound, which is thought to be biochemically active.

The perfluorinated heteroaromatic compounds are reactive to the nucleophilic attack because of their electron-deficient structure, and in principle, all fluorine atoms can be displaced by nucleophiles [19]. The nucleophilic substitution reaction is a class of organic reactions in which one nucleophile replaces another [20, 21]. On the other hand, it is known that nucleophilic substitution reactions originating from fluorine atoms in the structures of chemical compounds play a unique and important role in the addition of various groups to polyfluoroaromatic compounds. On the other hand, it is known that nucleophilic substitution reactions originating from fluorine atoms in the structures of chemical compounds

play a unique and important role in the addition of various groups to polyfluoroaromatic compounds [22]. In nucleophilic substitution reactions, the reactivity or strength of the nucleophile is called its nucleophilicity. Therefore, in the nucleophilic substitution reaction, a stronger nucleophile replaces a weaker nucleophile in the compounds.

The DFT method is a very popular method for the calculation of molecular structure and electronic properties of compounds recently [23-26]. In this study, the contributions of the regions in the N-(2,3,5,6-Tetrafluoropyridin-4-yl)formamide compound to the interactions resulting from the nucleophilic substitution reaction were investigated theoretically by Frontier Molecular Orbital (FMO) and Molecular electrostatic potential (MEP) analysis methods. Also, since intermolecular interactions are important parameters in crystal formation, the relative contributions of these interactions are investigated by Hirshfeld Surface (HS) analysis [27]. The analyzes made for the title compound gave very instructive results to explain the behaviour of the parts in the molecular structure.

In the last step of the study, ADME studies were conducted to elucidate the pharmacological properties of the title compound and drug similarity properties were revealed. In addition, potential targets of the N-(2,3,5,6-Tetrafluoropyridin-4-yl)formamide for molecular docking study were investigated by the virtual scanning method and their interactions with the target protein group were investigated by molecular docking study. Molecular docking studies are an effective in silico tool for interpreting interactions between receptors and ligands [28]. Recently, molecular docking studies have proven to be an important computer-based method for the screening, discovery, and prediction of receptor-ligand binding of bioactive compounds [29]. In this study, evidence for the contribution of nucleophilic substitution reactions of perfluorinated compounds to binding in molecular docking studies was investigated theoretically.

EXPERIMENTAL AND THEORETICAL SECTION

Experimental method

The N-(2,3,5,6-Tetrafluoropyridine-4-yl)formamide compound was synthesized and crystallized by *Newell et al.* [18]. Crystal data, data collection, and structure improvement details are summarized in the related article.

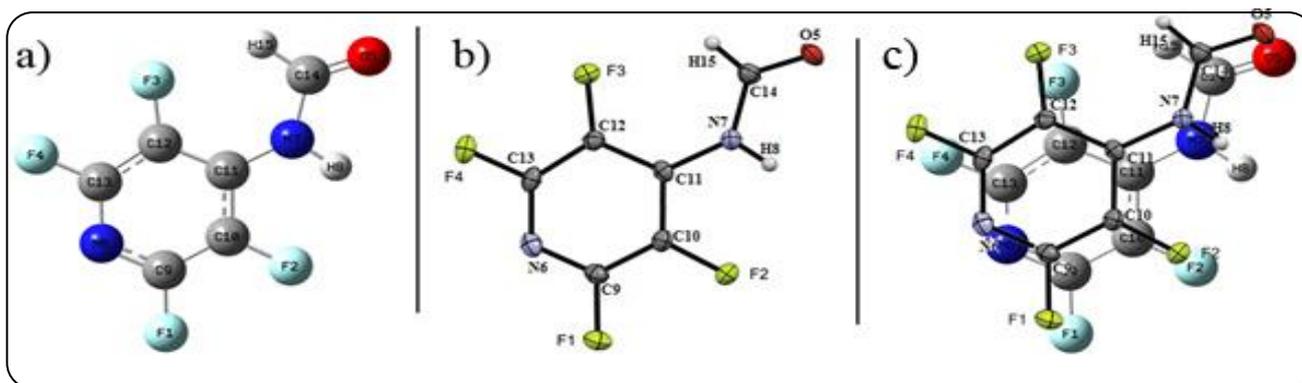


Fig. 1: a) Theoretically calculated optimized molecular structure. b) The most stable structural geometry obtained experimentally. c) The Comparison of the experimental and the optimized structure of the title compound

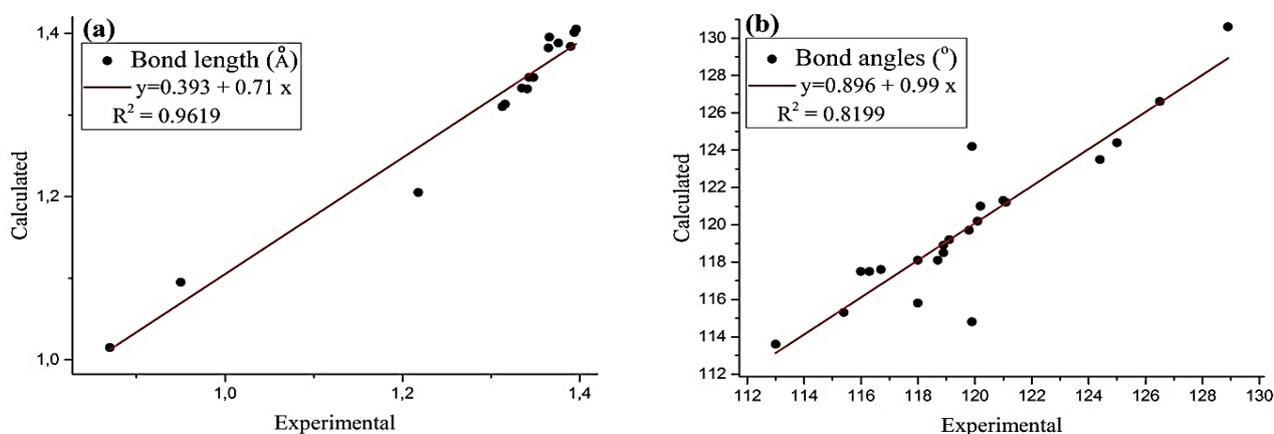


Fig. 2: The correlation graphs between theoretical and experimental optimized structure parameters of the title compound. a) correlations between bond lengths, b) correlations between bond angles

Computational method

In this study, theoretical calculations such as the determination of the most stable geometry of the title compound, the H-NMR study, the frontier molecular orbital analysis, and the molecular electrostatic potential map were calculated theoretically with the help of the Gaussian09 package program, which uses quantum chemical calculation methods [30]. Theoretical calculations were conducted at the B3LYP level of the Density Functional Theory (DFT) using the 6-311++G(d,p) basis set containing diffused and polarized functions. Hirshfeld surface analysis of the compound mentioned in the title was carried out using the Crystal Explorer 3.0 software [31]. The drug similarity properties for the title compound were examined via the SwissADME open-access web tool [32].

