Supporting Information

Design and Evaluation of Sulfadiazine Derivatives as Potent Dual Inhibitors of EGFR^{WT} and EGFR^{T790M}: Integrating Biological, Molecular Docking, and ADMET Analysis

Hany M. Abd El-Lateef[a,b*#], Hend A. A. Ezelarab[c*#], Ali M. Ali[b], Azhaar T. Alsaggaf[d], Wael A. Mahdi[e], Sultan Alshehri[e], Mohamed A. El Hamd*[f,g], Moustafa O. Aboelez[*].

[a] Department of Chemistry, College of Science, King Faisal University, Al-Ahsa 31982, Saudi Arabia

[b] Department of Chemistry, Faculty of Science, Sohag 82524, Egypt.

[c] Department of Medicinal Chemistry, Faculty of Pharmacy, Minia University, 61519-Minia, Egypt

[d] Department of Chemistry, Taibah University, Madinah, Saudi Arabia.

[e] Department of Pharmaceutics, College of Pharmacy, King Saud University, Riyadh 11451, Saudi Arabia.

[f] Department of Pharmaceutical Chemistry, College of Pharmacy, Shaqra University, Shaqra 11961, Saudi Arabia.

[g] Department of Pharmaceutical Analytical Chemistry, Faculty of Pharmacy, South Valley University, Qena 83523, Egypt.

[h] Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Sohag University, Sohag, Egypt.

Correspondence to:

[*]Prof. Dr Hany M. Abd El-Lateef, hmahmed@kfu.edu.sa, Dr Moustafa O. Aboelez, drmoustafaaboelez@yahoo.com, Dr Hend A. A. Ezelarab, hend_aly@minia.edu.eg, Dr Mohamed A. El Hamd, aboelhamdmohamed@su.edu.sa.

[#] Equal contributed

1. Experimental part

1.1. Chemistry

1.1.1. General Information

All commercially available reagents were purchased from Merck, Aldrich and Fluka and were used without further purification. All reactions were monitored by thin layer chromatography (TLC) using precoated plates of silica gel G/UV-254 of 0.25 mm thickness (Merck 60F254) using UV light (254 nm/365 nm) for visualization. Melting points were detected with a Kofler melting points apparatus and uncorrected. Infrared spectra were recorded with a FT-IR-ALPHBROKER-Platinum-ATR spectrometer and are given as cm^{-1} using the attenuated total reflection (ATR) method. ^1H NMR and ^{13}C NMR spectra for all new compounds were recorded in DMSO-d_6 on a Bruker Bio Spin AG spectrometer at 400 MHz and 100 MHz, respectively. For ^1H NMR, chemical shifts (δ) were given in parts per million (ppm) with reference to tetramethylsilane (TMS) as an internal standard ($\delta=0$); coupling constants (J) were given in hertz (Hz) and data are reported as follows: chemical shift, integration, multiplicity (s=singlet, d=doublet, t= triplet, m=multiplet). For ^{13}C NMR, TMS ($\delta=0$) or DMSO ($\delta=39.51$) was used as internal standard and spectra were obtained with complete proton decoupling. Elemental analyses were obtained on a Perkin-Elmer CHN-analyzer model.

1.2. Biological Evaluation

1.2.1. *In vitro* anti-proliferative activities

In vitro cytotoxicity activities of the target compounds were investigated quantitatively against three cancer cell lines, namely human epidermoid carcinoma cells (A431) and non-small cell lung cancer cells (A549 and H1975), applying the MTT method. Commercially available drugs (erlotinib, gefitinib, and osimertinib) are the standard references. The investigated cell lines were supplied by the American Type Culture Collection (ATCC) (Rockville, MD). The MTT assay is a common test for evaluating tumor growth and assessing the cytotoxicity of drug candidates and other toxic substances. In conclusion, yellow MTT is reduced through mitochondrial dehydrogenases in living tissue to generate purple formazan. A proper solvent dissolves the non-soluble purple formazan product into a colored solution. The absorbance of this purple formazan solution was evaluated at a specific wavelength. Once the portion of purple formazan released by cells allowed to be treated with an agent is compared to that of untreated control cells, the agent's efficiency in provoking cell death can be calculated by generating a dose-response relationship curve. Human cancer cell lines were implanted in 96-well plates at a dose of 3-8 x 10^3 cells per well. The

wells were then incubated at 37 °C for 12 hours in a 5% CO_2 incubator. To determine the DMSO level, each well's culture medium was replaced with 0.1 ml of new medium containing graded quantities of the target compounds. Following a two-day hatching time, the cells were matured in 100 µl MTT solution (5 µg ml^{-1}) for four hours in each well. After dissolving MTT-formazan crystals in 100 µl DMSO, the absorption intensity was determined photometrically at 490 nm using an automated ELISA reader system (TECAN, CHE). The IC_{50} values then were determined using nonlinear regression fitting models (Graph Pad, Prism Version 5) (n = 3, duplicate trials, reported as mean SD).

The investigated cell lines were incubated in RPMI-1640 media with 10% inactivated FBS, 50 µg/mL of gentamycin, 50 units/mL of penicillin, and 1 mmol/L of L-glutamine. The cultures were cultivated 2-3 times per week and kept at room temperature in a humidified environment with 5% carbon dioxide at densities of 3–8 × 10^3 cells/well on 96-well plates. After filling the fresh medium (0.1 mL) with the graduated concentrations of the target degraders very well, the culture medium was incubated for two days. The cultured cells on each plate received 100 µL of MTT solution (5.0 µg mL^{-1}) and were left for four hours. Employing an automated ELISA reader system (TECAN, CHE), the MTT-formazan crystals were dissolved in 100 µL of DMSO, and the absorbance of each collected well was detected at 490 nm. The formula employed to estimate surviving cells and inhibitory cells was as follows:

$$\% \, Surviving \, cell = \frac{Mean \, optical \, density \, (OD) of \, tested \, compound}{Mean \, OD \, of \, negative \, control} \times 100$$

$$\% \, Inhibiting \, cells = 100 - Surviving \, cells$$

Moreover, nonlinear regression fitting models were employed to compute the IC_{50} values (GraphPad, Prism 5). The obtained numerical data were calculated by using the average of three individual duplicate experiments and presented as the mean ± standard deviations (SD).

1.2.2. EGFR^{WT} and EGFR^{T790M} kinase inhibitory assay

When significant IC_{50} values versus target cell lines were identified, the inhibitory activity of derivatives versus both EGFR^{WT} and EGFR^{T790M} was studied more. In this study, the HTRF test with EGFR^{WT} and EGFR^{T790M} (Sigma) was performed. For the first 5 minutes, the compounds (**1-7**) were incubated in the enzymatic buffer with EGFR^{WT} and/or EGFR^{T790M} and their substrates. To start the enzymatic activity, 1.65 M ATP was allowed to react. The reaction runs for half an hour at 210 K. When EDTA-containing testing reagents were introduced, the procedure was halted. After a one-hour detection period, the IC_{50} values were computed by GraphPad Prism 5.0 program. Each concentration was evaluated using three different ways [37,38].

