crystals

MDPI

Article

Isolation, Identification, Spectral Studies and X-ray Crystal Structures of Two Compounds from *Bixa orellana*, DFT Calculations and DNA Binding Studies

Mehtab Parveen [1,*], Mohammad Azeem [1], Afroz Aslam [1], Mohammad Azam [2,*], Sharmin Siddiqui [3], Mohammad Tabish [3], Ali Mohammad Malla [1,4], Kim Min [5], Vitor Hugo Rodrigues [6], Saud I. Al-Resayes [2] and Mahboob Alam [5,*]

1 Division of Organic Synthesis, Department of Chemistry, Aligarh Muslim University, Aligarh 202002, India; azam23961@gmail.com (M.A.); afrozaslam10@gmail.com (A.A.); haiderchem09@gmail.com (A.M.M.)
2 Department of Chemistry, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia; sresayes@ksu.edu.sa
3 Department of Biochemistry, Aligarh Muslim University, Aligarh 202002, India; sharminbcb@gmail.com (S.S.); tabish.biochem@gmail.com (M.T.)
4 Government Degree College, Beerwah, Budgam, Kashmir 193411, India
5 Department of Safety Engineering, Dongguk University, 123 Dongdae-ro, Gyeongju 780714, Gyeongbuk, Korea; kimmin@dongguk.ac.kr
6 CFisUC, Department of Physics, University of Coimbra, 3004-516 Coimbra, Portugal; vhugo@uc.pt
* Correspondence: mehtab.organic2009@gmail.com (M.P.); mhashim@ksu.edu.sa (M.A.); mahboobchem@gmail.com (M.A.)

check for updates

Citation: Parveen, M.; Azeem, M.; Aslam, A.; Azam, M.; Siddiqui, S.; Tabish, M.; Malla, A.M.; Min, K.; Rodrigues, V.H.; Al-Resayes, S.I.; et al. Isolation, Identification, Spectral Studies and X-ray Crystal Structures of Two Compounds from *Bixa orellana*, DFT Calculations and DNA Binding Studies. *Crystals* **2022**, *12*, 380. https://doi.org/10.3390/cryst12030380

Academic Editors: Radu Claudiu Fierascu, Florentina Monica Raduly and Dinadayalane Tandabany

Received: 10 February 2022
Accepted: 7 March 2022
Published: 11 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Abstract: 4,6-Diacetylresorcinol (**1**) and 3-*O*-methylellagic acid dihydrate (**2**), both biologically significant compounds, were extracted from *Bixa orellana* and studied using IR, ^{1}H, and ^{13}C NMR, and UV-vis spectroscopic techniques. X-ray crystallographic techniques were also used to establish the molecular structure of the isolated compounds **1** and **2**. Geometric parameters, vibrational frequencies, and gauge including atomic orbital (GIAO) ^{1}H and ^{13}C NMR of **1** and **2** in the ground state were computed by the density functional theory (DFT) using B3LYP/6-311G(d,p) basis set backing up experimental studies and established the correct structure of isolated compounds. The parameters obtained from the combined DFT, and X-ray diffraction studies are mutually agreed to establish correct structures of **1** and **2**. In addition, an electrostatic potential map and HOMO−LUMO energy gap were made using the DFT calculation to determine the distribution of energy and the chemical reactivity region of the isolated compounds. The current study also provides further insights into the interaction of compound **2** with ct-DNA using numerous biophysical and *in silico* techniques. Moreover, *in silico* studies indicate that compound **2** binds to the DNA in the minor groove. Lipinski's rule of five revealed a higher tendency of compound **2** towards drug-likeness. The bioavailability and synthetic accessibility score for compound **2** was found to be 0.55 and 3.21, suggesting that compound **2** could serve as an effective therapeutic candidate.

Keywords: *Bixa orellana* (Family: Bixaceae); 4,6-diacetylresorcinol; 3-*O*-methylellagic acid dihydrate; X-ray diffraction; DFT; NMR; frontier molecular orbitals

1. Introduction

The Bixaceae family includes *Bixa orellana* L., also known as annatto. It is a Central and South American shrub that grows 3–6 m tall and is one of the oldest plants to produce natural colors. Originally, the herb was used as body paint, a treatment for heartburn, an insect repellant, a sunscreen, and to fight off evil. It was named after *Francisco de Orellana*, a Spanish conquistador [1]. Constipation, fevers, heartburn, asthma, scabies, ulcers, diarrhea, stomach upset, skin diseases, measles, anecdotal treatment of diabetes, allergy, leprosy, infectious diseases, burns, measles, gonorrhea, diarrhea, asthma, angina,

tumors, skin problems, and urinary infections have all been treated with *Bixa orellana* in different areas of the globe (oral and topic) [2,3] for centuries. Indigenous people have conventionally utilized the pulp from this plant's seeds externally to increase lips radiance, which is how *Bixa orellana* obtained its moniker, "lipstick tree" [4]. *Bixa orellana* is widely used in dairy food coloring and bleaching, specifically bakery products, cream desserts, buttermilk deserts, rice flour, and corn starch [5–7]. In recent decades, several distinct groups of phytoconstituents, including aliphatic compounds, carotenoids, apocarotenoids, sterols, volatile oils, monoterpenes, sesquiterpenes, triterpenoids, and other miscellaneous substances, have been isolated from all parts of this plant [8–10].

In continuance of our ongoing research [11], we herein report the extraction, isolation, spectroscopic, X-ray crystallographic techniques of two compounds [4-acetylresorcinol (**1**), 3-*O*-methylellagic acid dihydrate (**2**)], derived from the *Bixa Orellana* leaves. Physiochemical spectral data (FTIR, UV, ^{1}H NMR, ^{13}C NMR, and MS spectral analysis), including Single Crystal X-ray Diffraction was used to establish the chemical structure of two isolated phytoconstituents. Compound **1** has not yet been isolated out of this plant source, to our knowledge. The structure of **2**, which crystallized as white crystals, was validated by X-ray crystallography. Quantum chemical calculations using the DFT (B3LYP) theory have yet to be performed on these molecules **1** and **2**. The B3LYP functional was adopted because to its significance in quantum chemistry as well as its precision Geometry calculation, IR, NMR spectra, and a variety of other molecular properties for compounds were investigated in this study at the B3LYP/6-311(d,p) level of theory. The FTIR vibrational bands are interpreted using harmonic force field calculation without scaling, which has a lower computational cost and is sometimes important in extremely large chemical substances. Gauge independent atomic orbital (GIAO) approach is used to analyze NMR spectra (^{1}H and ^{13}C) in gaseous and solvent phases at the same level of theory. Theoretical results were compared to experimental data and found to be in good agreement.

2. Materials and Methods

2.1. Reagents and Apparatus

Silica gel (60–120 mesh) used for various chromatographic procedures was supplied by Merck (India) and Merck (Germany). The various solvents systems employed for thin layer chromatography were benzene-chloroform (8:2) and petroleum ether (60–80 °C)-benzene (1:1). A Kofler block was used to record melting points of compounds and were uncorrected. Carlo Erba analyzer model 1108 was used for elemental analysis (C, H, N). The IR spectra on KBr pellets were obtained using Interspec-2020 (FTIR) and Shimadzu IR-408 Perkin-Elmer 1800 (FTIR). UV-Vis recordings in methanolic solution were made with a Shimadzu UV-1800 UV-Vis spectrophotometer (Shimadzu, Japan). ^{1}H NMR and ^{13}C NMR spectra were performed in $CDCl_3$/DMSO-d_6 with TMS as an internal standard on Bruker Avance II 400 MHz spectrometers. Chemical shifts were expressed in ppm (δ). In the MS (EI) mode, mass spectra were recorded on JEOL D−300 mass spectrometer. Iodine vapor was used to verify purity for the compounds by observing spots on TLC.

2.2. Plant Material

Dr. Athar Ali Khan, Taxonomist, Department of Botany, Aligarh Muslim University, Aligarh, identified a specimen plant of *Bixa orellana* L. (Family: Bixaceae) taken from Allahabad University in Allahabad, India. A specimen with voucher number 243 was registered at the Botany Department, AMU, Aligarh.

2.3. Extraction and Isolation

The leaves of *Bixa orellana* (Family: Bixaceae) were shade-dried and ground into powder (2 kg). A dark green gummy mass was obtained after extracting the air-dried powdered leaves 3 times with 95% ethanol at reflux temperature and evaporating the solvent under reduced pressure. Petroleum ether (60–80), benzene, ethyl acetate, acetone, and methanol were applied to extract the dark green gummy mass. Distillation was used to extract the

solvent. Under reduced pressure, the petroleum ether and benzene extracts were concentrated, yielding a greenish sticky substance. The TLC study revealed that petroleum ether and benzene extracts acted the same way in different solvents, thus extracts were mixed. To elute the mixed petroleum ether–benzene extract from a silica gel column, a gradient of increasing solvents was used, including petroleum ether, petroleum ether–benzene, benzene, ethyl acetate, and eventually methanol. Those fractions with similar TLC characteristics and similar IR spectra were grouped. To obtain pure compounds **1** and **2**, they were purified using repeated column chromatography followed by fractional crystallization and obtained 4,6-diacetylresorcinol (compound 1) and 3-*O*-methylalgeic acid dihydrate (Compound **2**) was identified by comparison of m.p., TLC, Co-TLC and spectral data (IR, ^{1}H NMR, ^{13}C NMR and Mass) of authentic samples.

2.4. Spectral Analysis of Isolated Compounds

(1) 4,6-diacetylresorcinol: Colorless shining crystals (40 mg), m.p. 164–166 °C (lit. [12] 164–166 °C). Anal. Calc. for $C_{10}H_{10}O_4$; C, 61.85; H, 5.19; found: C, 61.83; H, 5.21. UV (MeOH) λ_{max}: 212, 275 and 312 nm. IR (KBr, ν_{max} cm^{-1}): 1589, 1490 (C=C), 1658 (C=O), 3430 (OH). ^{1}H NMR (400 MHz, DMSO-d_6, ppm) δ: 2.79 (s, 6H, 2 × $COCH_3$), 6.43 (s, 1H, H-2), 8.21 (s, 1H, H-5), 12.91 (s, 2H, 2 × OH); ^{13}C NMR (100 MHz, DMSO-d_6, ppm) δ: 26.12 (2 × CH_3), 104.99 (C-5), 113.66 (C-3), 136.31 (C-2), 168.93 (C-4, C-6), 202.52 (C-1); MS (ESI) (m/z): 194.06 [M$^+$] ($C_{10}H_{10}O_4$).

(2) 3-*O*-methylellagic acid dihydrate: Ethyl acetate fraction yielded compound **2**. Crystallization from chloroform and methanol gave greenish white crystals (60 mg), m.p. 350–352 °C (lit. [11]). Anal. Calc. for $C_{15}H_{12}O_{10}$; C, 51.14; H, 3.41; found: C, 51.13; H, 3.43. UV (MeOH) λ_{max}: 245–400 nm. IR (KBr) ν cm^{-1}: 3310 (OH), 1721 (C=O), 1450 (C=C), 1171 (C-O). ^{1}H NMR (400 MHz, DMSO-d_6, ppm) δ: 10.59 (s, 2H, C-4 OH, C-4′ OH), 8.72 (s, 1H, C-3′ OH), 7.51 (s, 1H, H-5), 7.46 (s, 1H, H-5′), 4.15 (s, 3H, OCH_3). ^{13}C NMR (100 MHz, DMSO-d_6) δ: 159.02–159.03 (C-7 and C-7′), 154.10–154.14 (C-4 and C-4′), 143.24 (C-3′), 142.02–142.04 (C-2 and C-2′), 141.35 (C-3), 113.20–113.22 (C-6 and C-6′), 112.02–112.04 (C-1 and C-1′), 111.10–111.12 (C-5 and C-5′), 62.56 (3′-OCH_3). MS (EI): (m/z) 352.41 [M$^{+\bullet}$] ($C_{15}H_{12}O_{10}$).

2.5. Crystallographic Analysis

A Bruker Kappa APEXII CCD X-ray diffractometer was used for X-ray diffraction of single crystals of compounds **1** and **2** at ambient temperature using graphite-based Mo-K radiation of monochromatic wavelength (λ = 0.71073 Å). SADABS [13] was used to make absorption adjustments. The direct technique of SHELXL−97 was used to identify the expected structure of compounds **1** and **2** and anisotropically refined (non-H atoms) using SHELXL−97's complete matrix least-squares on F2 [14]. The hydrogen atoms in the structures were placed in a superlative location and refined using isotropic displacement parameters. PLATON [15,16] and ORTEP−3 [17] were used to analyze the layout of the figures of the structures. The experimental parameters related to the monocrystalline X-ray examination of **1** and **2** are described in Table S1 (Supplementary Material).