In addition, target proteins with which the title

compound can interact were predicted by virtual scanning with the SwissTargetPrediction tool, an open-access web tool. Before the molecular docking study, AutoDock Tools-1.5.6 [33] was used in the preparation of target proteins. Molecular docking studies were carried out using PyRx version 0.8. The BIOVIA Discovery Studio software [34] was used to analyze the results after the molecular docking procedure. Molecular docking studies were carried out using PyRx version 0.8.

RESULTS AND DISCUSSION

Molecular structure

The most stable molecular structure of the N-(2,3,5,6-Tetrafluoropyridin-4-yl)formamide compound, which is the first step of the theoretical calculations, was calculated with the help of the B3LYP method and the 6-311++G(d,p) basis set. In the optimization process, the

Table 1: Optimized geometrical parameters of title compound: bond length (Å) ° and bond angles (°) and bond torsions (°) in comparison with experimental data

Bond Lengths (Å)			Bond Angles °			Dihedral Angles °		
Atoms	Cal.	Exp.*	Atoms	Cal.	Exp.*	Atoms	Cal.	Exp.*
F1-C9	1.332	1.341 (3)	C9-N6-C13	117.5	116.0 (2)	F1-C9-C10-F2	0	-0.6 (3)
F2-C10	1.346	1.348 (3)	H8-N7-C11	115.8	118 (2)	N6-C9-C10-C11	0	-0.7 (4)
F3-C12	1.346	1.343 (3)	H8-N7-C14	113.6	113 (2)	C13-N6-C9-C10	0	-0.8 (4)
F4-C13	1.333	1.335 (3)	C11-N7-C14	130.6	128.9 (2)	N7-C11-C12-F3	0	-1.1 (4)
O5-C14	1.205	1.218 (3)	F1-C9-N6	117.6	116.7 (2)	F2-C10-C11-N7	0	-1.3 (3)
N6-C9	1.313	1.316 (3)	F1-C9-C10	118.9	118.9 (2)	C14-N7-C11-C12	0	-15.2 (4)
N6-C13	1.31	1.313 (3)	N6-C9-C10	123.5	124.4 (2)	F3-C12-C13-N6	180	-178.2 (2)
N7-H8	1.015	0.87 (3)	F2-C10-C9	121.3	121.0 (2)	C11-C12-C13-F4	-180	-178.9 (2)
N7-C11	1.384	1.390 (3)	F2-C10-C11	118.5	118.9 (2)	N6-C9-C10-F2	-180	-179.1 (2)
N7-C14	1.395	1.366 (3)	C9-C10-C11	120.2	120.1 (2)	F2-C10-C11-C12	180	-179.3 (2)
C9-C10	1.382	1.365 (3)	N7-C11-C10	118.1	118.0 (2)	C13-N6-C9-F1	-180	-179.4 (2)
C10-C11	1.405	1.396 (3)	N7-C11-C12	126.6	126.5 (2)	C9-N6-C13-F4	180	-179.5 (2)
C11-C12	1.401	1.394 (3)	C10-C11-C12	115.3	115.4 (2)	C9-C10-C11-N7	180	-179.8 (2)
C12-C13	1.388	1.376 (4)	F3-C12-C11	121.2	121.1 (2)	C10-C11-C12-C13	0	-2.3 (3)
C14-H15	1.095	0.95	F3-C12-C13	119.7	119.8 (2)	C9-N6-C13-C12	0	0.7 (4)
			C11-C12-C13	119.2	119.1 (2)	C11-C12-C13-N6	0	0.9 (4)
			F4-C13-N6	117.5	116.3 (2)	C14-N7-C11-C10	180	167.0 (2)
			F4-C13-C12	118.1	118.7 (2)	C10-C11-C12-F3	180	176.8 (2)
			N6-C13-C12	124.4	125.0 (2)	F1-C9-C10-C11	180	177.9 (2)
			O5-C14-N7	121	120.2 (2)	C11-N7-C14-O5	-180	179.0 (2)
			O5-C14-H15	124.2	119.9	N7-C11-C12-C13	-180	179.8 (2)
			N7-C14-H15	114.8	119.9	F3-C12-C13-F4	0	2.0 (3)
						C9-C10-C11-C12	0	2.2 (3)

* Experimental data are taken from ref [18].

the initial geometry was created with the help of the Gaussview molecular imaging program based on the geometry obtained from the X-ray diffraction analysis and optimized by the DFT method with no restrictions. The calculated optimized structure of the molecule mentioned in the title is presented in Fig. 1a, the experimental structure found by X-ray results is presented in Fig. 1b, and the overlap state of the experimentally and theoretically found molecular structures is presented in Fig. 1c. Experimental and theoretical parameters, such as bond lengths, bond angles, and dihedral angles of the optimized structure, are presented in Table 1. The optimized structure of the title compound is nearly planar. Deviation values of F1, F2, F3, F4, and N6 atoms attached to the cyclic structure were found as 0.0557, -0.0529, -0.1030, 0.0383, and 0.0459 Å, respectively, with the help of SHELXL

software [35].

The correlation values between the experimental and theoretical optimization parameters of the title compound were found to be 0.099 for bond lengths and 0.81 for bond angles are given in Fig. 2a and Fig. 2b, respectively.

A potential energy surface analysis was performed to show the orientation of the polysubstituted pyridine and formamide groups in the structure of the title compound and to determine the conformational minimum and maximum energy of the molecule. Potential energy surface scanning around C14-N7-C11-C12 was theoretically performed with DFT/B3LYP level and 6-311++G(d,p) basis set. During the scanning process, the dihedral angle at C14-N7-C11-C12 was rotated in 36 steps from 0° to 360° in increments of 10°.

Fig. 3 provides information on the molecular energy

Table 2: Absorption wavelength (λ), excitation energies (E), and oscillatory strength (f) of the title compound and the major contributions for transitions derived from calculation in different media using DFT/B3LYP/6-311++G(d,p) method

N-(2,3,5,6-Tetrafluoropyridin-4-yl) formamid			
Parameters		Parameters	
E_{HOMO} (eV)	-8.04	χ (eV)	5.17
E_{LUMO} (eV)	-2.31	η (eV)	2.86
ΔE (eV)	5.73	S (eV^{-1})	0.18
I (eV)	8.04	ω (eV^{-1})	38.30
A (eV)	2.31	μ (eV^{-1})	-5.17

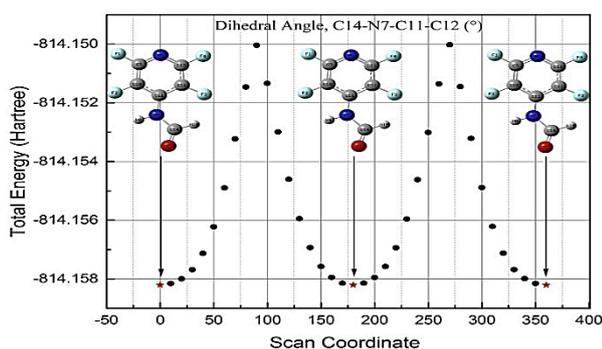


Fig. 3: Conformational analysis of the title compound by PES scan

distribution of rotation according to the selected torsion angle. As can be clearly seen from Fig. 3, the most stable conformations corresponding to the global minimums were calculated as 0° (-814.1582 Hartree), -180° (-814.1582 Hartree) and 360° (-814.1582 Hartree). On the other hand, the global maximums of the molecular structure were calculated as 90° (-814.150 Hartree) and 270° (-814.150 Hartree). It was observed that these values were consistent with the X-ray results given in Table 1.