1.3. *In silico* studies

1.3.1. Molecular Docking Study

Compounds **8**, **12**, and **14** were designed and docked in the ATP active site of both EGFR^{WT} and EGFR^{T790M} protein kinases (**PDB ID: 4HJO** and **3W2O**, respectively) (downloaded from the PDB website (Protein Data Bank)) using the Molecular Operating Environment (MOE 2019.) program as reported in the literature. The root mean square deviation (RMSD) values of erlotinib and TAK-285 secondly docked analogs and co-localized conformers, respectively, were 1.4 and 1.85 Å, respectively, illustrating the rationality of this docking protocol (**Fig. 1S** and **5S**, respectively, in supporting information). Additionally, gefitinib and osimertinib as reversible first- and irreversible third-generation EGFR inhibitors were docked versus wild and mutant EGFR, respectively, to elucidate their potential dual EGFR inhibitory activities. Erlotinib, the co-localized ligand, was principally dipped into its corresponding co-crystalized protein model (**PDB code: 4HJO**) to evaluate if MOE could replicate the native ligand superimposition to the wild EGFR protein active site (**Fig. 1S** and **2S**, and **Table 3**). Besides, the native and re-imposed derivatives are docked in the same way with the key amino acids. The energy score of the bonded re-docked drug (ΔG) was −6.83 Kcal/mol with a good RMSD value of 1.4 Å.

Regarding the co-crystallized ligand TAK-285, the co-localized ligand was primarily dipped into its corresponding EGFR^{T790M} adenine binding pocket of the co-crystal protein model (**PDB code:3W2O**) to assess if MOE could replicate the native ligand superimposition to the EGFR^{T790M} protein active site (**Fig. 5S** and **Table 4**). Moreover, the native and re-docked derivatives are docked in the same way with the key AAs. The bonded re-docked TAK-285 binding score (ΔG) equals to -7.35 Kcal/mol with a good RMSD value of 3.01 Å. Also, osimertinib re-docked in the same manner as the re-docked ligand TAK-285 and the other reference drug gefitinib,

1.3.2. ADMET Estimation

The physicochemical properties, lipophilicity (logP value), hydrophilicity calculations, pharmacokinetic characters such as GI absorption and CYP enzyme inhibition, drug-likeness and medicinal chemistry parameters such as the lead likeness of the most potent compounds **8** and **12** compared to the references Erlotinib and Gefitinib; were performed *via* using the SwissADME online website (http://www.swissadme.ch/) [57]. Toxicity parameters of these potent compounds were investigated through the pkCSM-pharmacokinetics website (http://biosig.unimelb.edu.au/pkcsm/) [57-61].
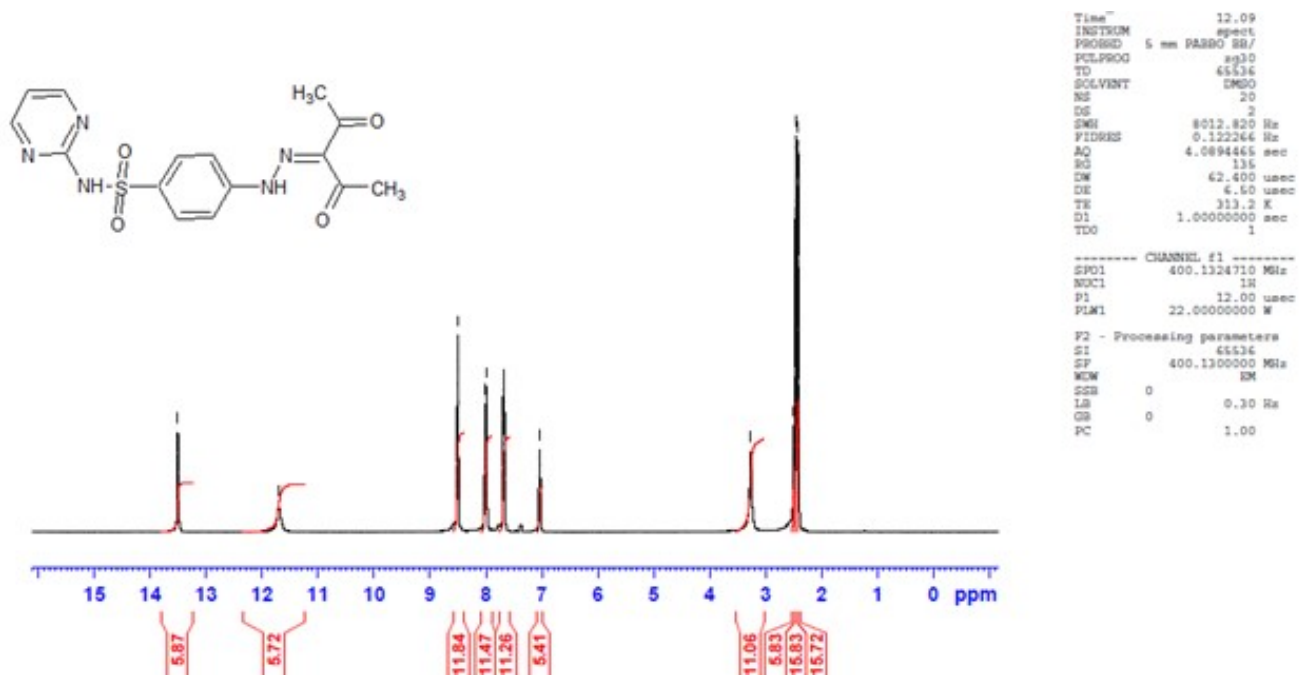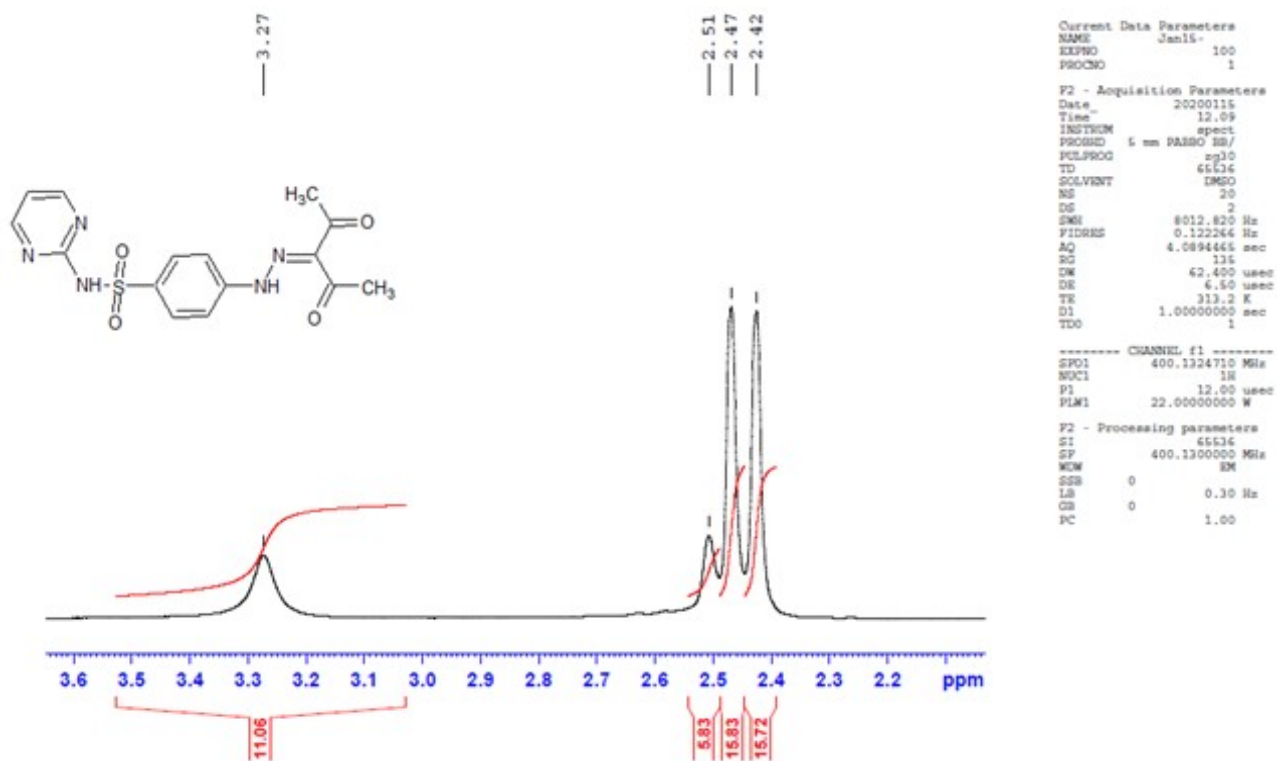
Figure s1: ^{1}H NMR spectrum of compound **6**
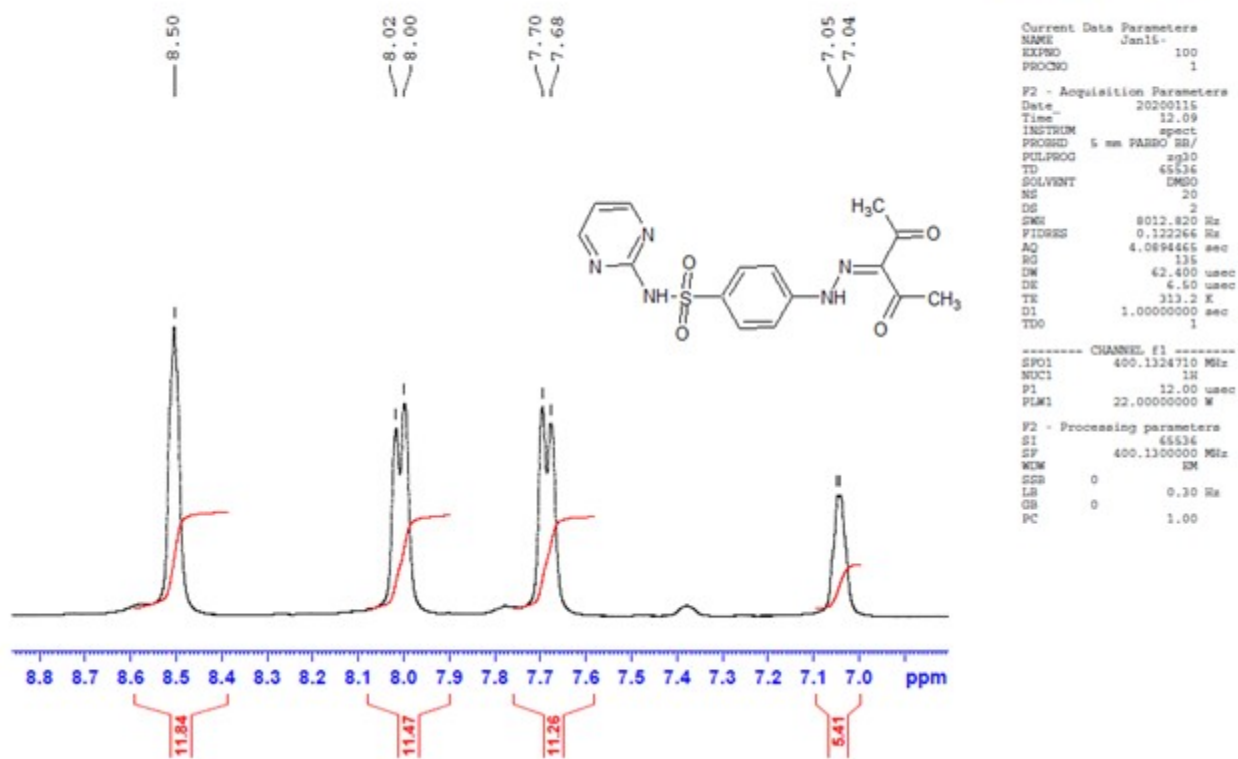
Figure s1: ^{1}H NMR spectrum of compound **6**