2.6. Computational Procedure

The isolated compounds (**1** and **2**) from *Bixa orellana* in their ground electronic states were subjected to density functional theory (DFT) calculations using density functional B3LYP-GD3 (Becke's 3 parameter hybrid functional for the exchange part, the Lee–Yang–Parr (LYP) correlation function, and Empirical D3 Dispersion [18]) and 6-311G(d,p) basis set by Gaussian package [19] using WebMO interface [20]. Under a strict convergence criterion, the isolated compound geometries in Cartesian representation were completely optimized in the ground electronic state. The ground state optimized geometry of the molecule was used to detect harmonic vibrational frequencies at the same level of theory. As a result, no imaginary frequencies were found, leading to the discovery of a true minima on the potential energy surface. The visualization programs GaussView [21], ChemCraft [22],

and GaussSum [23] were used to construct realistic images, electronic transitions and vibrational assignments using animated modes. In the solution phase ($CHCl_3$ for compound **1** and DMSO for compound **2**), the GIAO-B3LYP/6-311G(d,p) level of theory was employed to measure NMR spectra of compounds applying gauge independent atomic orbital (GIAO) approach. In addition, UV-Vis spectra of the isolated compounds **1** and **2** were simulated using polarizable continuum model (PCM). Excitation energy, oscillatory power, wavelength and HOMO–LUMO energy differences were obtained at the basis set of TD-B3LYP-D3/6-311G (d,p). Molecular electrostatic potential (MEP) diagrams of both isolated compounds were generated using same level theory applied for optimization of the isolated compounds in order to predict interactions with neighboring molecules in term of electrophilic and nucleophilic sites.

2.7. ct-DNA Binding Studies

2.7.1. Sample Preparation for DNA Binding Experiments

To make a 4 mM stock solution, compound **2** was dissolved in ethanol. The compound working solutions were formed in accordance with the requirements. To dissolve calf thymus DNA, 10 mM Tris HCl (pH 7.4) was added to the buffer (ct-DNA). The DNA was kept at 4 °C. To achieve a homogeneous solution, the solution was occasionally stirred. To evaluate dissolved calf thymus DNA purity, the (A_{260}/A_{280}) absorbance ratio was utilized. The absorbance ratio of solution was 1.81, indicating that it is clean and does not require to be purified further. DNA content was calculated using the extinction value of molar 6600 $M^{-1}cm^{-1}$ at 260 nm for a single nucleotide.

2.7.2. UV-Visible Absorption Spectroscopy

UV-visible absorption spectrum of compound **2** was measured using a Shimadzu UV-1800 spectrophotometer. At a fixed concentration of 200 M compound **2**, the reaction mixture was titrated with sequential concentrations of ct-DNA (0–50 M). A 10 mM Tris HCl buffer was used to adjust the baseline (pH 7.4).

2.7.3. Steady State Fluorescence

An RF-6000 Shimadzu spectrofluorometer was used to obtain steady-state fluorescence spectra of compound **2**. Spectra were obtained by adding ct-DNA (0–125 M) to a constant concentration of compound **2** (200 M). Compound **2** has a 273 nm excitation wavelength. The emission wavelength was adjusted accordingly, but the widths of the excitation and emission slits remained at 5 nm. Stern–Volmer constant (K_{sv}), binding constant (K_b) and constant (K_q) of bimolecular enhancement were all calculated using the data.

2.7.4. Circular Dichroism (CD)

A JASCO J-1500 circular dichroism spectrophotometer was used to collect and analyze the structural changes in ct-DNA produced by compound **2** interaction. The samples, which contained 200 µM of ct-DNA alone and 200 µM of compound **2** complexed with ct-DNA (1:1), were photographed at 200 nm/min in 190–340 nm wavelength range. In each batch of the experiment, a 10 mM Tris HCl (pH 7.4) buffer was utilized, as well as any necessary background adjustments. An average of 2 scans was used to produce the results.

2.7.5. Docking Studies

Docking analysis of 3-*O*-methylellagic (**2**) acid dihydrate was performed using Autodock 4.0 to describe the binding mode of compound **2** to DNA. The crystallographic information file (cif) of compound **2** was converted to PDB format using Mercury (Cambridge Crystallographic Data Centre). The RCSB database was used to retrieve the crystal structure of the synthetic DNA dodecamer d(CpGpCpGpApApTpTpCpGpCpG) with Pdb ID:1BNA. Avogadro was used to minimize the energy of the structures. Before docking, water molecules were stripped from the structures, Kollman charges were introduced, and polar hydrogen was supplied. Docking consisted of 10 runs using the Lamarckian

genetic approach. The default settings for all other parameters were used. The x, y, and z-axis dimensions of the grid box for compound **2** were set to 26, 30, and 48 Å, respectively. The docking was conducted with a grid spacing of 0.375. UCSF Chimera 1.01, PyMOL and Accelrys Discovery Studio 4.5 were used to investigate the docked pose in the lowest configuration in order to determine the likely binding mode of compound 2 with DNA.

2.7.6. Prediction of Drug-Likeness (Lipinski's Rule of Five)

A rule of thumb (Lipinski's rule) can be used to analyze the drug-likeness qualities of chemical compounds and aid in the differentiation of drug-like and non-drug molecules. The drug-likeness of compound **2** was determined by calculating physiochemical characteristics such as the octanol-water partition coefficient (log P), molecular weight (MW), molar refractivity, polar surface area, hydrogen bond donors and acceptors. The following requirements apply: (i) less than 500 Da molecular mass, (ii) less than 5 hydrogen bond donors, (iii) strong lipophilicity, (iv) fewer than 10 acceptors for hydrogen bonds, and (v) a molar refractivity of 40 to 130. A tested compound that meets 2 or more of these rules is referred to as a drug-like molecule. Using online software tools [24], Lipinski's rule of five was employed to profile molecules at pH 7.

2.7.7. Synthetic Accessibility and Bioavailability Score Prediction

The SwissADME web tool was used to predict the bioavailability and synthetic accessibility score of the isolated molecule **2**. A compound can only be an effective drug [25] if it has a high bioavailability score and a low synthetic accessibility score (1 = easy to synthesize; 10 = difficult to synthesize).

3. Results and Discussion

3.1. Crystal Structure, Molecular Geometry and IR Spectral Analysis of **1** *and* **2**

The isolated compounds (Figure 1) have four molecules per unit cell (Z = 4) in a P 21/c crystal structure. From single crystal X-ray diffraction data, it is found that the crystal of both compounds (Figure 2A,B belongs to monoclinic crystal structure with the following dimensions; a = 7.095(12) Å, b = 11.374(19) Å, c = 11.66(2) Å and the angle of $\alpha = 90^{\circ}$, β 100.46(3) $^{\circ}$ and $\gamma = 90^{\circ}$ for compound **1** that was reported synthetically in literature [11] and a = 10.1687(6) Å, b = 6.9037(5) Å, c = 21.3367(13) Å with the angle of $\beta = 107.668(3)$, and $\alpha = \gamma = 90^{\circ}$ for compound **2** that was documented elsewhere [26,27]. In the Supplementary Data (Table S1), full information on crystal structures of compounds **1** and **2** was provided. Optimization of isolated compounds **1** and **2** were performed using B3LYP-GD3 with theory level 6–311G (d,p) and optimized structure of both were displayed in Figure 2B and Figure 3B with numbering pattern of the investigated compounds. Geometric variables such as bond lengths, bond angle, and torsional angle obtained from optimized structures (**1** and **2**) by quantum chemical approach are shown in Table 1. The energy of optimized structures **1** and **2** has the energy of −1178.55087987 and −688.22154646 au, respectively. Compound **2** has lower energy (plus the negative) compared to compound **1**. It can also be assumed that compound **2** is more stable than compound **1**. In the optimized structure of 4,6-diacetylresorcinol (**1**), the bond lengths of C=O, C-O- and -OH were observed as 1.23, 1.24 and 1.05 Å, respectively, as given in Table 1 and these values were correlated to bond lengths obtained from X-ray single crystallography experiment (Table 1). Similarly, in the optimized of 3-*O*-methylellagic acid (**2**), the bond lengths of C12-O1, C13-O2, C17-O4, C24=O18, C24-O11, C25-O11, O2-H3, and C13-O1 were noticed as 1.34, 1.35, 1.19, 1.20, 1.39. 1.37, 0.96, and 1.43 Å, respectively, which displayed the presence of carboxylic, ester, and hydroxyl group in the structure of compound **2**. These bond length values were found to correspond with the values obtained as shown in Table 1. from a single X-ray study. Bond lengths and angles of certain atoms were found to marginally deviate from the corresponding experimental data. The deviation between certain atoms of the isolated compounds was found to be slightly large in the case of dihedral angles (°). Due to the non-covalent intermolecular interaction present in the crystal, these deviations may result

from the fact that the theoretical calculations were performed under vacuum for a single isolated compound. The molecular geometries of the isolated compounds **1** and **2** obtained using XRD and DFT computation with basis set B3LYP/6-311G(d,p) were superimposed atom by atom (Figure 2C and Figure 3C), and RMSE values of both compounds were found to be 0.1345 and 0.1063 Å, respectively. These low values indicated that theoretical and experimental structural parameters agreed well.

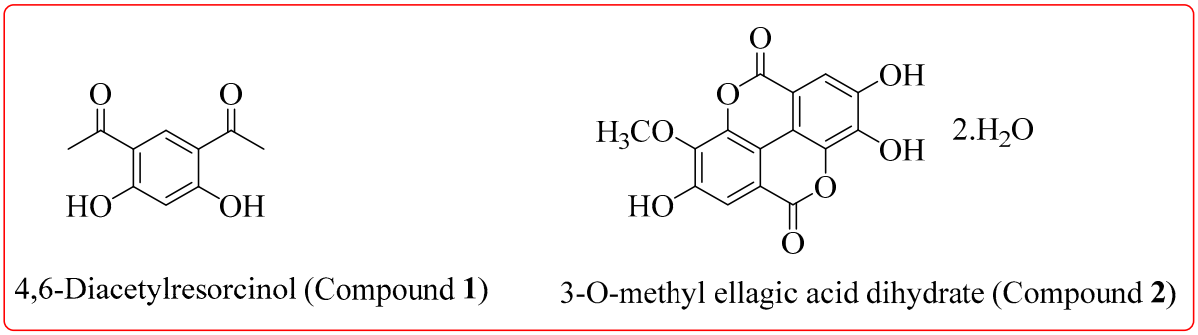

Figure 1. Isolated compounds (**1** and **2**) from *Bixa orellana*. leaves.

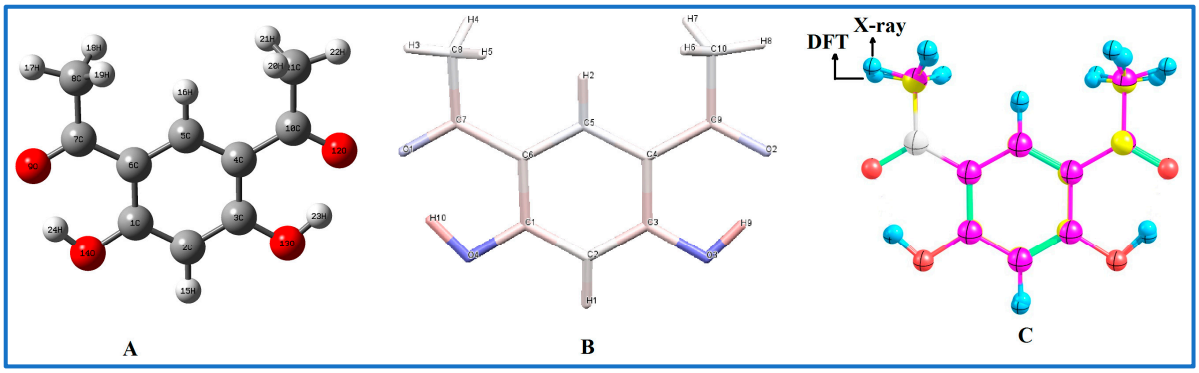

Figure 2. (**A**) Optimized geometry, (**B**) XRD of 4,6-diacetylresorcinol (**1**) and (**C**) superposition of the optimized structure on the experimental structure atom by atom.

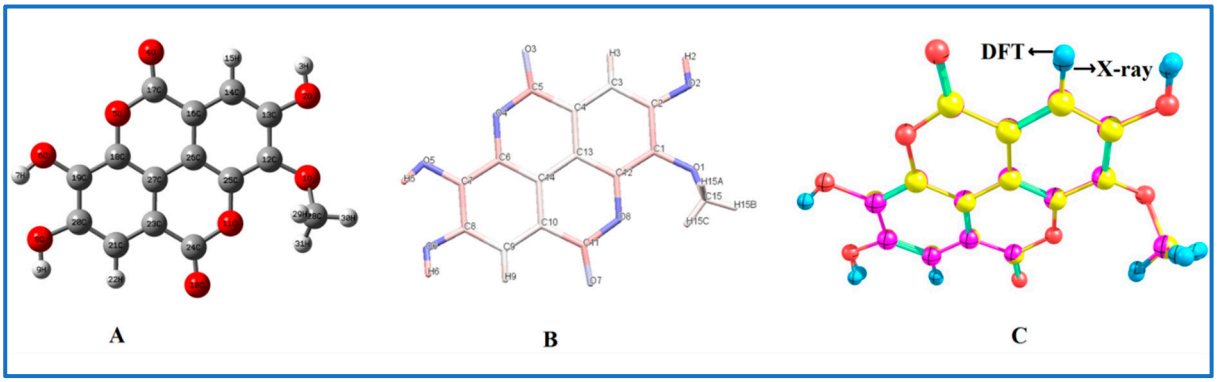

Figure 3. (**A**) Optimized geometry (**B**) XRD of 3-*O*-methylellagic acid dihydrate (**2**) and (**C**) superposition of the optimized structure on the experimental structure atom by atom.