Frontier molecular orbital analysis

Molecular orbital analysis plays a very important role in understanding the reactivity of chemical compounds. Especially the HOMO [Highest Occupied Molecular Orbital] and LUMO [Least Occupied Molecular Orbital] orbitals, also called frontier molecular orbitals, are very useful in understanding the characteristic and electronic properties of the compound. In molecules, the regions where the HOMO orbitals are localized act as an electron donor, while the regions where the LUMO orbitals are localized act as an electron acceptor. In other words, the HOMO orbital reflects the nucleophilic ability

of the molecule, and the LUMO orbital reflects the electrophilic ability of the molecule.

The small energy gap between the HOMO-LUMO orbitals means that the molecule has good polarizability and high chemical reactivity. When two reactants compound approach each other to react, the HOMO of one molecule symmetrically excites the LUMO of the other. As the reaction progresses, there will be a transfer of charge from the HOMO of one molecule to the LUMO of the other. Parr *et al.* proposed the electrophilicity index as a quantitative measure of energy reduction due to the maximum electron flow between the donor and acceptor orbitals [33]. The chemical reactivity and electronic properties of the title compound were investigated by frontier molecular orbital analysis. For the calculation, the optimized structure obtained using the DFT/B3LYP level and the 6-311++G(d,p) basis set was used. As a result of the theoretical calculation, E_{HOMO} and E_{LUMO} and the descriptive properties of the title compound calculated accordingly, such as chemical hardness (η), chemical potential (μ), chemical softness (S), electronegativity (χ), and electrophilic index (ω) are listed in Table 2.

The HOMO-LUMO orbital shapes of the title compound are mapped in three dimensions (3D) and where the molecular orbitals are localized is shown in Fig. 4a. The HOMO energy (E_{HOMO}) was calculated as -8.04 eV, the LUMO energy (E_{LUMO}) -2.31 eV, and the energy gap (ΔE) 5.73 eV. The HOMO-LUMO energy gap plays an important role in predicting the charge transfer within the molecule, chemical reactivity, biological activity, and stability of the compound [36-39]. The smaller the gap value is, the more chemically active the molecule is, and it is termed a soft molecule. The biological reactivity of soft molecules is higher than that of hard molecules [40]. As can be clearly seen from Table 2, the HOMO-LUMO energy range is less than 10 eV. This means that the title compound is a soft molecule [41]. Also, the energy gap is used to prove bioactivity from Intramolecular Charge Transfer (ICT) as it is a measure of electron conductivity [42]. As can be clearly seen from Fig. 4a, the HOMO and LUMO orbitals are localized almost throughout the molecule. Total Electronic State Density (TDOS) and Partial Density of State (PDOS) graph diagrams of the title compound calculated with the help of GaussSum 3.0 software [43] are given in Fig. 4b. PDOS essentially presents the composition of the fragment orbitals that

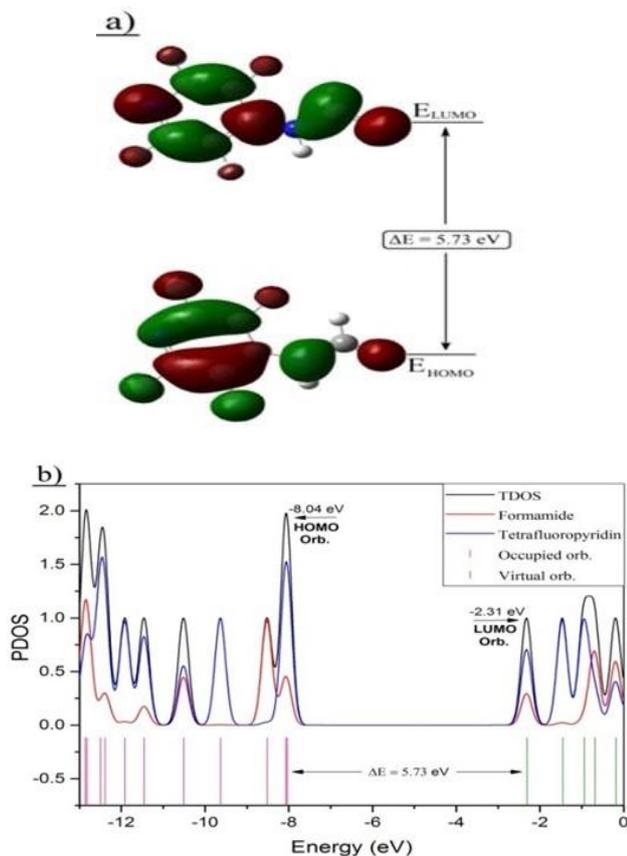


Fig. 4: a) Frontier molecular orbitals (HOMO and LUMO) for *N*-(2,3,5,6-Tetrafluoropyridin-4-yl)formamide. b) The calculated partial electronic density of the states plot for the title molecule

contribute to the molecular orbitals. As can be seen from the PDOS plot in Fig. 4b, 87% of the contributions to the HOMO orbitals came from the ring region, while 13% came from formamide. On the other hand, while 71% of the contributions to the LUMO orbitals came from the ring region, 29% came from formamide.

Mep's analysis

Molecular Electrostatic Potential (MEP) is a significant tool to predict electrophilic and nucleophilic attacks in biological interactions. Also, Molecular electrostatic potential maps are considered reliable hydrogen bond descriptors. The increasing order of electrostatic potential is defined by the color orange from red to blue. The green color indicates the region where the potential is neutral. In this work, calculations over the optimized structure were performed to interpret and predict regions of relative reactivity for nucleophilic and electrophilic attack by visualizing the MEP surfaces

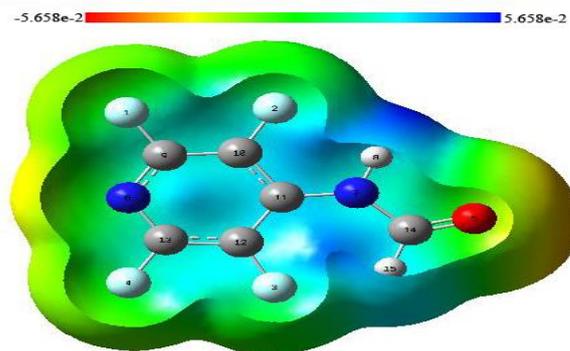


Fig. 5: MEP map of *N*-(2,3,5,6-Tetrafluoropyridin-4-yl)formamide

of the title compound in the gas phase. The MEP diagram of the title compound three-dimensionally is shown in Fig. 5. Electron density values on the MEP range from -5.658×10^{-2} to $+5.658 \times 10^{-2}$ from the negative region to the positive region of the compound. Negative regions (red) of the MEP map are related to electrophilic attack reactivity, while positive (blue) regions are related to nucleophilic attack reactivity. As can be clearly seen in the MEP pattern of the title compound presented in Fig. 5, the potential site for nucleophilic attack is on the hydrogen atom (H8) attached to the nitrogen atom in the formamide group, which has a positive potential (blue region). The most negative regions (red-yellow regions) that determine the nucleophilic region of the title compound are located on the carbonyl oxygen atom (O5) in the formamide group and the nitrogen atom (N6) in the polyfluoropyridine ring. Therefore, fluorine atoms on the electron-poor aromatic ring prepare the ring for nucleophilic attack due to its electron-withdrawing properties. The calculated MEP pattern revealed the negative and positive portions of the title compound. These reactive areas provide evidence of the position to be considered for molecular docking of the studied molecule.