Figure s1: ^{1}H NMR spectrum of compound **6**
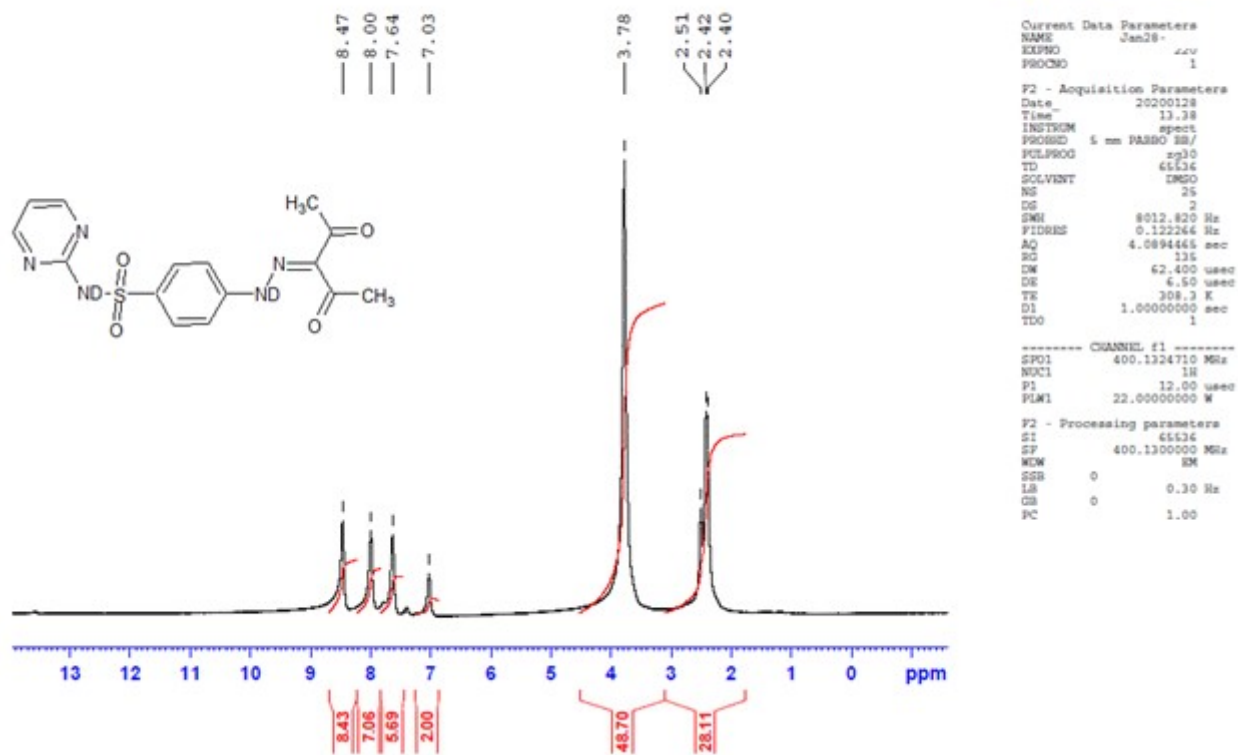
Figure s1: ^1H NMR (D2O) spectrum of compound **6**

Figure s2: C^{13} NMR spectrum of compound **6**

BRUKER

Current Data Parameters
NAME   Feb09-
EXPNO   41v
PROCNO   1

F2 - Acquisition Parameters
Date_   20200209
Time   13.01
INSTRUM   spect
PROBHD   5 mm PABBO BB/
PULPROG   zg30
TD   65536
SOLVENT   DMSO
NS   20
DS   2
SWH   8012.820 Hz
FIDRES   0.122266 Hz
AQ   4.0894465 sec
RG   68.22
DW   62.400 usec
DE   6.50 usec
TE   308.2 K
D1   1.00000000 sec
TD0   1