Table 1. Comparison of important bond lengths [Å] bond angles [°] and dihedral angles [°] obtained by X-ray study and DFT/B3LYP/6-311G(d,p) for molecules **1** and **2**.

Bonds (Å)	Expt.	DFT	Angles (°)	Expt.	DFT	Dihedral Angles (°)	Expt.	DFT
4,6-Diacetylresorcinol (1)								
C1-C2	1.38	1.39	C2-C1-C6	121.0	120.2	C6-C1-C2-H15	−175.2	−179.9
C1-C6	1.41	1.43	C2-C1-O14	117.9	118.1	O14-C1-C2-C3	179.0	180.0
C1-O14	1.34	1.33	C6-C1-O14	121.0	121.6	O14-C1-C6-C5	−179.5	−180.0
C2-C3	1.37	1.39	C1-C2-C3	119.8	120.8	O14-C1-C6-C7	−0.916	−0.004
C2-H15	0.99	1.08	C1-C2-H15	116.9	119.5	C1-C2-C3-O13	−178.7	−179.9
C3-C4	1.42	1.43	C3-C2-H15	122.8	119.5	C15-C2-C3-O13	−5.6	−0.0143
C3-O13	1.34	1.33	C3-C4-C4	117.9	120.2	C2-C3-O13-H23	174.6	179.9
C4-C5	1.38	1.39	C2-C3-O13	118.1	118.1	C4-C3-O13-H23	−5.0	−0.0051
C4-C10	1.46	1.46	C4-C3-O13	120.9	121.6	C3-C4-C10-O12	1.3	0.00076
C5-C6	1.39	1.39	C3-C4-C10	120.0	119.7	C5-C4-C10-O12	−178.2	179.9
C5-H16	0.99	1.08	C5-C6-C7	121.7	122.4	C5-C6-C7-O9	176.9	179.9
C6-C7	1.46	1.46	C6-C7-C8	120.9	120.0	O9-C7-C8-H17	10.5	0.0172
C7-C8	1.48	1.51	C6-C7-O9	119.8	121.0	O9-C7-C8-H18	−106.3	−120.0
C7-O9	1.23	1.23	C8-C7-O9	119.1	118.8	O9-C7-C8-H19	135.7	120.0
O14-H24	1.05	0.99	C1-C14-O24	101.6	106.9	O12-C10-C11-H20	−125.3	−119.9
3-*O*-Methylellagic acid (2)								
O1-C12	1.35	1.34	C12-O1-C28	115.9	120.4	C28-O1-C12-C13	115.1	146.0
O1-C28	1.44	1.43	H3-O2-C13	109.5	108.9	C28-O1-C12-C25	−68.7	−36.5
O2-H3	0.82	0.96	C17-O5-C18	122.1	122.7	C12-O1-C28-H29	−53.0	−42.5
O2-C13	1.34	1.35	H7-O6-C19	109.4	108.0	H3-O2-C13-C12	172.8	179.2
O4-C17	1.20	1.19	C24-O11-C25	122.4	123.6	H3-O2-C13-C14	−7.3	−1.1
O5-C17	1.37	1.39	O1-C12-C13	118.7	116.0	C18-O5-C17-O4	−178.0	179.5
O5-C18	1.38	1.36	O1-C12-C25	122.0	126.0	C18-O5-C17-C16	1.9	−0.4
O6-H7	0.81	0.96	C13-C12-C25	119.0	117.7	C17-O5-C18-C19	178.1	−179.3
O6-C19	1.33	1.34	O2-C13-C12	114.8	115.7	H7-O6-C19-C18	−170.3	179.8
C18-C27	1.38	1.39	O2-C13-C14	124.3	122.9	H7-O6-C19-C20	10.2	−0.1
C23-C24	1.45	1.46	C13-C14-H15	120.4	120.8	O10-C24-O11-C25	178.9	−178.8
C19-C20	1.40	1.41	C13-C14-C16	119.1	120.0	O10-C24-C23-C21	0.9	−0.4
C28-H31	0.95	1.09	O8-C20-C19	113.9	113.6	H15-C14-C16-C17	−0.04	0.2
C21-H22	0.93	1.08	C19-C18-O5	118.3	118.4	H15-C14-C16-C26	179.4	−179.6
C16-C25	1.39	1.40	C18-C27-C26	118.7	118.5	C18-C27-C23-C24	178.5	179.6

The IR frequencies of the optimized geometry of the isolated compounds **1** and **2** were computed at the same basis set using a double harmonic approximation. Experimental and simulated IR spectra of compounds **1** and **2** are demonstrated in Figures 4 and 5. Table 2 displays the positions of some significant infrared (FTIR) vibrational bands of both compounds as well as their theoretical values (in cm^{-1}), intensities, and assignments. In general, harmonic frequencies overestimate the experimental frequency due to systematic errors triggered by basis set incompleteness error, different approximations, and lack of vibrational anharmonicity used in both compounds in the current DFT calculations. The IR data of both compounds **1** and **2** obtained from theoretical and experimental were compared with the literature [11].

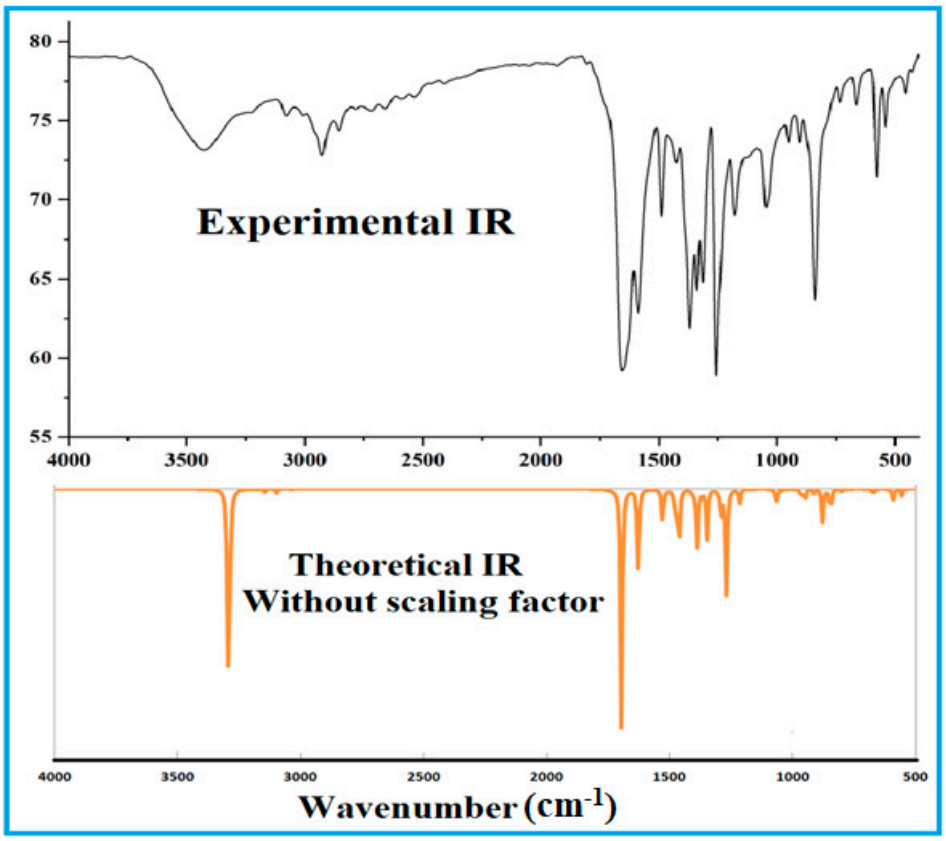

Figure 4. Theoretical and experimental IR spectra of compound **1**.

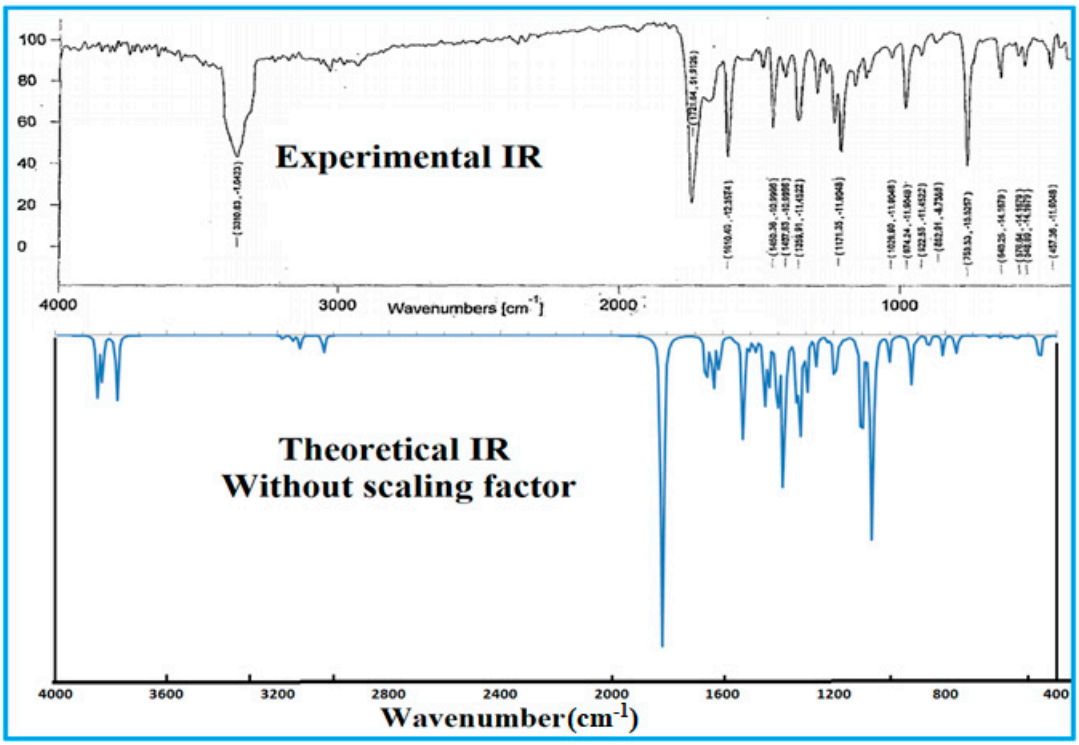

Figure 5. Theoretical and experimental IR spectra of compound **2**.

Table 2. Experimental and theoretical IR frequencies (in cm^{-1}) along with IR intensity (in km/mol) and vibrational assignments of compound **1** and **2**.

	FTIR	$\omega_{harmonic}$	Assignments		FTIR	$\omega_{harmonic}$	Assignments
4,6-Diacetylresorcinol (1)	3427	3295	νOH	**3-O-Methylellagic acid dihydrate (2)**	3433	3848–3778	νOH
		3291	ν_{asy}OH		3159	3184–3179	νCH (ar)
		3212	νCH (ar)		2927	3146, 3119, 3035	ν_{asy}CH, νCH
	3082	3207	νCH (ar)		1723	1820, 1813	νC=O
		3147	νCH (ar)		1600	1659	νCC, βHOC
		3146	νCH (ar)			1650, 1634, 1612, 1560, 1549, 1526	νCC, ν_{asy}CC γHCC
	2925 2851	3097	νCH		1518	1504	βCOC, βHCH, τHCOC
		3036	νCH		1422	1489	βHCH,τHCOC
	1652	1696	ν_{asy}C=O			1450	νCC, βHOC
		1694	νC=O		1361	1382	νO-C, ν_{asy}C-O
		1661	ν_{asy}C=O, Ar-ring		1340	1372, 1322	ν_{asy}CC, βHOC
	1588	1627	νCC, ν_{asy}CC, βHOC		1293	1319	ν_{asy}O-C, βHCC
		1527	βHOC			1297	νO-C, γHOC, βHCC
	1486	1482	βHCH, τHCCC		1190	1221	βHOC, βHCH, τHCOC
		1472	βHCH, τHCCC			1175	βHOC, βHCC
	1425	1398	βHCH		1060	1167	βHCH, τHCOC
		1290	νO-Cand βHCC (ar)		966	1053, 1000	νO-C
		1266	νCC, βHOC, βCCC		912	921	γCCC
		1214	νO-C, βHCC (ar)		867	861	τHCCC
		1092	βCCC, τHCCC			859	outCCCC
		964	γ HCH, τHCCC		800	807	mix vibration
		914	τHCCC			759	outOCOC
		762	βCCC			699	βCOC, outOCCC
		725	outCCCC			640	βCOC, βOCC,
		674	βCCC		575	614	βCOC, βOCC
		556	βCCC, βCCO			539	νCC, βCCC,

ν-stretching, ν_{asy} asymmetric stretching, β-in plane bending, γ-out of plane bending vibrations, t-torsional vibration.