Hirshfeld surface analysis

Hirshfeld Surface Analysis (HSA) is an analysis method that visualizes intermolecular short and long interactions with a color scale on the molecule [44]. The color codes on the surface in red, white, and blue represent contact zones that are shorter, equal, and longer than the sum of the Van der Waals radii, respectively. In addition, two-dimensional dnorm fingerprint patterns prepared by using atomic coordinates with the help of HSA allow the determination of contribution rates from atoms to intermolecular interactions. de and di axes appearing

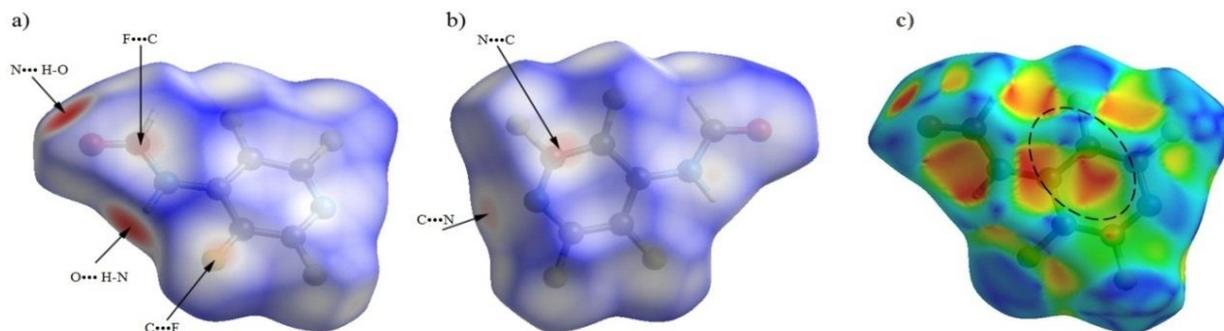


Fig. 6: Front (a), rear (b) views of Hirshfeld surface mapped on $dnorm$ and (c) Shape index

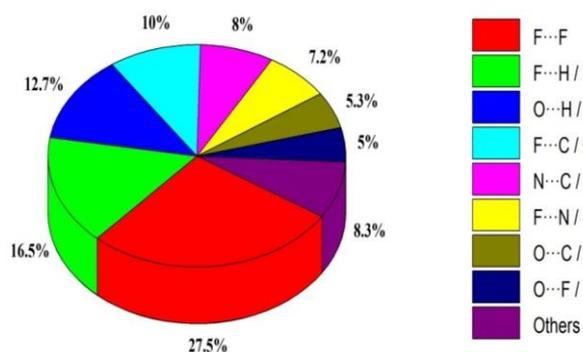


Fig. 7: Percentage contributions of interatomic contacts to the Hirshfeld surface

in the $dnorm$ surface plot represent the outer and inner distances, respectively, of a point on the HSA to the nearest atoms.

In this study, Hirshfeld surface analysis including $dnorm$ surfaces, shape index, and 2D Fingerprint Plots (FP) was performed by the Crystal Explorer 3.0 program in order to verify intermolecular interactions and provide quantitative data on the relative contributions to surface, and are given in Fig. 6a-c and 7, respectively. The dark red dots seen in the Hirshfeld surface map in Fig. 6a and b represent strong hydrogen bonds caused by acceptor and donor atoms, while other intermolecular interactions are represented by light red dots. The Hirshfeld surface formed using the shape index property was employed for the determination of the $\pi \cdots \pi$ stacking interactions in the crystal packing. Fig. 6c shows the shape index contour surface plot of the title compound generated in the range of -1 to 1 Å. The shape index is a dimensionless surface feature used to identify complementary pairs of red-hollows and blue bumps in maps where molecular surfaces converge [45, 46]. The blue triangle regions symbolize hydrogen donor groups and the red triangle

regions symbolize hydrogen acceptor groups. π - π interactions are usually indicated by adjacent red and blue triangles on the shape index map [47], as is the case with the cap molecule. These triangles form a kind of "bow-tie" on the map, indicating the presence of π - π interactions (marked in the ellipse, Fig. 6c).

As seen in Fig. 6a, dark red dots on the $dnorm$ surface of the title compound indicate that strong hydrogen bonds have occurred between $N \cdots H-O$ and $O \cdots H-N$ atoms. On the other hand, in Figs 6a and 6b, it is seen that the interactions between $F \cdots C$ and $N \cdots C$ cause light red dots on the $dnorm$ surface. Hydrogen bonds and other contacts made by atoms in the structure of the title compound are shown in the two-dimensional fingerprint patterns presented in Fig. 7a-i, along with their contribution to the Hirshfeld surface. $F \cdots F$ short contact made the greatest contribution to the Hirshfeld surface with 27.5% (Fig. 7b). The contributions from the $F \cdots F$ contact and other contacts to the Hirshfeld surface are given in Fig 8. The data presented in Figs. 7a-i and 8 provide qualitative evidence for the existence of various hydrogen bonds and intermolecular interactions in the crystal structure of the title compound. A pair of sharp spikes appearing in Fig. 7d indicate the presence of hydrogen bonds between $N7-H8 \cdots O5'$ and $O5 \cdots H8'-N7'$ atoms.

H-NMR

NMR spectroscopy is a useful technique for the structural analysis of organic compounds [48]. The 1H -NMR spectrum of the title compound was calculated by the B3LYP/6-31G++(d,p) method using TMS as a standard. The exhibition of two signals at δ 7.7 and 9 ppm in the 1H NMR ($CDCl_3$) spectrum of the title compound was assigned to the NH and OH protons, respectively. On the other hand, the theoretical 1H -NMR shift values were