-------- CHANNEL f1 --------
SFO1   400.1324710 MHz
NUC1   1H
P1   12.00 usec
PLW1   22.00000000 W

F2 - Processing parameters
SI   65536
SF   400.1300000 MHz
WDW   EM
SSB   0
LB   0.30 Hz
GB   0
PC   1.00

Figure s3: ^{1}H NMR spectrum of compound **7**

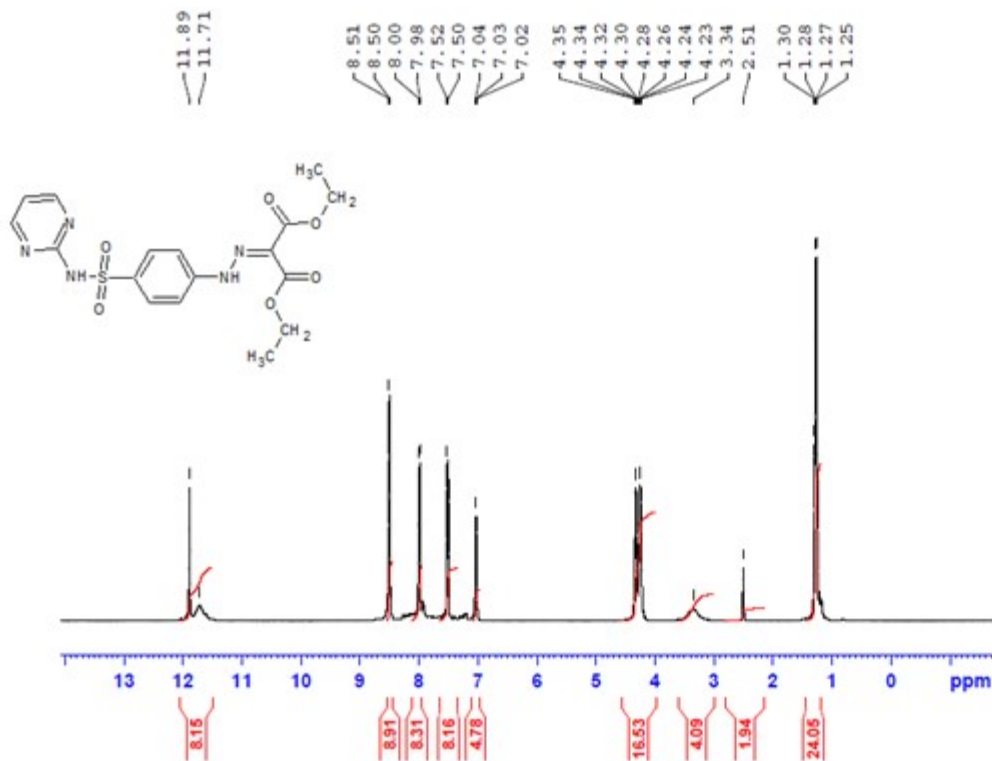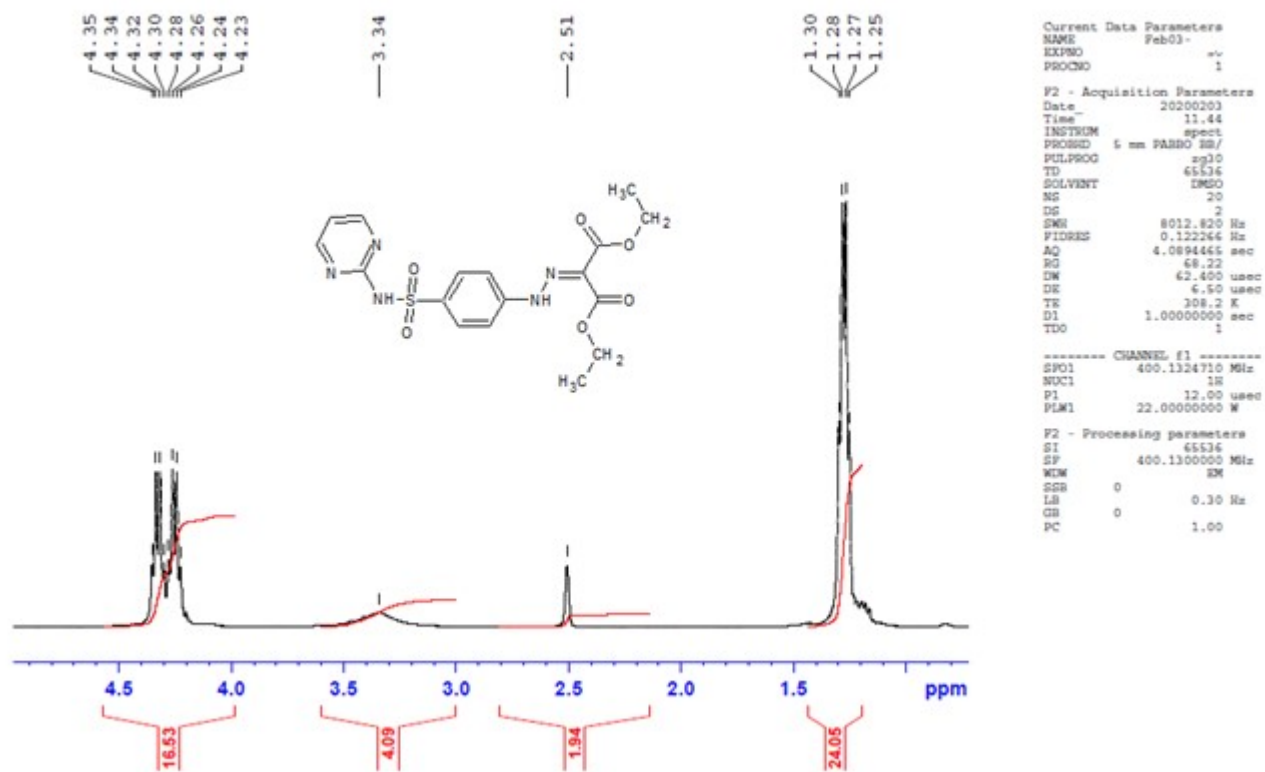**Figure s3:** ^{1}H NMR spectrum of compound **7**