Both compounds **1** and **2** have groups of C-H, aromatic, OH, and C=O in their structure. Simulated IR spectra using B3LYP/6-311G(d,p) basis set were compared with experimental IR spectra to explain vibrations of functional groups, as shown in Figures 4 and 5. All assignments set out in Table 2 are in accordance with the literature [28]. The bands due to stretching C-H, C=O, C=C, C-C, ring, C-O as well as in-plane and out-of-plane deformation vibrations are shown in Table 2. The bands that appeared in the region of the lower wavenumber were mainly due to mixed vibrations of torsion and out-of-plane

deformation of the rings. In general, the IR bands vibrate in the range of 3700–2700 cm^{-1} due to OH groups (IR Spectrum Table and Chart available on Sigma Aldrich website), with different intensities showing the nature of –OH groups such as free, inter, or intramolecular bonding in chemical compounds. The IR bands displayed in a broad peak in compound **1** and **2** at 3427 and 3433 cm^{-1}, respectively in IR spectra, assigning OH groups to hydrogen bonding interactions. In the theoretical spectra of 4,6-Diacetylresorcinol (**1**) and 3-*O*-methylellagic acid dihydrate (**2**), phenolic hydroxyl groups were vibrated at 3295–3291cm^{-1} and 3848–3778 cm^{-1}, respectively, without applying the scale factor. In the aromatic structure, the characteristic region for C-H stretching bands usually falls just above 3000 cm^{-1}. The stretching bands occurred in multiple weak and narrow peaks at 3082 cm^{-1} for **1** and 3159 cm^{-1} for compound **2** of the FTIR spectra while theoretical values of aromatic C-H were observed at 3212–3146 cm^{-1} and 3184–3179 cm^{-1}, respectively. In-plane and out-of-plane C-H bending vibrations exhibit their characteristic bands in the region 1100–1500, 800–1000, and 650–1000 cm^{-1}. In present compounds **1** and **2** in-plane deformation vibrations were observed mixing with vibrations of C-C, C-O-C stretching bands at 1482, 1472, 1290, and 1214 cm^{-1} for **1** and at 1489, 1319, 1297, 1221, 1175, and 1167 cm^{-1} for compound **2**, while out of plane bending vibration were noticed mixing with torsional and out vibration as shown in table. The asymmetric and symmetric stretching bands of the methyl group (C-H) of chemical compounds in FTIR have been documented below 3000 cm^{-1} in the 3000–2840 cm^{-1} range. FTIR spectra of both compounds exhibited bands at 2925, 2851, and 2927 cm^{-1}, representing asymmetric and symmetric stretching bands suggested $–CH_3$ groups in both compounds **1** and **2**, while the calculated values at 3097, 3146, 3119, and 3035 cm^{-1} (without scaling factor) were assigned to methyl groups in the present compounds. In general, the stretching vibration of the normal carbonyl group referred to without hindrance environment appears with a high intensity peak of about 1700 cm^{-1}. Strong bands visualized at 1652 and 1721 cm^{-1} with overlapping shoulders were assigned to C=O stretching vibrations while calculated bands of carbonyl groups using B3LYP/6–311G(d,p) basis set in IR spectra were identified at 1696, 1694, 1661 cm^{-1} for **1** and 1820, 1813 cm^{-1} for **2**. In most cases, the C-O stretching vibration of the aromatic ether, alkyl aryl ether, tertiary alcohol, ester, aliphatic ether, secondary alcohol, and primary alcohol attached to the heterocyclic moiety can cause a strong to medium intensity band in the 1310–850 cm^{-1} region. In the present study, stretching vibration bands at 1425, 1361, and 1293 cm^{-1} were observed in FTIR spectra of compounds (**1** and **2**). Bands at 1290, 1214, 1382, 1319, 1297, 1053, and 100 cm^{-1} may be due to the various nature C-O stretching vibrations with mixing of other vibrations in the theoretical study, as shown in table. The carbon-carbon stretching vibrations of aromatic ring crop up in the region 1625–1430 cm^{-1}. The C=C stretching vibration of aromatic ring present in compounds (**1** and **2**) display their characteristics bands at 1588 and 1425 for **1**, 1600, 1422 and 1293 for **2** in FTIR while calculated bands at 1627 and 1266 for **1**, 1659, 1450 and 1297 for **2** were noted with mixing of other vibrations of various intensity in IR spectra of the isolated compounds.

3.2. Frontier Molecular Orbital and UV-Visible Spectral Analysis

Frontier molecular orbitals (FMOs) referred to HOMO and LUMO involved in chemical reactions. Figure 6 displays the highest occupied molecular orbitals (HOMO) and lowest unoccupied molecular orbital (LUMO) for compounds **1** and **2** isolated from plant. HOMO is the orbital, which primarily acts as an electron donor, and LUMO is the orbital, which primarily acts as an electron acceptor, and molecular chemical stability is defined by the difference between HOMO and LUMO. It has been noted that the highest occupied molecular orbitals (HOMOs) are largely distributed over aromatic rings, including C=O and –OH functional groups excluding C2H15, C5H16 and methyls in compound **1** and O5 in compound **2** whereas LUMOs is completely delocalized on all of the molecules except OH groups in compound **1** and two OH groups in compound **2** as seen in Figure 6. The energy gap (Eg) between HOMO and LUMO is a critical parameter to study the nature of chemical reactivity because it is a parameter of electron transfer from HOMO to LUMO.

The large value of the gap specifies the most stability and the least reactivity and *vice versa*. The energy gaps of compounds **1** and **2** were found to be 4.73 and 3.97 eV, respectively. The low energy gaps of the two compounds provide an indication of the efficient electronic transition as well as their high reactivity. The HOMO energy describes the susceptibility of the molecule towards electrophilic attacks, while the LUMO energy represents the susceptibility of the molecule to nucleophilic attacks. The lowest singlet to singlet spin allowed transitions, oscillator strengths, absorption wavelengths, and excitation energies for 4,6-diacetylresorcinol (**1**) and 3-*O*-methylellagic acid dihydrate (**2**) in solution phase and comparison of experimental and theoretical UV-vis absorption spectra of both compounds **1** and **2** are presented in Table 3 and Figures 7 and 8. The UV-vis bands observed at 272 and 323 for compound **1** and 255 and 365 nm for compound **2** in solution were assigned to H→L + 1, H-1→L + 1, H→L, and H-1→L + 1, H→L, H-3→L + 1, H-2→L transitions for **1** and H-3→L, H-5→L, H-4→L and H→L, H-1→L, H→L + 1 transitions for compound **2** as shown in Table 3. The corresponding calculated wavelengths were observed to be 268.5, 251.9, 302.6, and 300.1 nm for compound **1** and 259.7, 252.9, 351.7, and 313.6 nm for compound **2** in solution. The alteration of the predicted wavelengths was due to the solvent effect as observed in the two compounds (**1** and **2**).

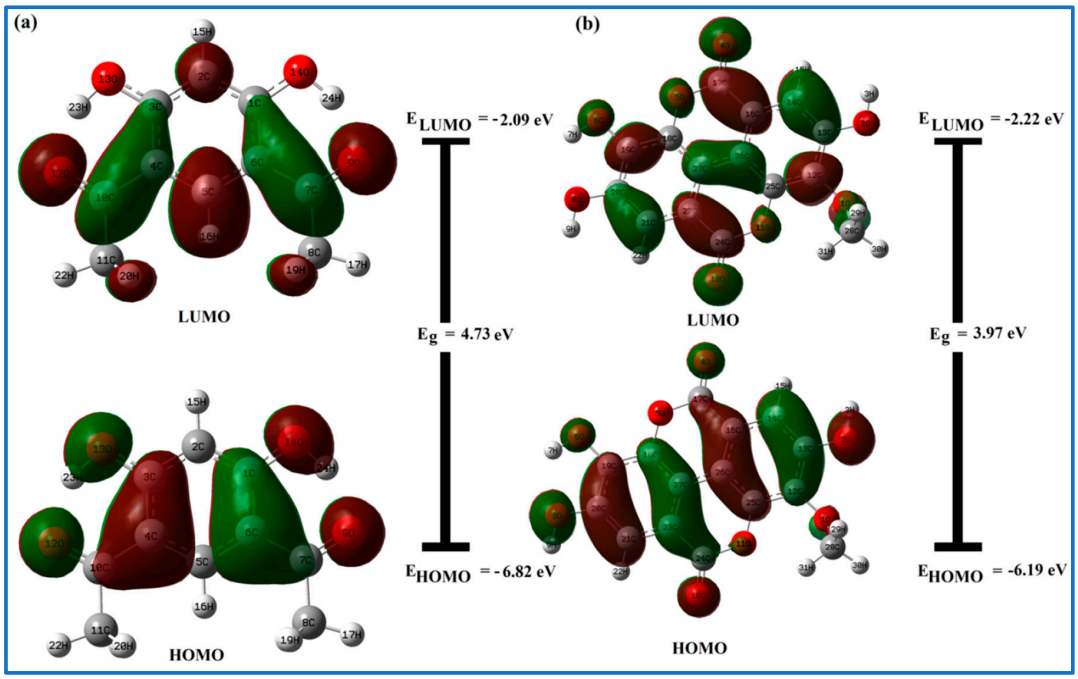

Figure 6. Energy levels, gaps and localization distribution of HOMO and LUMO (frontier molecular orbitals) of the (**a**,**b**) for compounds **1** and **2**, (energies are given in eV).

Table 3. UV/Vis spectral characteristic of compounds **1** and **2** in the solution phase.

	λ_{exp}	λ_{theo}	E(eV)	f	Composition (%)
4,6-Diacetylresorcinol (1)	273.6	268.5	4.6167	0.1915	H→L + 1 (99)
		251.9	4.9212	0.7777	H-1→L + 1 (69), H→L (25)
	323.7	302.6	4.0968	0.1209	H-1→L + 1 (26), H→L (73)
		300.1	4.1314	0.0366	H-3→L + 1 (19), H-2→L (79)
3-*O*-Methylellagic acid dihydrate (2)	255.6	259.7	4.77	0.0498	H-3→L(30)
		252.9	4.90	0.0366	H-5→L (25), H-4→L (50)
	365.3	351.7	3.52	0.2290	H→L(94)
		313.6	3.95	0.0269	H-1→L (51), H→L + 1 (4)

Abbreviation used: H-HOMO, L-LUMO, λ-wavelength (in nm), E-excitation energy, f-oscillator strength.

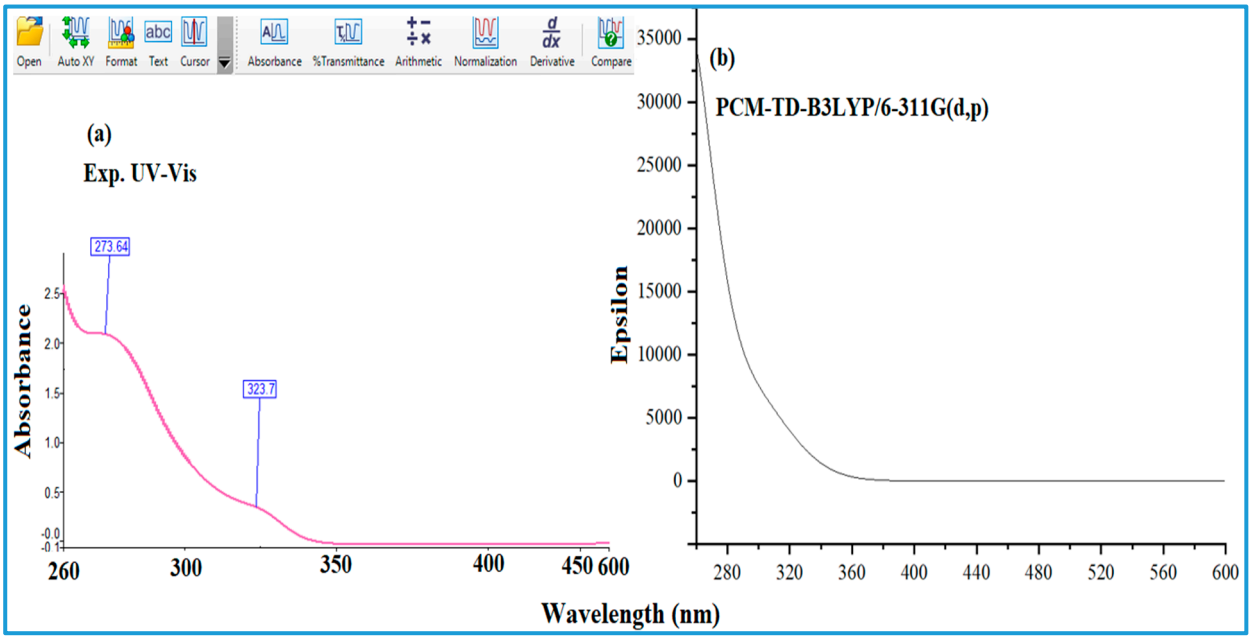

Figure 7. Experimental (**a**) and calculated (**b**) UV-vis absorption spectra of the 4,6-Diacetylresorcinol.