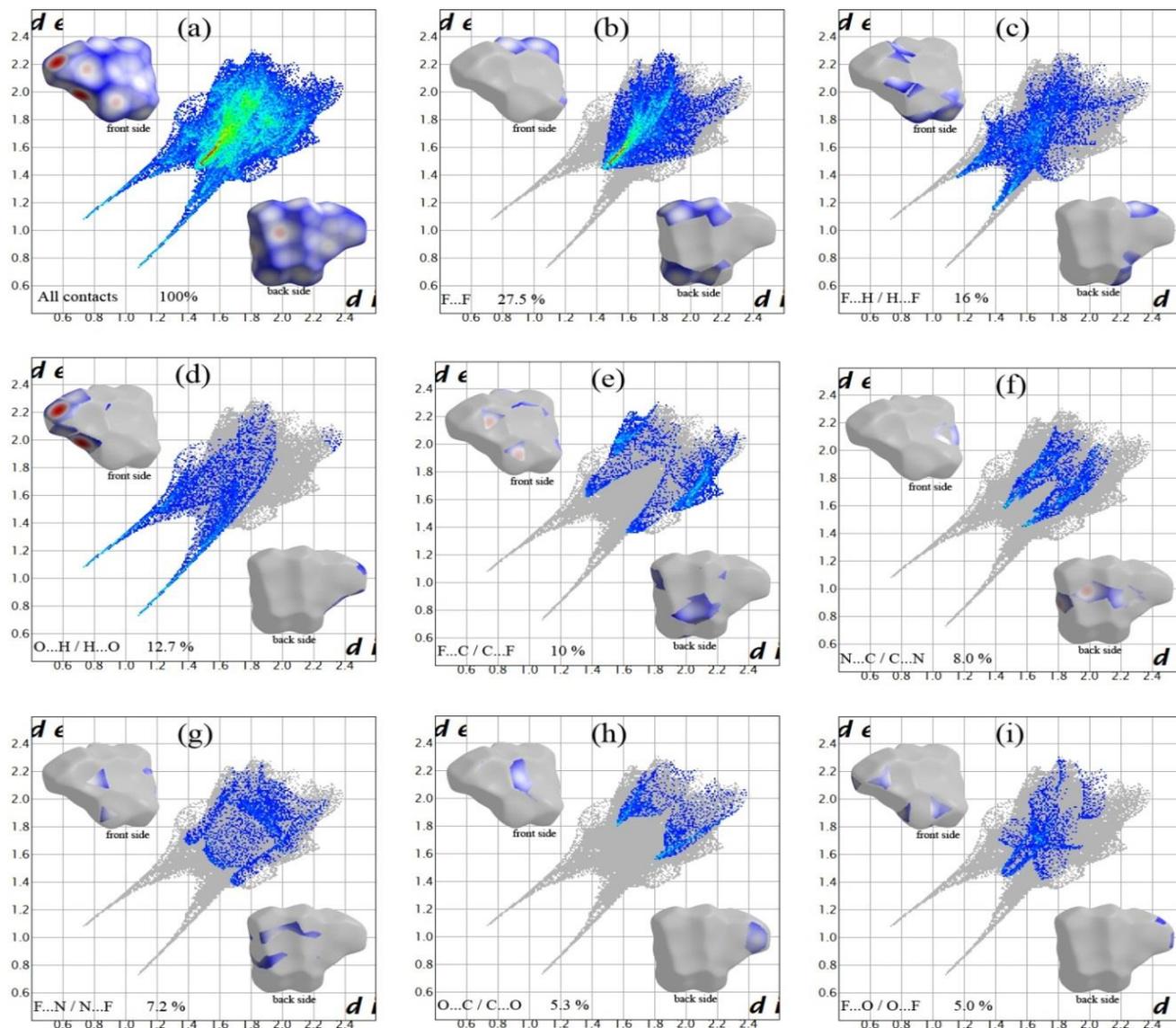


Fig. 8: Two dimensional fingerprints graphs for the title compound. The d_i and d_e along x and y represent the nearest adjacent nuclei internal and external from the three dimensional Hirshfeld surface. (a) Total contribution, (b) $F\cdots F$, (c) $F\cdots H/H\cdots F$, (d) $C\cdots H/H\cdots C$, (e) $O\cdots H/H\cdots O$, (f) $N\cdots C/C\cdots N$, (g) $F\cdots N/N\cdots F$, (h) $O\cdots C/C\cdots O$ and (i) $F\cdots O/O\cdots F$ inter-atomic contacts

calculated as 7.3 ppm for NH and 9.4 ppm for OH by the DFT method. The correlation between the calculated and experimentally obtained spectrum values was calculated as $R^2 = 0.987$. This result indicates a consistency between experimental and theoretical $^1\text{H-NMR}$ results.

Druglikeness and ADME studies

For a compound to be evaluated as a drug, its interaction with the target protein and its activity in living organisms should be examined. Therefore, an Absorption, Distribution, Metabolism, and Excretion (ADME) study

is performed to examine the ligands' interaction with living organisms. The physicochemical data of the title compounds for the ADME study were searched using the SwissADME online server. SwissADME is a free, user-friendly web tool that is widely used for the analysis of physicochemical properties, pharmacokinetics, and relevance to medicinal chemistry in the synthesis of existing drugs or new drug molecules.

There are different criteria such as Lipinski, Ghose, Veber, and Egan in the literature regarding whether or not a compound shows drug similarity. In this study, we will

Table 3: Lipinski's Rule of Five for ADME analysis of the title compounds

Lipinski's five criteria	Accepted range	N-(2,3,5,6-Tetrafluoropyridin-4-yl)formamid	
		Calculated	Decision
Molecular mass (Da) (MW)	≤500	194.09	✓
Hydrogen bond donor (HBD)	≤5	1	✓
Hydrogen bond acceptor (HBA)	≤10	6	✓
LogP	≤5	1.16	✓

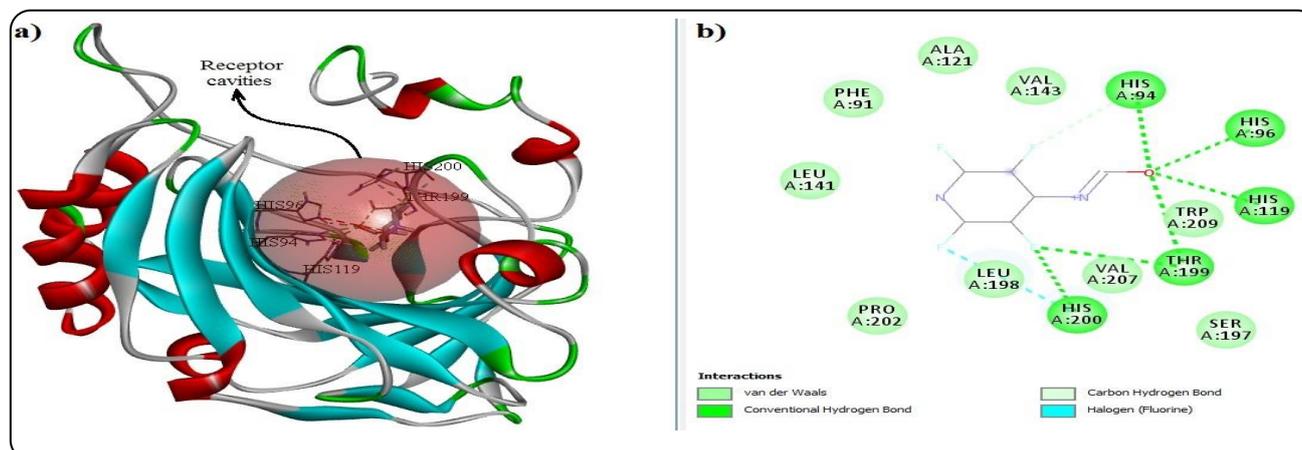


Fig. 9: a) Molecular docking interactions between N-(2,3,5,6-Tetrafluoropyridin-4-yl)formamide and the binding sites of 4WR7. b) Residual interactions between N-(2,3,5,6-Tetrafluoropyridin-4-yl)formamide and 4WR7

evaluate drug similarity characteristics according to Lipinski's criteria, which is one of the most frequently encountered criteria in the literature. Lipinski's five criteria were developed by Christopher Lipinski to predict the oral bioactivity of molecules. According to the Lipinski criteria, it can be defined as the chemical structure limitations on the molecular weights (≤ 500), hydrogen-bond acceptor (≤ 10) and hydrogen-bond donor (≤ 5) numbers, and lipophilicity ($\log P$ or $\text{clog } P$) (≤ 5) of compounds [49]. The drug similarity properties calculated for the compound are given in Table 3. As can be clearly seen in Table 3, there is no violation of the Lipinski criteria. This means that there is nothing to prevent the title compound from being an orally active drug in humans.