BRUKER

Figure s3: ^{1}H NMR spectrum of compound **7**

Figure s4: ^{13}C NMR Spectrum of compound 7
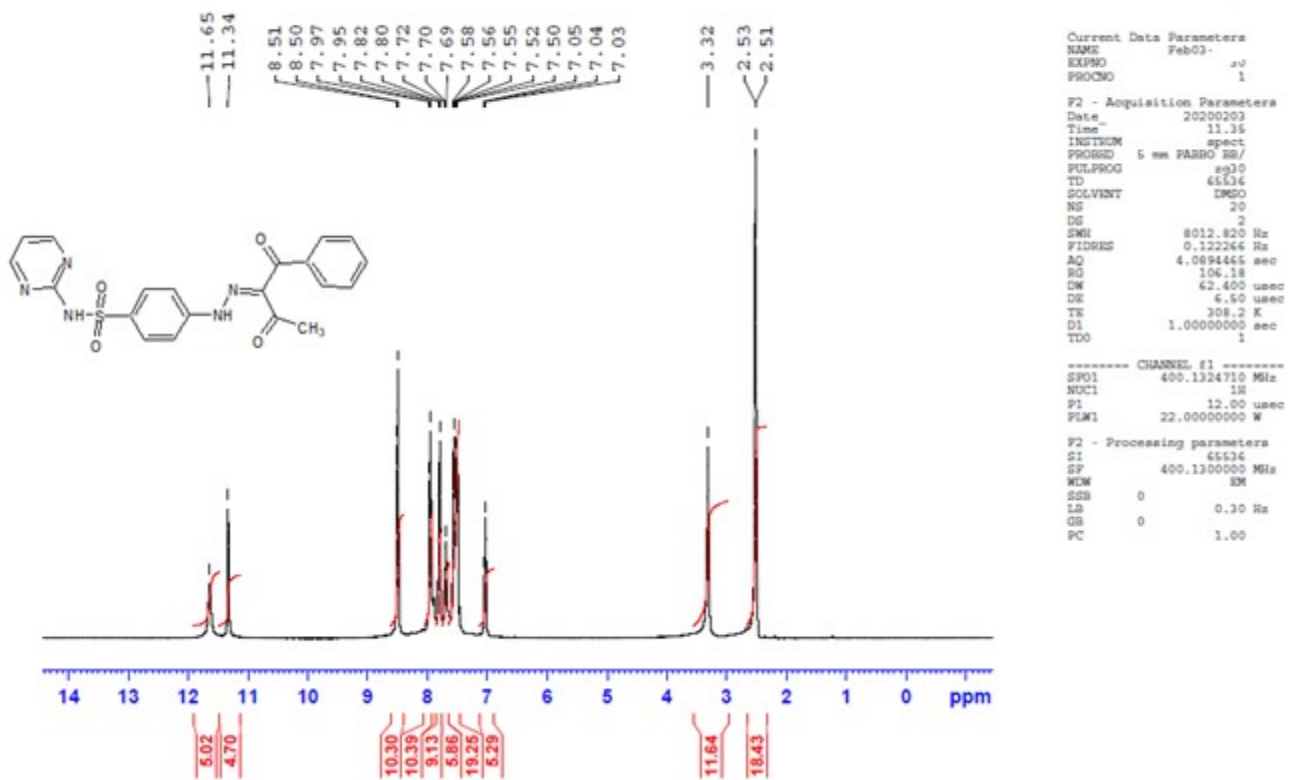
Figure s5: ^{1}H NMR spectrum of compound **8**

Figure s5: ^{1}H NMR spectrum of compound **8**

Figure s6:^{13}**C NMR Spectrum of compound 8**

BRUKER

Figure s7: ^{1}H NMR spectrum of compound **9**

BRUKER
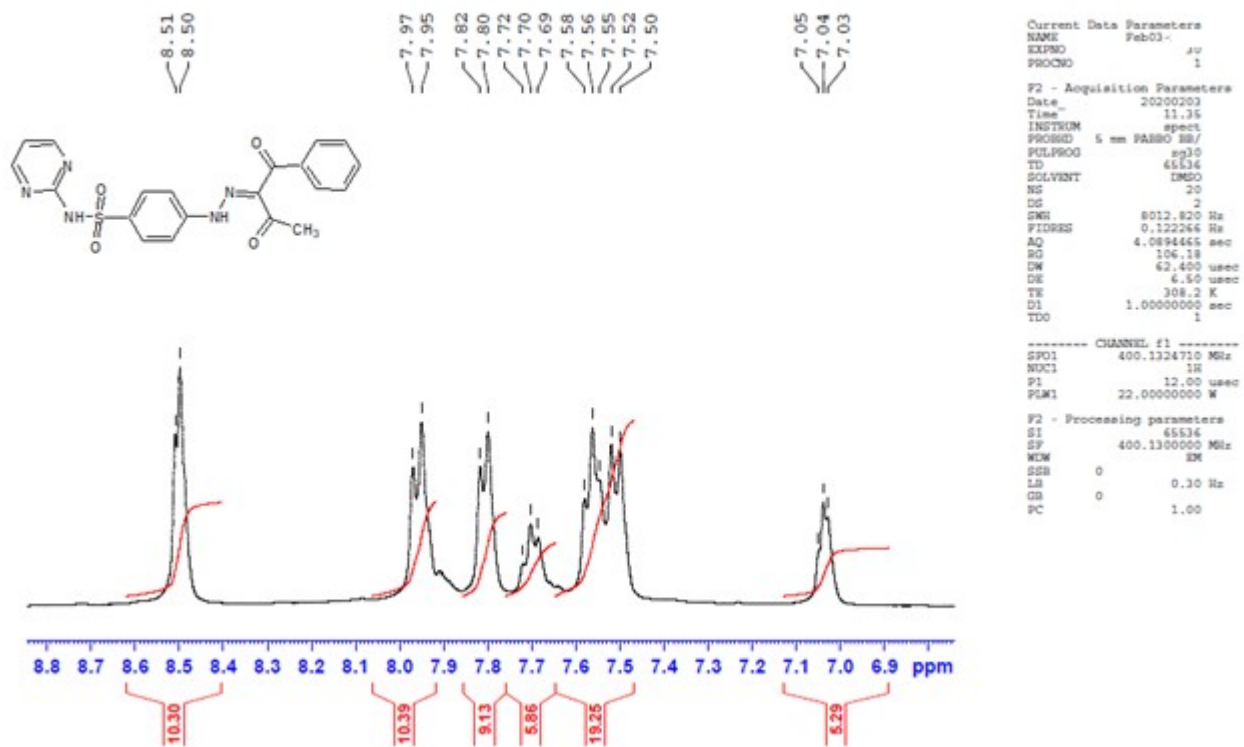
8.51 8.50
8.03 8.01
7.61 7.59
7.06 7.04 7.03
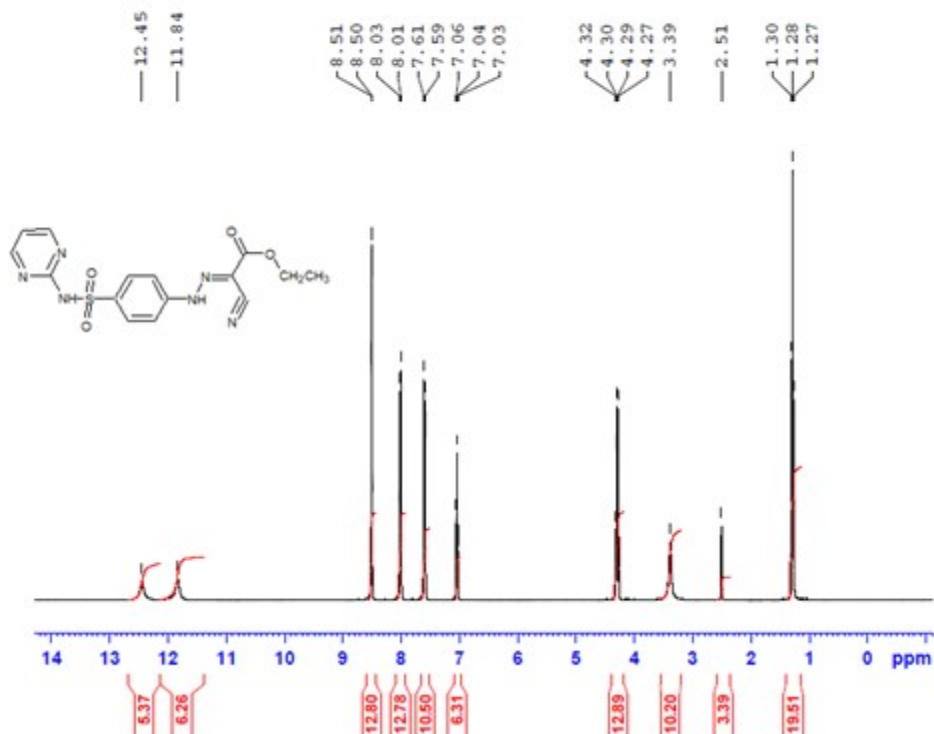
Current Data Parameters
NAME   Jan19-
EXPNO   12U
PROCNO   1