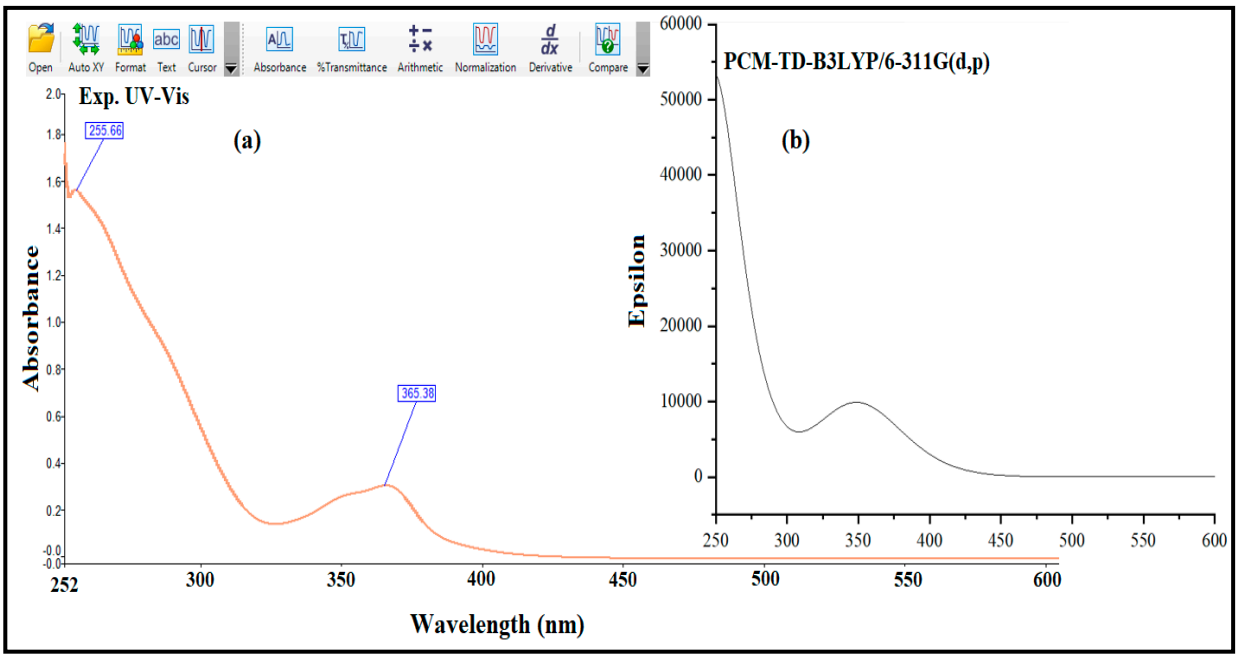

Figure 8. Experimental (**a**) and calculated (**b**) UV-vis absorption spectra of the 3-*O*-methylellagic acid dihydrate.

3.3. NMR Spectral Analysis

The ^{1}H and ^{13}C-NMR spectra of compounds (**1** and **2**) and their chemical shifts were given in ppm downfield form of tetramethylsilane (TMS) and illustrated in Figures 9–12. Figure 9 & Figures 10 and 11 & Figure 12 depict the ^{1}H and ^{13}C NMR spectra of compounds 1 and 2, respectively. Theoretical calculations were performed to validate the accuracy of the experimental data collected. Prior to calculating the theoretical chemical shifts in term of NMR, the molecular geometry of the compounds was optimized using GIAO-B3LYP with 6–311G(d,p) level of theory. Deuterium exchangeable proton of the –OH group in **1** showed a chemical shift at 129 ppm based on its integration of two protons as a singlet peak in ^{1}H-NMR spectrum of compound **1**. Its corresponding calculated chemical

shift was observed at 12.87 ppm. Other singlets resonated at 8.2 and 6.4 ppm showed aromatic protons flanked by methyl ketone and hydroxyl groups, respectively, whereas theoretical singlets of aromatic protons were resonated at 8.1 and 6.3 ppm in solution phase suggesting that chemical shifts were quite similar to experimental values leading to the identification of exact structure of compound **1**. Chemical shifts at 2.1 and 1.6 ppm may contribute to the methyls attached to the ketone groups, corresponding singlets at 2.7 and 1.9 ppm in theoretical calculations were very close to experimental studies as shown in Figure 9. As shown in Figure 11 of compound **2**, singlets refer to methyl hydrogens divided into three and observed in the theoretical spectrum at 3.5, 4.1 and 4.8 ppm, which were observed as a peak at 4.2 ppm in experimental spectrum because the methyl protons experienced the same chemical environment that caused their overlaps. At 7.4 and 7.6 ppm, aromatic protons were noted as singlets of two aromatic ring protons and corresponding peaks observed in the theoretical spectrum at 7.60 and 7.63 ppm were very close to the experimental spectrum. The deuterium exchangeable proton of the three –OH groups in compound **2** displayed chemical shifts at 4.8, 5.0, and 6.0 ppm as three singlet peaks in theoretical ^{1}H NMR spectrum of the compound **2**. However phenolic groups attached to various chemical environments on aromatic rings were observed at 8.7 ppm for one proton and 10.5 ppm assigned to two protons. The theoretical values of NMR displayed good correlation with experimental values for compounds **1** and **2** except discrepancy was noted for phenolic protons because of lability of the hydroxyl proton and proximity of another phenolic group for compound **2**.

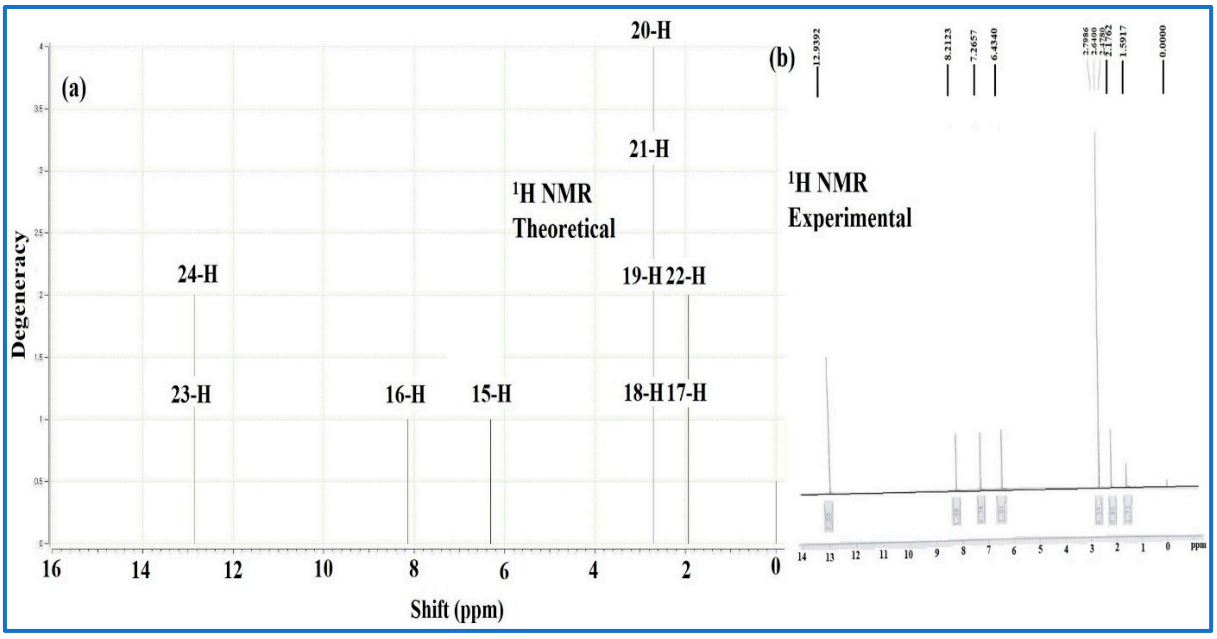

Figure 9. Calculated (**a**) and experimental (**b**) ^{1}H-NMR spectra of the isolated compound **1**.

The signals characterized in the ^{13}C NMR spectra of molecules **1** and **2** (Figures 10 and 12) showed resonance at 206 and 175 for 1 and 162, 162, 159, 154.4, and 144.1 ppm for compound **2,** indicating the presence of carbonyl carbons and carbons attached –OH groups present in aromatic rings in various chemical environments with different chemical shifts in isolated compounds **1** and **2**. Corresponding signals in experimental ^{13}C NMR of compounds **1** and **2** were observed at 202 and 168.9 ppm for **1** and 159, 154, and 143.2 ppm for **2** are closed to theoretical ^{13}C NMR spectra of both compounds. Other important signals observed in the theoretical ^{13}CNMR spectrum at 106.4, 115.3, and 141 ppm were attributed to carbon flanked by −OH groups, carbons attached to carbonyl groups, and carbon flanked by carbonyl groups for compound **1** were found to correlate with the signals at 105, 113.6, and 136.2 ppm in experimental ^{13}C NMR of compound **1**. Signals resonated at 26 ppm

was attributable to carbon atom of the methyl groups bonded to the oxygen atom in experimental spectrum of compound **1**. The estimated theoretical value of 26.4 ppm of methyl carbon is in accordance with the experimental results. Some other ^{13}C NMR signals of aromatic carbons directly influenced by functional groups and substituted carbon atoms including methyl carbon were observed in theoretical spectrum at 140 (-OH-C-C18-O-), 146 (-C12-O-CH3) and 149 ppm (C25-O-C=O), 112.3 (H-C21), 114.2 ppm (H-C14) and 61.2 ppm (O-C28-) that were resonated in similar pattern with slightly deviation in experiment ^{13}C NMR of compound **2**.

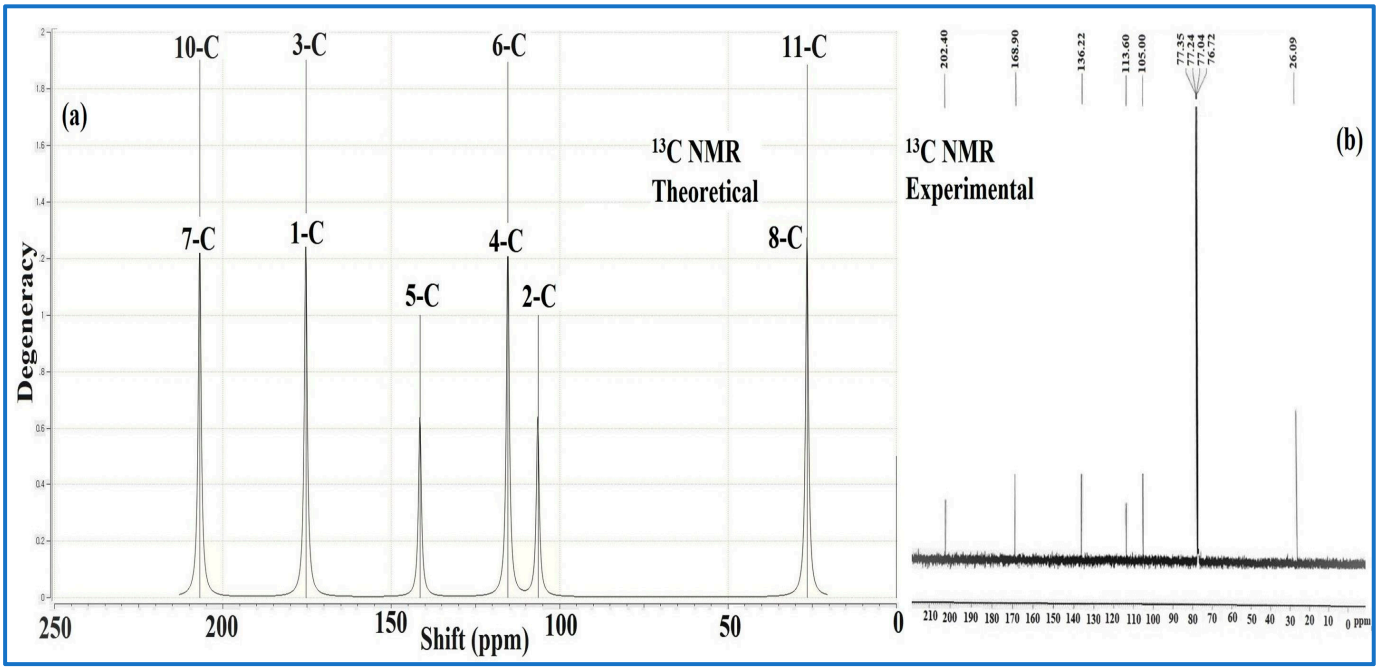

Figure 10. Calculated (**a**) and experimental (**b**) ^{13}C-NMR spectra of the isolated compound **1**.