Molecular docking study

The purpose of a molecular docking study is to determine the optimal binding positions between Protein and ligand by examining the interactions between proteins, which are large molecules, and ligands, which are small

molecules. Ligands generally interact with amino acid residues located in the active cavity of target proteins. For this reason, it is very important to identify the amino acids in the active site of the target proteins and to define the region where these amino acids are located as the study area. Since there is no molecular docking study in the literature on N-(2,3,5,6-Tetrafluoropyridin-4-yl)formamide and its derivatives, potential target proteins were determined by virtual scanning through the SwissTarget Prediction tool, a web tool with an open-access, user-friendly interface. In the virtual scan, a protein with PDB code 4WR7 from the Carbonic Anhydrase I (CAI) family was selected as the target protein for the title molecule. The carbonic Anhydrase has three conserved histidine residues (His94, His96, and His119) and a single active site with a zinc atom coordinated by a water molecule [50].

The optimized structure of the title compound was used as a ligand in the molecular docking study with PyRx software. On the other hand, the water molecules, hetatm, and ligand groups in the structure of the 4WR7 target protein were cleaned by AutoDockTools-1.5.6 before

Table 4: Protein-ligand interactions of the title compound with 4WR7

Protein	PDB code	Ligand	ΔG (kcal.mol ⁻¹)	Conventional H-Bonds (Å)	Carbonl H-Bonds (Å)	Halogen (Fluorine) (Å)
Carbonic Anhydrase I	4WR7	N-(2,3,5,6-Tetrafluoropyridin-4-yl)formamide	-5.8	His94(2.30) His96(2.31) His119(2.24) Thr199(2.91-2.94) His200 (2.59)	His94(3.18) His200 (3.30)	His200(3.70)

the molecular docking study. In the next step, polar hydrogen atoms and Kollman charges were added to the target protein. Finally, the pre-docking process was completed by entering the grid parameters to include the active cavity of the target protein.

Protein-ligand interactions obtained by molecular docking studies performed with the help of Pyrex software are given in Table 5 and Fig. 8. It is clearly seen in both Table 4 and Fig. 9 that the title compound performs a conventional hydrogen bond interaction with His94, His96, His119, Thr199, and His200 in the active cavity of the 4WR7 target protein. Two of the conventional hydrogen bonds that take place here are between the active residues Thr199 and His200 and fluorine atoms. In addition, the other hydrogen bond made with Thr99 was realized with the O atom. On the other hand, the active residue His200 formed a halogen bond with the fluorine atom.

The presence of the halogen bond is clear evidence of a net attractive interaction between an electrophilic region associated with a halogen atom in the structure and a nucleophilic region in another molecular structure.

CONCLUSIONS

In this study, molecular geometry was successfully calculated and compared with experimental results. The optimized structure was used in other theoretical calculations. From the HOMO-LUMO analysis, the title compound was found to be a chemically soft and bioactive molecule. The predicted MEP diagram presented information about the reactive sites in the molecular docking study. The non-covalent interactions of the titular compound are detailed using Hirshfeld surface analysis; from which it is concluded that H-F/F-H and its contacts are a notable donor for crystal packaging. The drug-likeness study demonstrated that the title compound had a good pharmacokinetic profile. In addition, in silico molecular docking results showed a good association between N-(2,3,5,6-Tetrafluoropyridin-4-yl)formamide

ligand and selected 4WR7 protein, with ligand docking into the active site of the protein. The presence of a halogen bond between these interactions was also observed. In this study, it was emphasized that the title compound is promising in drug designs because compounds with halogen bonds increase the selectivity and efficiency at the active sites of target proteins. This result will make perfluorinated structures interesting in the synthesis of biologically active compounds and in pharmacological and medical research.

Received : Oct. 26, 2022 ; Accepted : Jan. 21, 2023

REFERENCES

- [1] Griswold W.R., Toney M.D., [Role of the Pyridine Nitrogen in Pyridoxal 5'-Phosphate Catalysis: Activity of Three Classes of PLP Enzymes Reconstituted with Deazapyridoxal 5'-Phosphate](#), *J. Am. Chem. Soc.*, **133(37)**: 14823-14830 (2011).
- [2] Han H., Kovtonyuk V.N., Gatilov Y.V., Andreev R.V., [Directed Synthesis of Isomeric Polyfluorinated Dinitrotetraoxacalixarenes and Bicyclooxacalixarenes](#), *J. Fluorine Chem.*, **261**: 110022 (2022).
- [3] Ishchenko R.A., Kargapolova I.Y., Orlova N.A., Shelkownikov V.V., Maksimov A.M., Ryazanov N.D., Berezhnaya V.N., Chemonosov A.A., [Polyfluorinated Triphenyl-4, 5-Dihydro-1H-Pyrazoles with Dendroid Arylsulfanyl Moieties as Donor Blocks in Donor-Acceptor Chromophores](#), *J. Fluorine Chem.*, **248**: 109841 (2021).
- [4] Cao T., Zhou Y., Wang H., Qiao R., Zhang X., Liu L., Tong Z., [Preparation of Polyfluorinated Azobenzene/Niobate Composite as Electrochemical Sensor for Detection of Ascorbic Acid and Dopamine](#), *Microchem. J.*, **179**: 107422 (2022).