F2 - Acquisition Parameters
Date_   20200119
Time   10.41
INSTRUM   spect
PROBHD   5 mm PABBO BB/
PULPROG   zg30
TD   65536
SOLVENT   DMSO
NS   20
DS   2
SWH   8012.820 Hz
FIDRES   0.122266 Hz
AQ   4.0894465 sec
RG   68.22
DW   62.400 usec
DE   6.50 usec
TE   294.4 K
D1   1.00000000 sec
TDO   1

-------- CHANNEL f1 --------
SFO1   400.1324710 MHz
NUC1   1H
P1   12.00 usec
PLW1   22.00000000 W

F2 - Processing parameters
SI   65536
SF   400.1300000 MHz
WDW   EM
SSB   0
LB   0.30 Hz
GB   0
PC   1.00

8.7 8.6 8.5 8.4 8.3 8.2 8.1 8.0 7.9 7.8 7.7 7.6 7.5 7.4 7.3 7.2 7.1 7.0 ppm

12.80   12.78   10.50   6.31

Figure s7: ^{1}H NMR spectrum of compound **9**
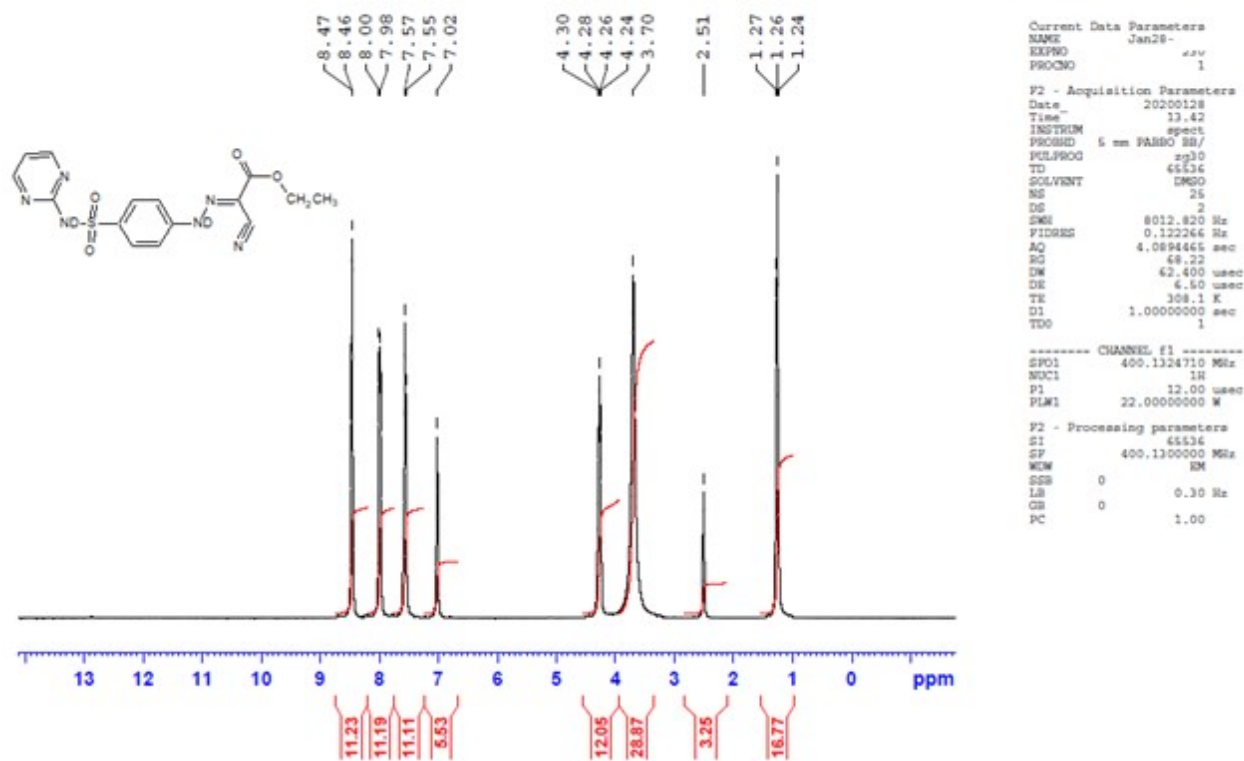
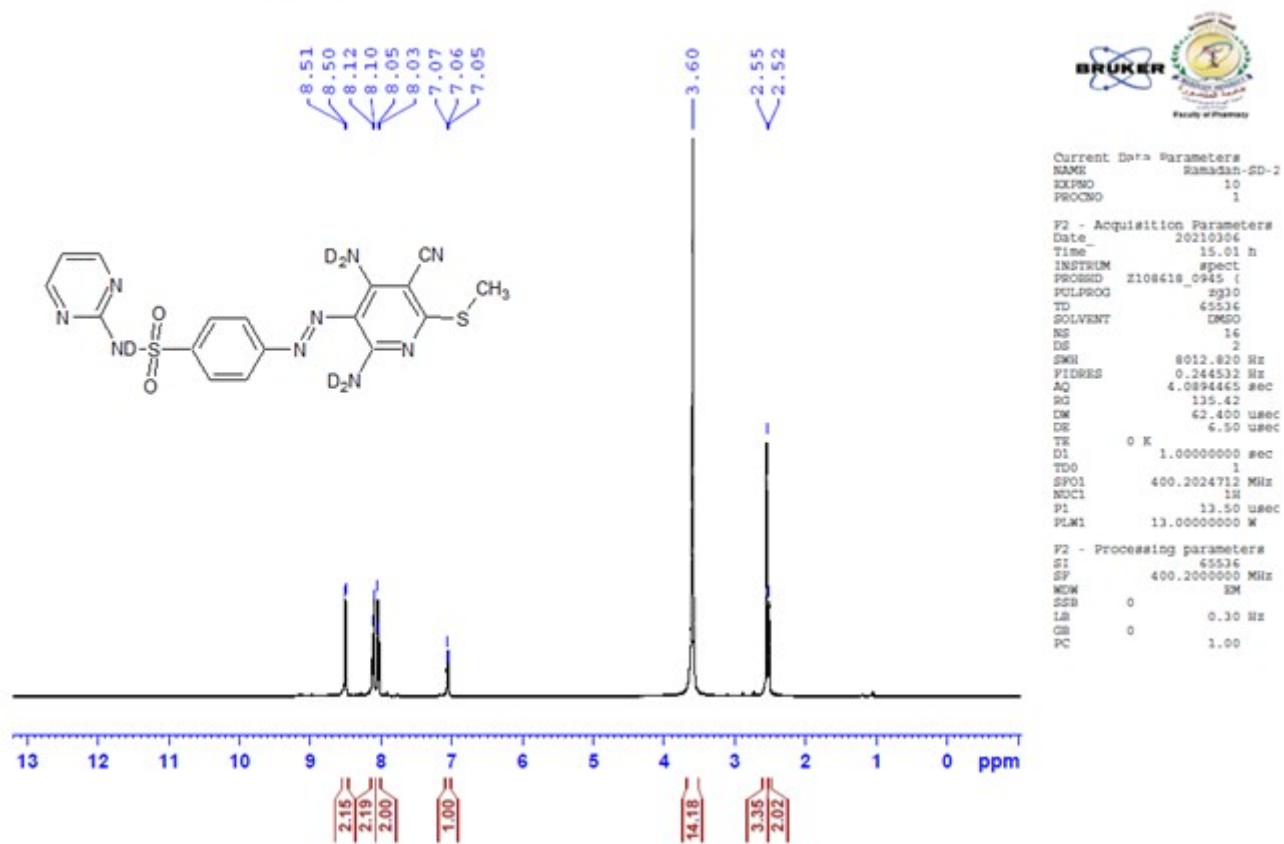**Figure s7:** ^{1}H NMR (D2O) spectrum of compound **9**