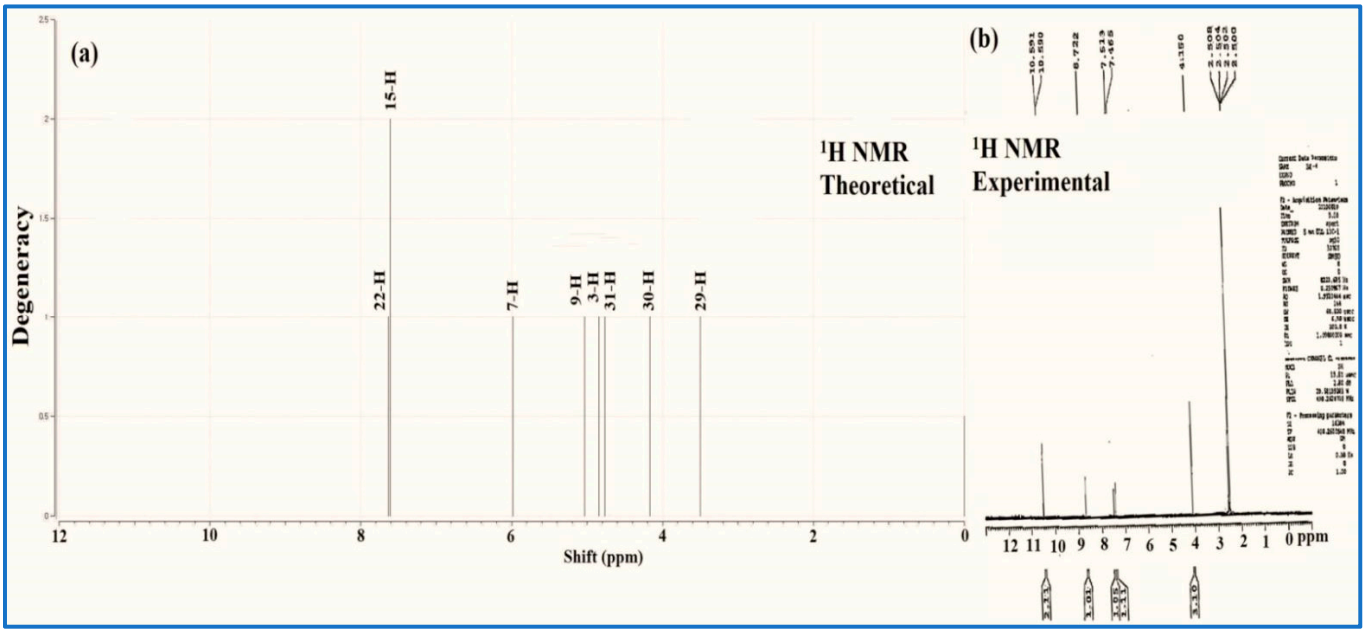

Figure 11. Calculated (**a**) and experimental (**b**) ^{1}H-NMR spectra of the isolated compound **2**.

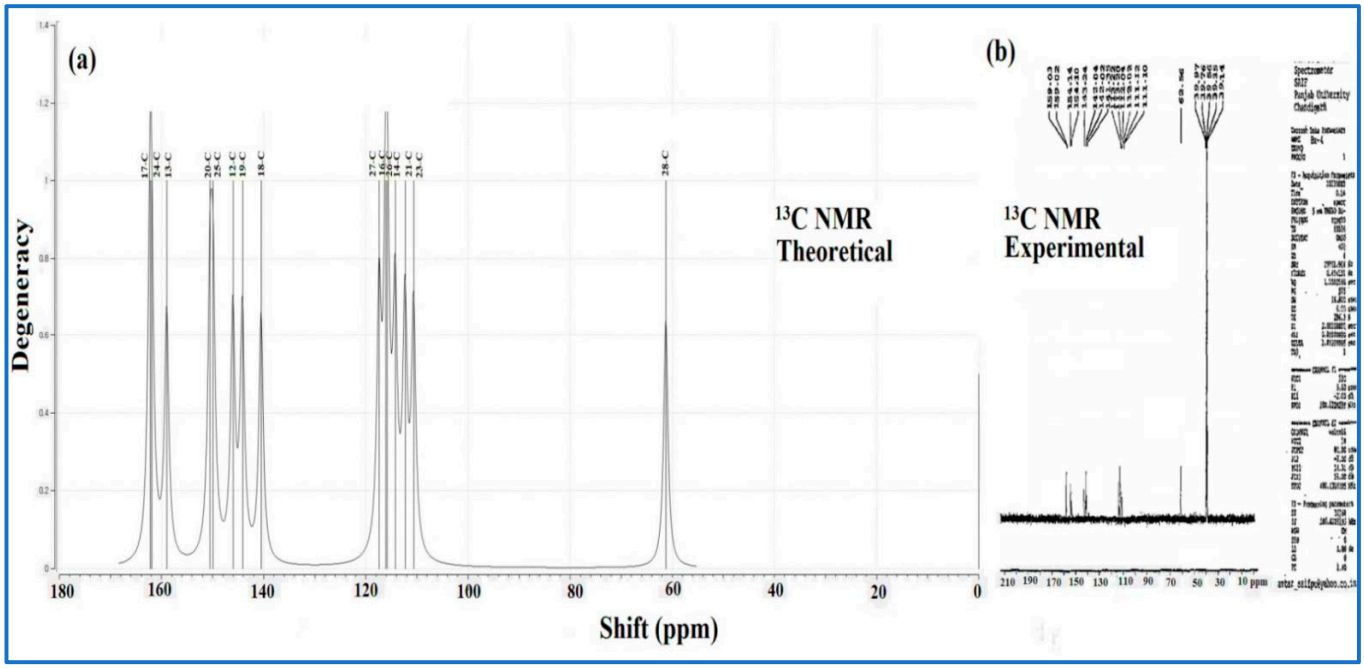

Figure 12. Calculated (**a**) and experimental (**b**) ^{13}C-NMR spectra of the isolated compound **2**.

3.4. Molecular Electrostatic Potential (MEP)

In the study of biological recognition processes and hydrogen bonding interactions, the MEP map is a useful tool for qualitatively interpreting electrophilic and nucleophilic reactions [29–31]. Molecular electrostatic potential of 4,6-diacetylresorcinol (**1**) and 3-*O*-methylellagic acid dihydrate (**2**) are calculated using the optimized structures of both compounds applying at B3LYP/6-311G(d,p) basis set, and their plots are shown in Figure 13. Colored figures of compounds **1** and **2** were a significant sign of inter- and intramolecular interactions and reactivity of molecules. The red to blue color of the graphs reflects the electron-rich to electron-poor regions. Carbonyl oxygen and hydroxyl oxygen atoms (O13, O14, for **1** and O2, O6, O8 with ionizable protons) at the bottom of the reddish region have an electron rich area of **1** and **2**, potentially the most aggressive nucleophilic attack region. On the other hand, an electron poor region with a bluish color overlying a phenolic proton of the isolated compounds that may demonstrate an electrophilic behavior.

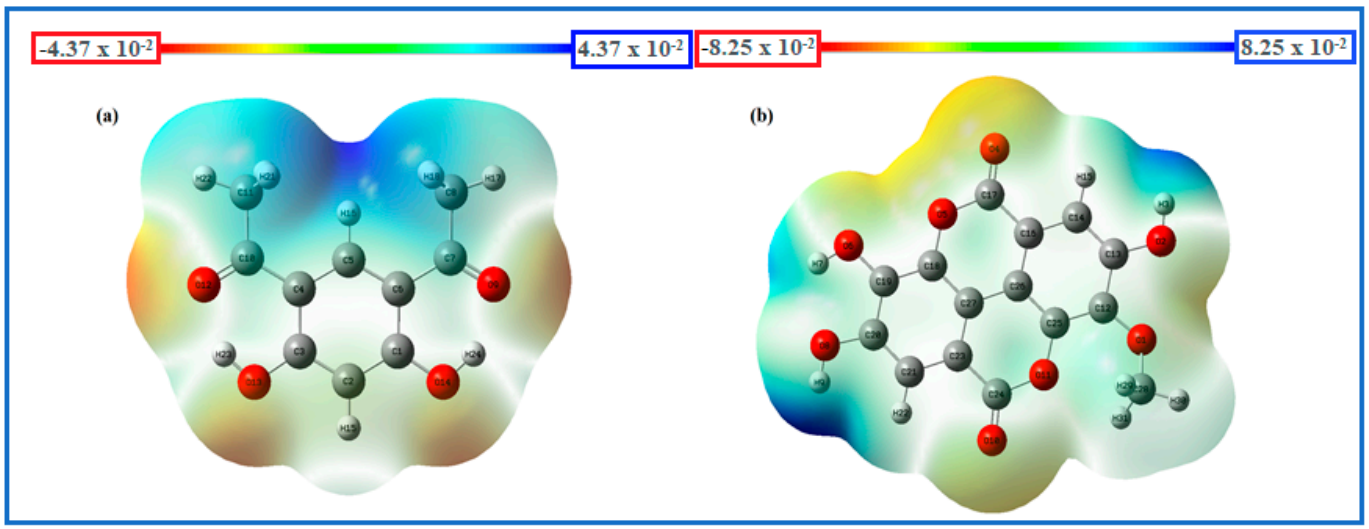

Figure 13. Molecular electrostatic potential surface of the compound **1** (**a**) and compound **2** (**b**) having regions rich and poor in electrons, as indicated by the colors red and blue.

3.5. UV-Visible Absorption Spectroscopy

UV–Visible absorption spectroscopy is one of the most fundamental and extensively used experimental techniques for evaluating the stability of ct-DNA and interactions with small ligand molecules [32]. Complex formation is usually accompanied by changes in the intensity and position of the absorption spectra [33]. The absorption maximum of compound **2** was centered on 273 nm. The maximum absorption band of compound **2** increased with increasing ct-DNA concentrations, accompanied by a shift of 11 nm, indicating the formation of a complex, as illustrated in Figure 14. Intercalation as a possible method of binding was discarded because it typically involves hypochromic or isosbestic points [32].

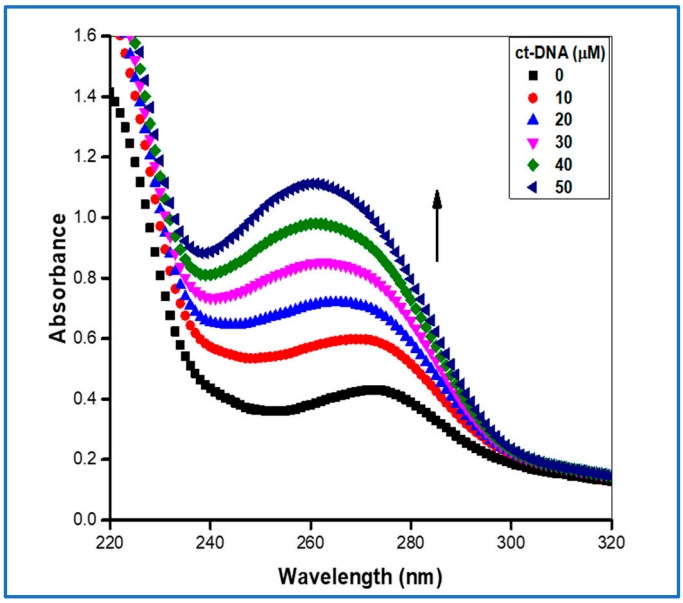

Figure 14. UV-absorption spectra of compound **2** (200 μM) in the presence of increasing concentration of ct-DNA (0–50 μM).

3.6. Steady-State Fluorescence

Fluorescence spectroscopy is undoubtedly one of the extensively used techniques for studying the interactions between small ligand molecules and DNA. High sensitivity, large linear concentration range, and selectivity are its advantages over other techniques. It can also be used to determine the pattern of compound binding to ct-DNA, as well as providing a variety of other details about the binding mode, strength, and number of binding sites in ct-DNA [34]. Compound **2** (200 μM) was excited at 273 nm and showed emission maxima at 367 nm. As demonstrated in Figure 15a, after consecutive additions of ct-DNA (0–125 μM), the intensity of compound **2** emission grew dramatically. This suggests the possibility of a complex-forming between them [35].