- [5] Ranjbar-Karimi R., Poorfreidoni A., Masoodi H.R., Survey Reactivity of Some N-aryl Formamides with Pentafluoropyridine, *J. Fluorine Chem.*, **180**: 222-226 (2015).
- [6] Shi J., Lou Z., Yang M., Zhang Y., Liu H., Meng Y., Theoretical Characterization of Formamide on the Inner Surface of Montmorillonite, *Surf. Sci.*, **624**: 37-43 (2014).
- [7] Bommuraj V., Chen Y., Birenboim M., Barel S., Shimshoni J.A., Concentration-and Time-Dependent Toxicity of Commonly Encountered Pesticides and Pesticide Mixtures to Honeybees (*Apis Mellifera* L.), *Chemosphere*, **266**: 128974 (2021).
- [8] CARTER B.G., Chamberlain K., Wain R., Investigations on Fungicides: XV. The Fungitoxicity and Systemic Antifungal Activity of N-(2, 2, 2-Trichloro-1-Methoxyethyl) Formamide and Related Compounds, *Ann. Appl. Biol.*, **75**(1): 49-55 (1973).
- [9] Alla K., Vijayakumar V., Sarveswari S., Synthesis and in Vitro Antimicrobial Evaluation of New Quinolone Based 2-Arylamino Pyrimidines, *Polycyclic Aromat. Compd.*: 1-22 (2022).
- [10] Ibrahiem E.H.I., Nigam V.N., Brailovsky C.A., Madarnas P., Elhilali M., Orthotopic Implantation of Primary N-[4-(5-Nitro-2-Furyl)-2-Thiazolyl] Formamide-Induced Bladder Cancer in Bladder Submucosa: An Animal Model for Bladder Cancer study, *Cancer Res.*, **43**(2): 617-622 (1983).
- [11] Iliés M.A., Vullo D., Pastorek J., Scozzafava A., Iliés M., Caproiu M.T., Pastorekova S., Supuran C.T., Carbonic Anhydrase Inhibitors. inhibition of Tumor-Associated Isozyme IX by Halogenosulfanilamide and Halogenophenylaminobenzolamide Derivatives, *J. Med. Chem.*, **46**(11): 2187-2196 (2003).
- [12] Winum J.-Y., Dogné J.-M., Casini A., de Leval X., Montero J.-L., Scozzafava A., Vullo D., Innocenti A., Supuran C.T., Carbonic Anhydrase Inhibitors: Synthesis and Inhibition of Cytosolic/Membrane-Associated Carbonic Anhydrase Isozymes I, II, and IX with Sulfonamides Incorporating Hydrazino Moieties, *J. Med. Chem.*, **48**(6): 2121-2125 (2005).
- [13] Krishnamurthy V.M., Bohall B.R., Kim C.Y., Moustakas D.T., Christianson D.W., Whitesides G.M., Thermodynamic Parameters for the Association of Fluorinated Benzenesulfonamides with Bovine Carbonic Anhydrase II, *Chemistry—An Asian Journal*, **2**(1): 94-105 (2007).
- [14] Samanta P.N., Das K.K., Prediction of Binding Modes and Affinities of 4-Substituted-2, 3, 5, 6-Tetrafluorobenzenesulfonamide Inhibitors to the Carbonic Anhydrase Receptor by Docking and ONIOM Calculations, *J. Mol. Graphics Modell.*, **63**: 38-48 (2016).
- [15] Shchegol'kov E., Shchur I., Burgart Y.V., Saloutin V., Trefilova A., Ljushina G., Solodnikov S.Y., Markova L., Maslova V., Krasnykh O., Polyfluorinated Salicylic Acid Derivatives as Analogs of Known Drugs: Synthesis, Molecular Docking and Biological Evaluation, *Bioorg. Med. Chem.*, **25**(1): 91-99 (2017).
- [16] Harris J.D., Coon C.M., Doherty M.E., McHugh E.A., Warner M.C., Walters C.L., Orahod O.M., Loesch A.E., Hatfield D.C., Sitko J.C., Engineering and Characterization of Dehalogenase Enzymes from *Delftia Acidovorans* in Bioremediation of Perfluorinated Compounds, *Synth. Syst. Biotechnol.*, **7**(2): 671-676 (2022).
- [17] Etsè K.S., Etsè K.D., Zaragoza G., Mouithys-Mickalad A., Structural Description, IR, TGA, Antiradical, HRP Activity Inhibition and Molecular Docking Exploration of N-Cyclohexyl-N-Tosylformamide, *J. Mol. Struct.*, **1269**: 133731 (2022).
- [18] Newell B.D., McMillen C.D., Lee J.P., N-(2, 3, 5, 6-Tetrafluoropyridin-4-yl) Formamide, *IUCrData*, **7**(8): x220804 (2022).
- [19] Cartwright M.W., Parks E.L., Pattison G., Slater R., Sandford G., Wilson I., Yufit D.S., Howard J.A., Christopher J.A., Miller D.D., Annulation of Perfluorinated Heteroaromatic Systems by 1, 3-Dicarbonyl Derivatives, *Tetrahedron*, **66**(17): 3222-3227 (2010).
- [20] Panova M.A., Shcherbakov K.V., Burgart Y.V., Saloutin V.I., Selective Nucleophilic Aromatic Substitution of 2-(Polyfluorophenyl)-4H-Chromen-4-Ones with Pyrazole, *J. Fluorine Chem.*, **263**: 110034 (2022).
- [21] Shabalin A.Y., Adonin N.Y., Bardin V.V., Taran O.P., Ayushev A.B., Parmon V.N., Synthesis of Potassium 4-(1-Azol-1-yl)-2, 3, 5, 6-Tetrafluorophenyltrifluoroborates from K [C6F5BF3] and Alkali Metal Azol-1-Ides. The Dramatic Distinction in Nucleophilicity of Alkali Metal Azol-1-Ides and Dialkylamides, *J. Fluorine Chem.*, **156**: 290-297 (2013).