Figure s8:¹³C NMR Spectrum of compound 9
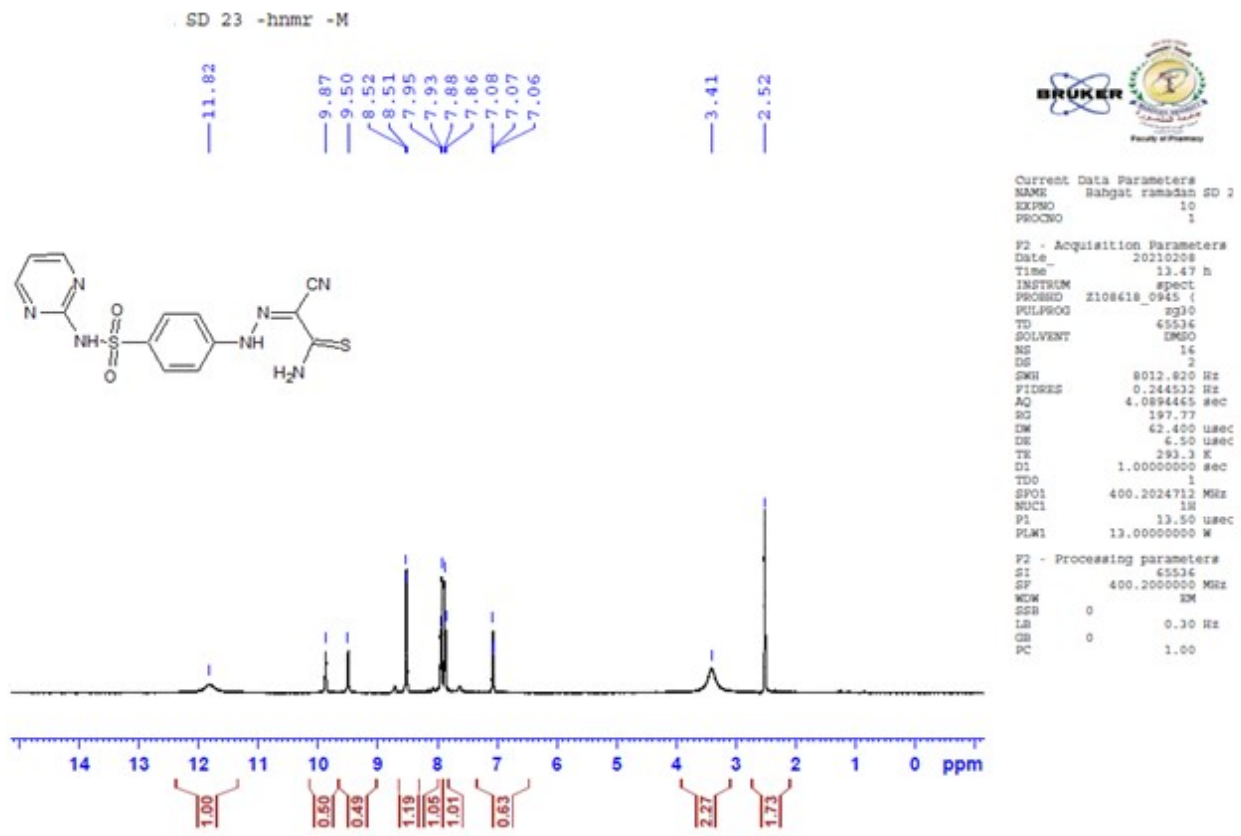
Figure s9: ^{1}H NMR spectrum of compound **10**

Figure s9: ^{1}H NMR spectrum of compound **10**

Figure s10:^{13}C NMR Spectrum of compound 10

BRUKER
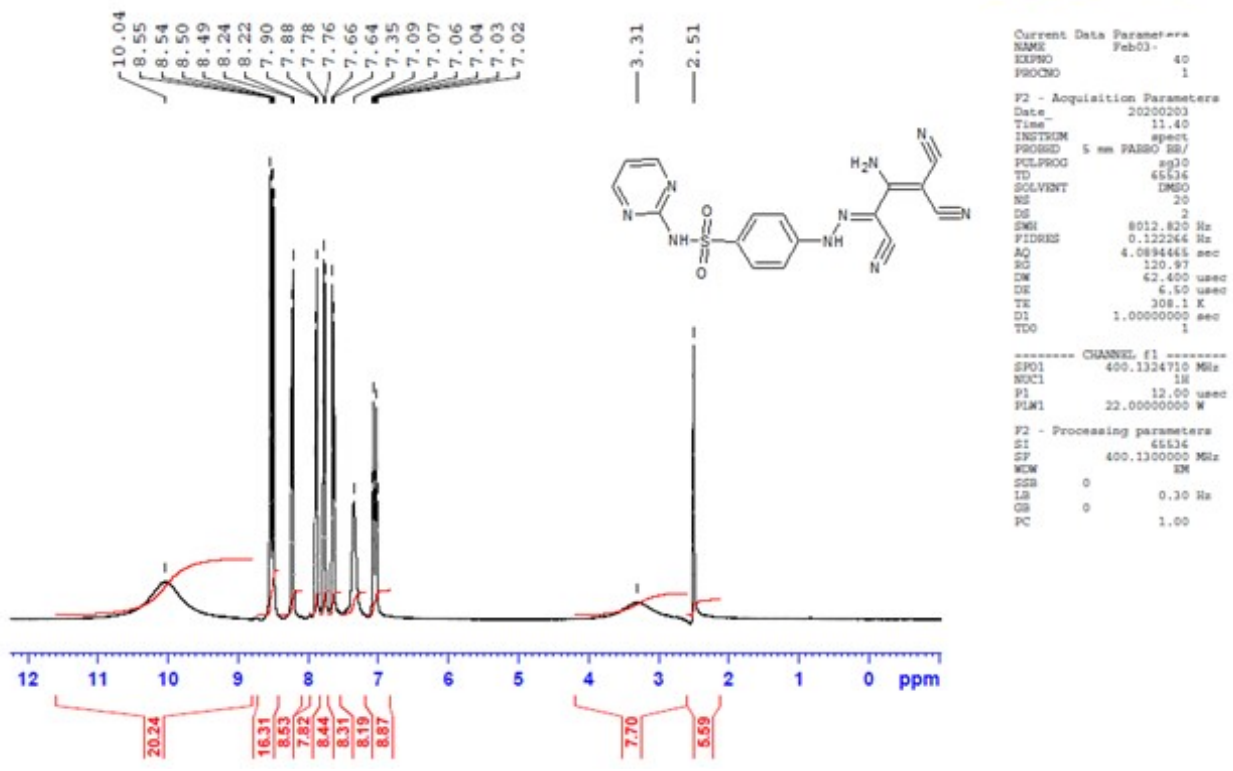
Figure s11: ^{1}H NMR spectrum of compound **11**