The Stern–Volmer equation [36] was used to estimate the enhancing mechanism involved in the process of binding between ct-DNA and compound **2**:

$$\frac{F_0}{F} = 1 - K_{SV}[Q] \quad (1)$$

$$K_q = \frac{K_{SV}}{\tau_0} \quad (2)$$

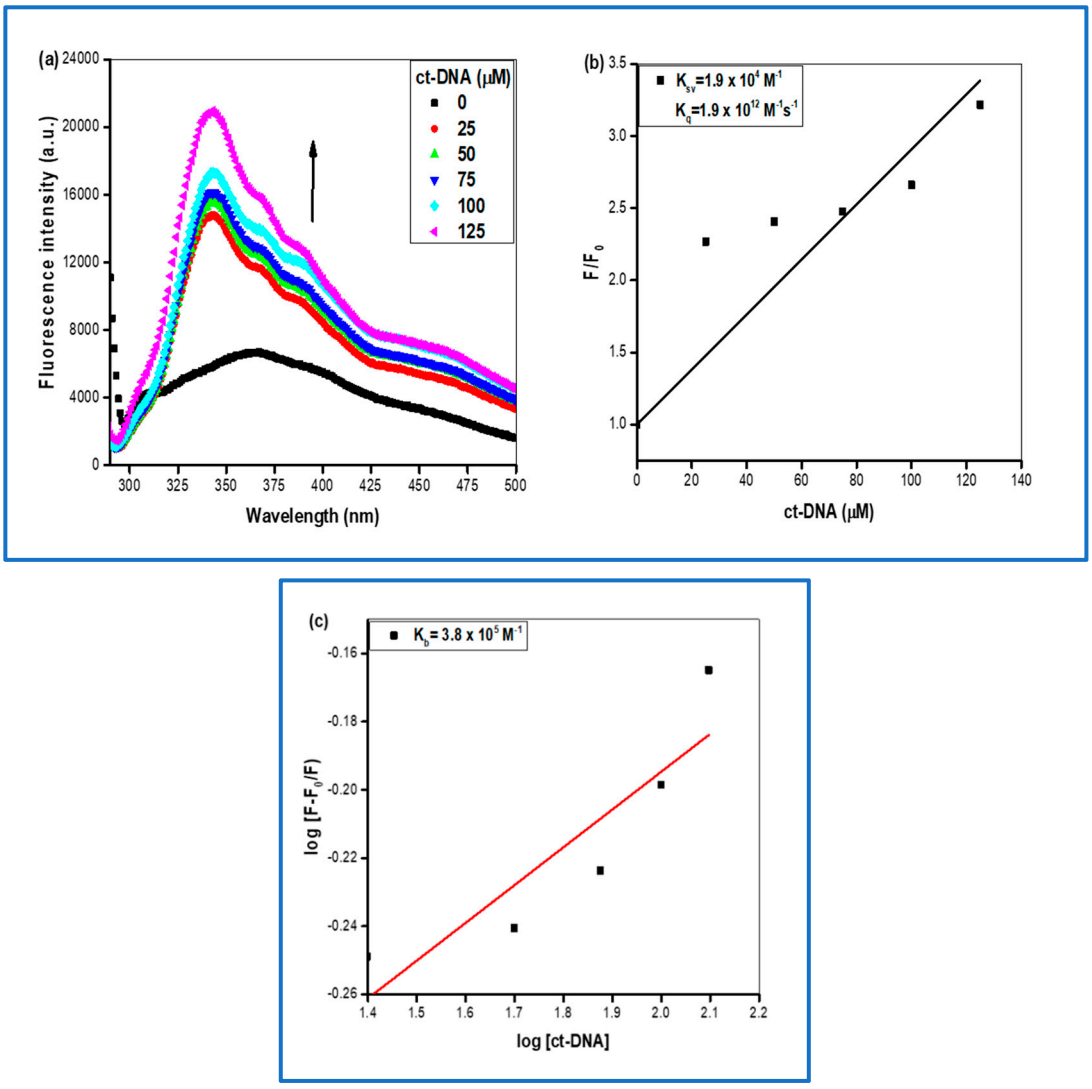

Figure 15. (**a**) Steady-state fluorescence spectra of compound **2** in the presence of increasing concentration of ct-DNA (0–125 μM). (**b**) Stern–Volmer plot of compound **2** complexed with ct-DNA to calculate K_{sv} and K_q. (**c**) Double logarithmic regression plot to calculate the binding constant (K_b).

The fluorescence intensities in the absence and presence of the quencher (ct-DNA) are represented by F_0 and F, respectively. The Stern–Volmer and the bimolecular enhancement constant are represented as K_{sv} and K_q, respectively. In the absence of the quencher, τ_0 is the fluorophore's average lifespan ($\tau_0 \approx 10^{-8}$ s) [37].

The binding process can be improved in one of two ways: by using a static or dynamic method. Static enhancement involves the formation of a ground state complex [38], while dynamic enhancement is largely associated with molecular diffusion.

The slope of the Stern–Volmer plot was used to obtain the K_{sv} value for compound **2** in Figure 15b. Equation (2) was used to calculate K_q using the value of K_{sv}, and both results are given in Table 1. Since the values of K_q exceeds the threshold value for dynamic enhancement 2×10^{11} M^{-1} s^{-1}, this suggests that rather than a dynamic approach [32], a static improvement technique should be used.

The binding constant (K_b) and interaction stoichiometry (n) were calculated using a double logarithmic regression plot. [32]:

$$\frac{\log(F - F_0)}{F} = \log K_b + n \log[DNA] \quad (3)$$

where K_b is the binding constant, and n is the number of binding sites of compound **2** in ct-DNA double helix. As shown in Figure 15c, the log $[(F - F_0)/F]$ vs. log [ct-DNA] plot's slope and intercept were calculated using Equation (3).

The value of n was discovered to be **1** when compound **2** was complexed with the ct-DNA interaction. As a result, compound **2** appears to have only one binding site in ct-DNA. The values of K_b for compound **2** were also determined and are presented in Table 4.

Table 4. From fluorescence measurements, values of Stern–Vomer, bimolecular enhancement constant, binding site number and binding constant achieved.

System	K_{sv} (M^{-1})	K_q (M^{-1} s^{-1})	n	K_b (M^{-1})
Compound 2	1.9×10^4	1.9×10^{12}	1.0	3.8×10^5

3.7. Circular Dichroism (CD)

Modifications in the secondary structure of ct-DNA [35] can be tracked using a useful tool, CD spectroscopy. By monitoring variations in the CD spectra of ct-DNA, binding mechanism of compound **2** to ct-DNA was also confirmed. Base stacking interactions cause a negative peak at 245 nm, and helicity induces a positive peak at 275 nm in the CD spectra of ct-DNA [39]. Intercalation entails the creation of an intercalation cavity in which the compounds can bind, as well as the development of new molecular connections that anchor the complex. Groove binding, on the other hand, does not require such conformational changes [37]. Intercalative molecules impair the ct-DNA interaction with base stacking, altering the negative peak as a result. Groove binders have no effect on the structure of ct-DNA; hence no such modifications are apparent [35]. No substantial change in the CD spectra of ct-DNA (200 µM) was found when compound **2** (200 µM) was added, indicating a groove binding interaction as seen in Figure 16.

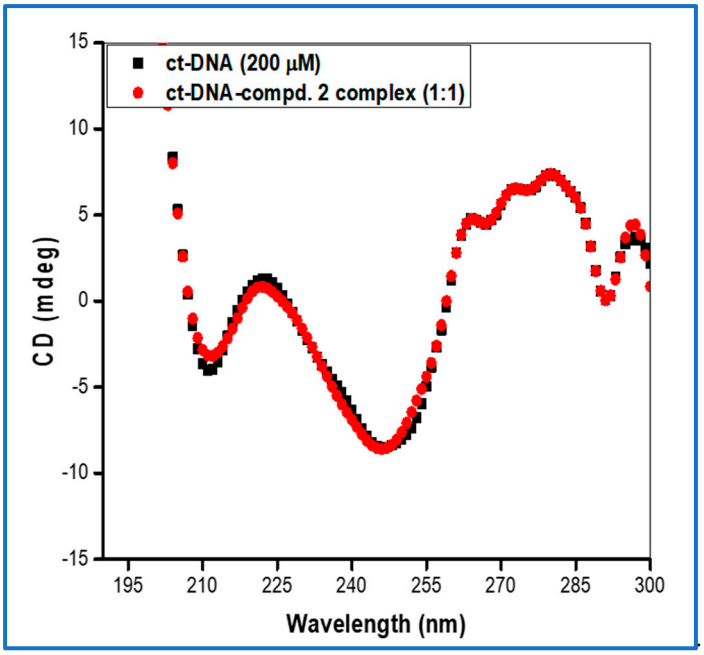

Figure 16. The CD spectra of ct-DNA (200 µM) alone and ct-DNA complexed with compound **2** in the molar ratio 1:1.

3.8. Docking

Molecular docking is a scoring function-based technique for predicting the interaction mechanism of a small molecule in a protein's binding site with the optimal orientation

and affinity [36]. It is also utilized to back up the findings obtained through spectroscopy and fluorescence-based experiments. Compound **2** was docked against B-DNA dodecamer d(CGCGAATTCGCG)2 (PDB ID:1BNA). Docking studies reveal that molecule **2** binds in DNA's minor grooves with a binding energy of -8.7 kcal mol^{-1}, as illustrated in Figure 17a. Compound **2** engages with DNA via two hydrogen bonds (DG) with bond lengths of 2.34 and 2.31 Å, which help compound **2** binding to DNA to be stable.

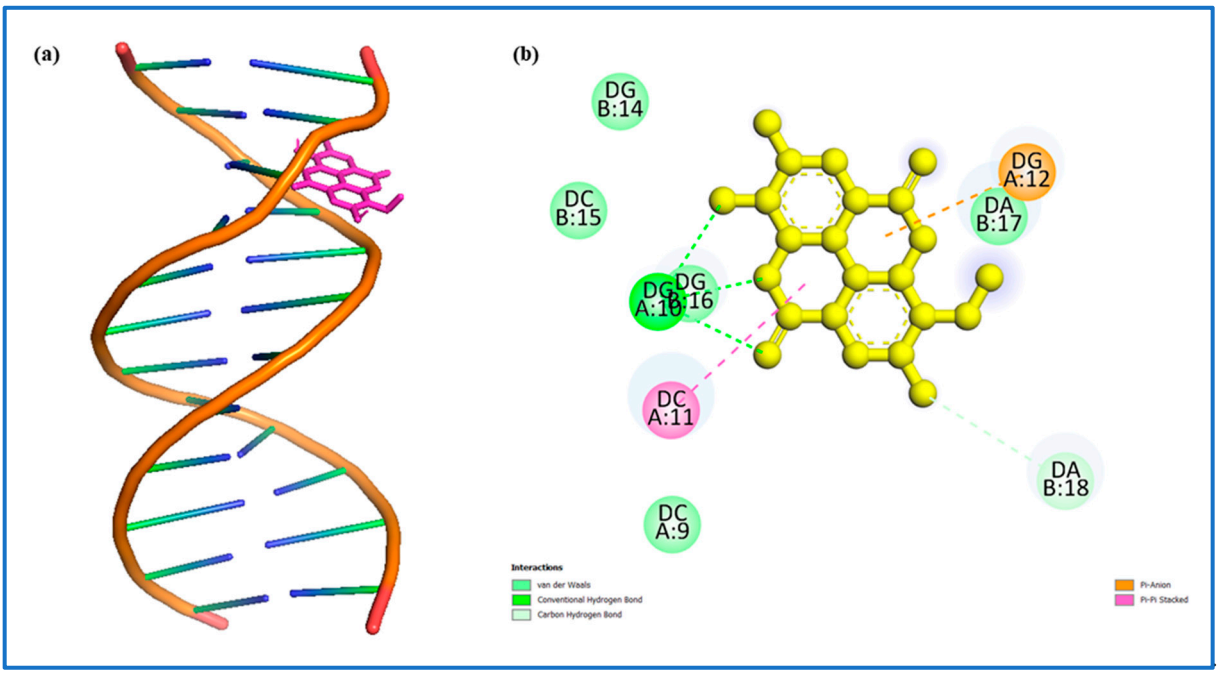

Figure 17. (**a**) Docked pose of compound **2** binding with DNA. (**b**) 2D-plot of the interaction.

3.9. Prediction of DRUG-likeness (Lipinski's Rule of Five)

The early preclinical analysis is greatly assisted by drug-likeness filters, which help to avoid costly late-stage preclinical and clinical failure. The drug-likeness properties of the compound were studied by using the Lipinski rule of five. It is a preliminary criterion for assessing its structural resemblance to an ideal drug [40,41]. Compound **2** under investigation follows all of the parameters under the 'Rule of Five', with no violations calculated via an online server, and revealed a higher tendency of compound **2** towards drug-likeness.

3.10. Bioavailability Score and Synthetic Accessibility

Bioavailability difficulties linked with a compound might slow down the therapeutic development process. For a compound to exhibit its pharmacological action on the body and be effective as a drug, it must be available in sufficient concentration in the systemic circulation for a specific duration [42]. To avoid unfavorable outcomes in the future, it is important to evaluate the compound's bioavailability early in the therapeutic development process. Therefore, the bioavailability of compound **2** was determined with the SwissADME web tool [25]. This web tool determines the bioavailability of the compound based on the molecular properties and lipophilicity by applying through different principles such as Lipinski's rule of five, Ghose filter, Veber filter, Egan filter, and Muegge filter. The bioavailability score of the isolated molecule (**2**) was found to be 0.55, which implies that there is a 55% probability of being bioavailable.

To identify a promising lead compound during the process of virtual screening, it is preferred to filter out a non-toxic, physiologically active compound with good bioavailability. In addition, the degree of complexity in synthesizing a compound is a factor that should be considered while choosing the most promising compound [43]. The SwissADME

web tool was also used to measure the degree of difficulty in synthesizing compound **2**. Synthetic accessibility is a fingerprint-based computational approach for determining how difficult it is to synthesize a compound. A synthetic accessibility value of **1** indicates that the compound is relatively easy to synthesize, while the synthetic accessibility score of 10 indicates that the compound is extremely difficult to synthesize. Compound **2** was found to have a synthetic accessibility score of 3.21, indicating that it will be easier to synthesize.