- [22] Morgan P.J., Saunders G.C., Macgregor S.A., Marr A.C., Licence P., [Nucleophilic Fluorination Catalyzed by a Cyclometallated Rhodium Complex](#), *Organometallics*, **41**(7): 883-891 (2022).
- [23] Barim E., Akman F., [Synthesis, Characterization and Spectroscopic Investigation of N-\(2-Acetylbenzofuran-3-yl\) Acrylamide Monomer: Molecular Structure, HOMO–LUMO Study, TD-DFT and MEP Analysis](#), *J. Mol. Struct.*, **1195**: 506-513 (2019).
- [24] Akman F., Kaçal M.R., Akman F., Soylu M.S., [Determination of Effective Atomic Numbers and Electron Densities from Mass Attenuation Coefficients for Some Selected Complexes Containing Lanthanides](#), *Can. J. Phys.*, **95**(10): 1005-1011 (2017).
- [25] SenthilKannan K., Sivaramakrishnan V., Kalaipoonguzhali V., Chinnadurai M., Kannan S., [Electronic Transport, HOMO–LUMO and Computational Studies of CuS Monowire for Nano Device Fabrication by DFT Approach](#), *Mater. Today: Proc.*, **33**: 2746-2749 (2020).
- [26] Mumit M.A., Pal T.K., Alam M.A., Islam M.A.-A.-A.-A., Paul S., Sheikh M.C., [DFT Studies on Vibrational and Electronic Spectra, HOMO–LUMO, MEP, HOMA, NBO and Molecular Docking Analysis of Benzyl-3-N-\(2, 4, 5-Trimethoxyphenylmethylene\) Hydrazinecarbodithioate](#), *J. Mol. Struct.*, **1220**: 128715 (2020).
- [27] Khan B.A., Ashfaq M., Muhammad S., Munawar K.S., Tahir M.N., Al-Sehemi A.G., Alarfaji S.S., [Exploring Highly Functionalized Tetrahydropyridine as a Dual Inhibitor of Monoamine Oxidase A and B: Synthesis, Structural Analysis, Single Crystal XRD, Supramolecular Assembly Exploration by Hirshfeld Surface Analysis, and Computational Studies](#), *ACS Omega*, **7**(33): 29452-29464 (2022).
- [28] Yu Z., Kang L., Zhao W., Wu S., Ding L., Zheng F., Liu J., Li J., [Identification of Novel Umami Peptides from Myosin via Homology Modeling and Molecular Docking](#), *Food chemistry*, **344**: 128728 (2021).
- [29] Yu Z., Ji H., Shen J., Kan R., Zhao W., Li J., Ding L., Liu J., [Identification and Molecular Docking Study of fish Roe-Derived Peptides as Potent BACE 1, AChE, and BChE Inhibitors](#), *Food Funct.*, **11**(7): 6643-6651 (2020).
- [30] Frisch M.J., Trucks G.W., Schlegel H.B., Scuseria G.E., Robb M.A., Cheeseman J.R., Scalmani G., Barone V., Mennucci B., Petersson G.A., Nakatsuji H., Caricato M., Li X., Hratchian H.P., Izmaylov A.F., Bloino J., Zheng G., Sonnenberg J.L., Hada M., Ehara M., Toyota K., Fukuda R., Hasegawa J., Ishida M., Nakajima T., Honda Y., Kitao O., Nakai H., Vreven T., Montgomery J.A., Peralta Jr. J.E., Ogliaro F., Bearpark M., Heyd J.J., Brothers E., Kudin K.N., Staroverov V.N., Kobayashi R., Normand J., Raghavachari K., Rendell A., Burant J.C., Iyengar S.S., Tomasi J., Cossi M., Rega N., Millam J.M., Klene M., Knox J.E., Cross J.B., Bakken V., Adamo C., Jaramillo J., Gomperts R., Stratmann R.E., Yazyev O., Austin A.J., Cammi R., Pomelli C., Ochterski J.W., Martin R.L., Morokuma K., Zakrzewski V.G., Voth G.A., Salvador P., Dannenberg J.J., Dapprich S., Daniels A.D., Farkas O., Foresman J.B., Ortiz J.V., Cioslowski J., Fox D.J., [Gaussian 09, Inc, Wallingford CT](#), **121**: 150-166 (2009).
- [31] Wolff S., Grimwood D., McKinnon J., Turner M., Jayatilaka D., Spackman M., [“Crystal Explorer”](#), University of Western Australia Crawley, Australia, 2021.
- [32] 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.*, **7**(1): 1-13 (2017).
- [33] Trott O., Olson A.J., [AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading](#), *J. Comput. Chem.*, **31**(2): 455-461 (2010).
- [34] Systèmes D., Free Download: BIOVIA Discovery Studio Visualizer-Dassault Systèmes, in, 2020.
- [35] Sheldrick G.M., [A Short History of SHELX](#), *Acta Crystallogr., Sect. A: Found. Crystallogr.*, **64**(1): 112-122 (2008).
- [36] Parr R.G., Szentpály L.v., Liu S., [Electrophilicity Index](#), *J. Am. Chem. Soc.*, **121**(9): 1922-1924 (1999).
- [37] Fukui K., [Role of Frontier Orbitals in Chemical Reactions](#), *Science*, **218**(4574): 747-754 (1982).
- [38] Samsonowicz M., Regulska E., Kowczyk-Sadowy M., Butarewicz A., Lewandowski W., [The Study on Molecular Structure and Microbiological Activity of Alkali Metal 3-Hydroxyphenylacetates](#), *J. Mol. Struct.*, **1146**: 755-765 (2017).

- [39] Samsonowicz M., Regulska E., Świsłocka R., Butarewicz A., [Molecular Structure and Microbiological Activity of Alkali Metal 3, 4-Dihydroxyphenylacetates](#), *J. Saudi Chem. Soc.*, **22(8)**: 896-907 (2018).
- [40] Merza M.M., Hussein F.M., AL-ani R.R., [Physical Properties and Biological Activity of Methyldopa Drug Carrier Cellulose Derivatives. Theoretical Study](#), *Egypt. J. Chem.*, **64(8)**: 4081-4090 (2021).
- [41] Yadav B., Yadav R.K., Srivastav G., Yadav R., [Experimental Raman, FTIR and UV-Vis Spectra, DFT Studies of Molecular Structures and Barrier Heights, Thermodynamic Functions and Bioactivity of Kaempferol](#), *J. Mol. Struct.*, **1258**: 132637 (2022).
- [42] Bendjeddou A., Abbaz T., Gouasmia A., Villemin D., [Molecular Structure, HOMO-LUMO, MEP and Fukui Function Analysis of Some TTF-Donor Substituted Molecules Using DFT \(B3LYP\) Calculations](#), *Int. Res. J. Pure Appl. Chem.*, **12(1)**: 1-9 (2016).
- [43] O'boyle N.M., Tenderholt A.L., Langner K.M., [Cclib: A Library for Package-Independent Computational Chemistry Algorithms](#), *J. Comput. Chem.*, **29(5)**: 839-845 (2008).
- [44] Spackman M.A., Jayatilaka D., [Hirshfeld Surface Analysis](#), *Cryst. Eng. Comm.*, **11(1)**: 19-32 (2009).
- [45] Madni M., Ahmed M.N., Abbasi G., Hameed S., Ibrahim M.A., Tahir M.N., Ashfaq M., Gil D.M., Gomila R.M., Frontera A., [Synthesis and X-ray Characterization of 4, 5-Dihydropyrazolyl-Thiazoles Bearing a Coumarin Moiety: On the Importance of Antiparallel \$\pi\$ -Stacking](#), *Chemistry Select*, **7(36)**: e202202287 (2022).
- [46] Ashfaq M., Ali A., Tahir M.N., Khalid M., Assiri M.A., Imran M., Munawar K.S., Habiba U., [Synthetic Approach to Achieve Halo Imine Units: Solid-State Assembly, DFT Based Electronic and Non Linear Optical Behavior](#), *Chem. Phys. Lett.*, **803**: 139843 (2022).
- [47] Ali A., Ashfaq M., Din Z.U., Ibrahim M., Khalid M., Assiri M.A., Riaz A., Tahir M.N., Rodrigues-Filho E., Imran M., [Synthesis, Structural, and Intriguing Electronic Properties of Symmetrical Bis-Aryl- \$\alpha\$, \$\beta\$ -Unsaturated Ketone Derivatives](#), *ACS Omega*: (2022).
- [48] Akoka S., Remaud G.S., [NMR-Based Isotopic and Isotopomic Analysis](#), *Prog. Nucl. Magn. Reson. Spectrosc.*, **120**: 1-24 (2020).
- [49] Lipinski C.A., [Lead-and Drug-Like Compounds: The Rule-of-Five Revolution](#), *Drug Discovery Today: Technol.*, **1(4)**: 337-341 (2004).
- [50] Zubrienė A., Smirnovienė J., Smirnov A., Morkūnaitė V., Michailovienė V., Jachno J., Juozapaitienė V., Norvaišas P., Manakova E., Gražulis S., [Intrinsic Thermodynamics of 4-Substituted-2, 3, 5, 6-Tetrafluorobenzenesulfonamide Binding to Carbonic Anhydrases by Isothermal Titration Calorimetry](#), *Biophys. Chem.*, **205**: 51-65 (2015).