4. Conclusions

In this paper, two compounds named as 4,6-Diacetylresorcinol (**1**) and 3-*O*-methylellagic acid dihydrate (**2**), were isolated from the leaves of *Bixa orellana*, although they have previously been described in other sources and synthetic methods. The density functional theory was performed to calculate the electronic structure in support of experimental results. In order to compare spectral data of experimental studies (IR, ^{1}H, ^{13}C NMR, UV and parameters obtained from single X-ray diffraction) with theoretical exploration, B3LYP/6-311G(d,p) was used to optimize geometric parameters in the gas phase, while the NMR and UV-Vis studies of the optimized compounds performed by GIAO-B3LYP/6-311G(d,p) and PCM-B3LYP/6-311G(d,p) level theory were applied in the respected solution phase, respectively. The theoretical and experimental observations of the isolated compound were in good agreement with the establishment of the chemical structures of both molecules. In addition, energy of HOMO, LUMO and its gap including the MEP for compound **1** was calculated that were found to be −6.82 eV, 2.09 eV, and 4.73 eV, respectively, while compound **2** was −6.19 eV, 2.22 eV, and 3.97 eV respectively, these energy gaps of 4,6-diacetylresorcinol and 3-*O*-methylellagic acid dihydrate, including MEP, were an indicator of successful electronic transition and their high reactivity including chemical nature of molecules having various functional groups for the future course of the drug development response. Using multiple biophysical and in-silico approaches, the present study adds to our understanding of the interaction of molecule 2 with ct-DNA. UV-absorption and fluorescence spectroscopy confirm the complex formation between compound **2** with ct-DNA. The value of binding constant (K_b) obtained was in the order of 10^5 M^{-1}. Circular dichroism results of ct-DNA complexed with compound **2** suggest groove binding mode. Furthermore, *in silico* studies also suggest that the binding of compound **2** occurs in the minor groove of the DNA. Lipinski's rule of five revealed a higher tendency of compound **2** towards drug-likeness. The bioavailability score and synthetic accessibility score of compound (**2**) was found to be 0.55 and 3.21, suggesting that compound **2** could serve as effective therapeutic candidate.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/cryst12030380/s1. Table S1. The experimental parameters related to the monocrystalline X-ray examination of **1** and **2.**

Author Contributions: Conceptualization, M.P., M.A. (Mohammad Azeem), A.M.M. and M.A. (Mahboob Alam); methodology, M.A. (Mahboob Alam), V.H.R. and A.A.; software, K.M., S.S., S.I.A.-R. and M.T.; validation, M.A. (Mahboob Alam) and A.A.; formal analysis, M.A. (Mohammad Azeem) and V.H.R.; investigation, M.A. (Mohammad Azam) and S.I.A.-R.; resources, M.A. (Mahboob Alam) and A.A.; writing—original draft preparation, M.A. (Mohammad Azam), S.I.A.-R., A.A. and K.M.; writing—review and editing, S.S., M.T. and M.P.; visualization, M.A. (Mahboob Alam); supervision, M.P., M.A. (Mohammad Azam) and K.M.; writing—original draft preparation, validation, K.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors acknowledge the financial support through Researchers Supporting Project number (RSP-2021/147), King Saud University, Riyadh, Saudi Arabia. M. Azeem thanks Department of Chemistry, A.M.U, Aligarh, for providing the necessary research facilities. UGC is acknowledged for financial support. The authors acknowledge the Laboratory for Advanced Computing at University of Coimbra (http://www.uc.pt/lca) for providing computing resources that have contributed to the research results reported within this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Silva, S.N.S.; Amaral, C.L.F.; Reboucas, T.N.H. Adoption of conservation practices on farm and selection of varieties by producers of annatto in the city of Vitoria da Conquista-BA. *Rev. Bras. Agroecologica* **2010**, *5*, 106.
2. Correa, M.P. *Dicionario das Plantas ´Uteis do Brasil e das ´ Exoticas Cultivadas*; Ministerio da Agricultura/IBDF: Rio ´de Janeiro, Brasil, 1978; Volume 4.
3. Villar, R.; Calleja, J.M.; Morales, C.; Caceres, A. Screening of 17 Guatemalan medicinal plants for platelet antiaggregant activity. *Phytother Res.* **1997**, *11*, 441. [CrossRef]
4. Aher, A.A.; Bairagi, S.M. Formulation and evaluation of herbal lipstick from colour pigments of *Bixa Orellana* (Bixaceae) seeds. *Int. J. Pharm. Biol. Sci.* **2012**, *4*, 357.
5. Kang, E.J.; Campbell, R.E.; Bastian, E.; Drake, M.A. Invited review: Annatto usage and bleaching in dairy foods. *J. Dairy Sci.* **2010**, *93*, 3891. [CrossRef] [PubMed]
6. Venugopalan, A.; Giridhar, P.; Ravishankar, G.A. Food, ethanobotanical and diversified applications of *Bixa orellana* L.: A scope for its improvement through biotechnological mediation. *Ind. J. Fundament. Appl. Life Sci.* **2011**, *1*, 9.
7. Scotter, M. The chemistry and analysis of annatto food colouring: A review. *Food Addit. Contam.* **2009**, *26*, 1123. [CrossRef]
8. Mercadante, A.Z.; Steck, A.; Pfander, H. Three minor carotenoids from annatto (*Bixa orellana*) seeds. *Phytochemistry* **1999**, *52*, 135–159. [CrossRef]
9. Pino, J.A.; Correa, M.T. Chemical Composition of the Essential Oil from Annatto (*Bixa orellana* L.) Seeds. *J. Essent. Oil Res.* **2003**, *15*, 66. [CrossRef]
10. Yong, Y.K.; Zakaria, Z.A.; Kadir, A.A.; Somchit, M.N.; Lian, G.E.C.; Ahmad, Z. Chemical constituents and antihistamine activity of *Bixa orellana* leaf extract. *BMC Complement. Altern. Med.* **2013**, *13*, 32. [CrossRef]
11. Parveen, M.; Malla, A.M.; Ali, A.; Nami, S.A.A.; Silva, P.S.P.; Silva, M.R. Isolation, Characterization, Bioassay and X-ray Crystallographic Study of Phytoconstituents from *Bixa orellana* Leaves. *Chem. Nat. Compd.* **2015**, *51*, 62. [CrossRef]
12. Khan, M.S.Y.; Sharma, S.; Husain, A. Synthesis and antibacterial evaluation of new flavonoid derivatives from 4,6-diacetyl resorcinol. *Sci. Pharm.* **2002**, *70*, 287. [CrossRef]
13. SADABS. *Area-Detector Absorption Correction*; Siemens Industrial Automation, Inc.: Madison, WI, USA, 1996.
14. Sheldrick, G.M. *SHELXL-97: Program for Crystal Structure Refinement*; University of Göttingen: Göttingen, Germany, 1997.
15. Platon, S.A. An Integrated Tool for the Analysis of the Results of a Single Crystal Structure Determination. *Acta Crystallogr. Sect. A* **1990**, *46*, C34.
16. Platon, S.A. *A Multipurpose Crystallographic Tool*; Utrecht University: Utrecht, The Netherlands, 1998.
17. Farrugia, L.J. WinGX and ORTEP for Windows: An update. *J. Appl. Cryst.* **2012**, *45*, 849. [CrossRef]
18. Grime, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104. [CrossRef] [PubMed]
19. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Montgomery, J.A., Jr.; Vreven, T.; Kudin, K.N.; Burant, J.C.; et al. *Gaussian 03, Revision E.01*; Gaussian Inc.: Wallingford, CT, USA, 2004.
20. Schmidt, J.R.; Polik, W.F. *WebMO Enterprise, 18.1.001*; WebMO LLC: Holland, MI, USA, 2016.
21. Dennington, R.; Keith, T.; Millam, J. *Gauss View Version 5*; Semichem Inc.: Shawnee Mission, AR, USA, 2009.
22. Chemcraft-Graphical Software for Visualization of Quantum Chemistry Computations. Available online: https://www.chemcraftprog.com (accessed on 1 November 2021).
23. O'boyle, N.M.; Tenderholt, A.L. cclib: A library for package-independent computational chemistry algorithms. *J. Comput. Chem.* **2008**, *29*, 839. [CrossRef] [PubMed]
24. Jayaram, B.; Singh, T.; Mukherjee, G.; Mathur, A.; Shekhar, S.; Shekhar, V. Sanjeevini: A freely accessible webserver for target directed lead molecule discovery. *BMC Bioinform.* **2012**, *13*, 1. [CrossRef] [PubMed]
25. 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. [CrossRef]
26. Kokila, M.K.; Nirmala, K.A.; Shamala, P.N. Structure of 4,6-di acetylresorcinol. *Acta. Cryst.* **1992**, *C48*, 1133.
27. Rossi, M.; Erlebacher, J.; Zacharias, D.E.; Carrell, H.L.; Iannucci, B. The crystal and molecular structure of ellagic acid dihydrate: A dietary anti-cancer agent. *Carcinogenesis* **1991**, *12*, 2227. [CrossRef]
28. Socrates, G. *Infrared Characteristic Group Frequencies*, 3rd ed.; Wiley Interscience Publications: New York, NY, USA, 1980.
29. Kosar, B.; Albayrak, C. Spectroscopic investigations and quantum chemical computational study of (E)-4-methoxy-2-[(p-tolylimino)methyl]phenol. *Spectrochim. Acta A* **2011**, *78*, 160. [CrossRef]

30. Alam, M.J.; Ahmad, S. Quantum chemical and spectroscopic investigations of 3-methyladenine. *Spectrochim. Acta A* **2014**, *128*, 653. [CrossRef] [PubMed]
31. Tabatchnik, A.; Blot, V.; Pipelier, M.; Dubreuil, D.; Renault, E.; Le Questel, J.-Y. Theoretical Study of the Structures and Hydrogen-Bond Properties of New Alternated Heterocyclic Compounds. *J. Phys. Chem. A.* **2010**, *114*, 6413. [CrossRef] [PubMed]
32. Afrin, S.; Rahman, Y.; Sarwar, T.; Husain, M.A.; Ali, A.; Tabish, M. Molecular spectroscopic and thermodynamic studies on the interaction of anti-platelet drug ticlopidine with calf thymus DNA. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2017**, *186*, 66. [CrossRef] [PubMed]
33. Parveen, M.; Aslam, A.; Siddiqui, S.; Tabish, M.; Alam, M. Structure elucidation, DNA binding and molecular docking studies of natural compounds isolated from Crateva religiosa leaves. *J. Mol. Struct.* **2021**, *1251*, 131976. [CrossRef]
34. Sirajuddin, M.; Ali, S.; Badshah, A. Drug-DNA interactions and their study by UV-Visible, fluorescence spectroscopies and cyclic voltametry. *J. Photochem. Photobiol. B Biol.* **2013**, *124*, 1–19. [CrossRef]
35. Hussain, I.; Fatima, S.; Siddiqui, S.; Ahmed, S.; Tabish, M. Exploring the binding mechanism of β-resorcylic acid with calf thymus DNA: Insights from multi-spectroscopic, thermodynamic and bioinformatics approaches. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2021**, *260*, 119952. [CrossRef]
36. Siddiqui, S.; Ameen, F.; Jahan, I.; Nayeem, S.M.; Tabish, M. A comprehensive spectroscopic and computational investigation on the binding of the anti-asthmatic drug triamcinolone with serum albumin. *New J. Chem.* **2019**, *43*, 4137. [CrossRef]
37. Siddiqui, S.; Mujeeb, A.; Ameen, F.; Ishqi, H.M.; Rehman, S.U.; Tabish, M. Investigating the mechanism of binding of nalidixic acid with deoxyribonucleic acid and serum albumin: A biophysical and molecular docking approaches. *J. Biomol. Struct. Dyn.* **2021**, *39*, 570. [CrossRef]
38. Siddiqui, S.; Ameen, F.; ur Rehman, S.; Sarwar, T.; Tabish, M. Studying the interaction of drug/ligand with serum albumin. *J. Mol. Liq.* **2021**, *336*, 116200. [CrossRef]
39. Ameen, F.; Siddiqui, S.; Jahan, I.; Nayeem, S.M.; ur Rehman, S.; Tabish, M. A detailed insight into the interaction of memantine with bovine serum albumin: A spectroscopic and computational approach. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2022**, *265*, 120391. [CrossRef]
40. Enmozhi, S.K.; Raja, K.; Sebastine, I.; Joseph, J. Andrographolide as a potential inhibitor of SARS-CoV-2 main protease: An in silico approach. *J. Biomol. Struct. Dyn.* **2021**, *39*, 3092. [CrossRef] [PubMed]
41. Benet, L.Z.; Hosey, C.M.; Ursu, O.; Oprea, T.I. BDDCS, the Rule of 5 and drugability. *Adv. Drug Deliv. Rev.* **2016**, *101*, 89. [CrossRef] [PubMed]
42. Pathak, K.; Raghuvanshi, S. Oral bioavailability: Issues and solutions via nanoformulations. *Clin. Pharmacokinet.* **2015**, *54*, 325. [CrossRef] [PubMed]
43. Zothantluanga, J.H.; Gogoi, N.; Shakya, A.; Chetia, D.; Lalthanzara, H. Computational guided identification of potential leads from *Acacia pennata* (L.) Willd. as inhibitors for cellular entry and viral replication of SARS-CoV-2. *Future J. Pharm. Sci.* **2021**, *7*, 201. [CrossRef]