Figure s11: ^{1}H NMR spectrum of compound **11**

Figure s12:^{13}**C NMR Spectrum of compound 11**

BRUKER

Figure s13: ^{1}H NMR spectrum of compound **12**

Figure s13: ^{1}H NMR spectrum of compound **12**

Figure s14:¹³C NMR Spectrum of compound 12
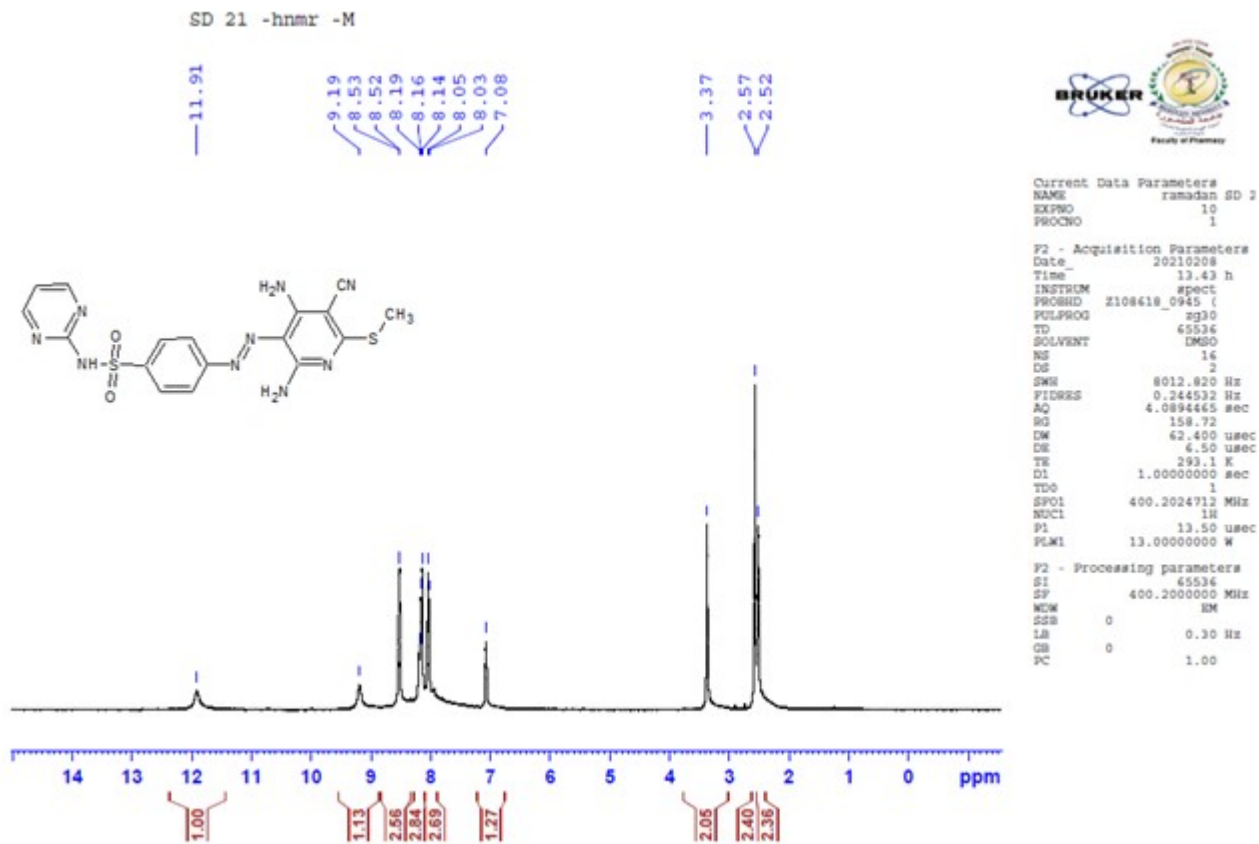
Figure s15: ^{1}H NMR spectrum of compound **13**

Figure s15: ^{1}H NMR (D2O) spectrum of compound **13**
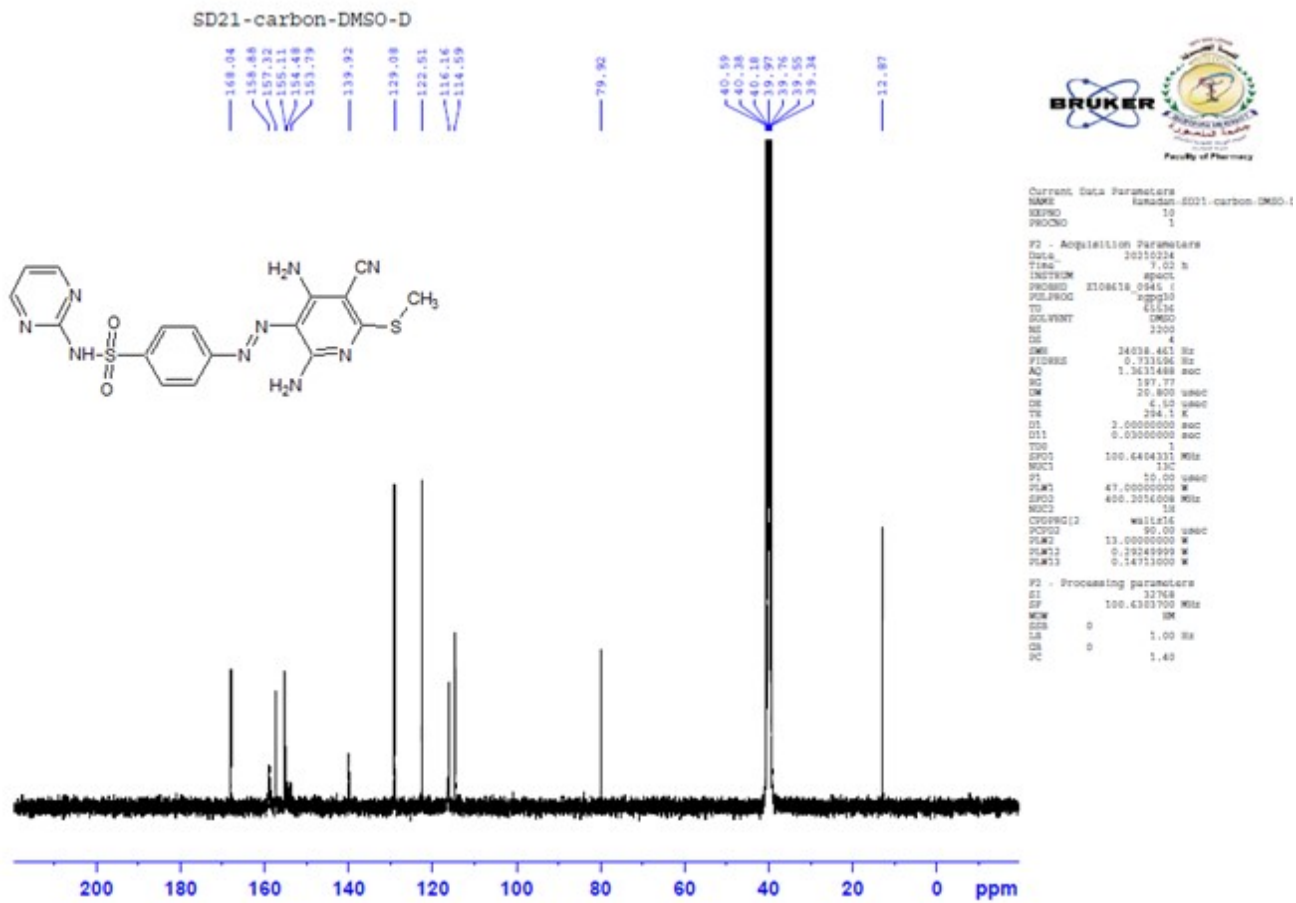
Figure s16: ^{13}C NMR Spectrum of compound **13**
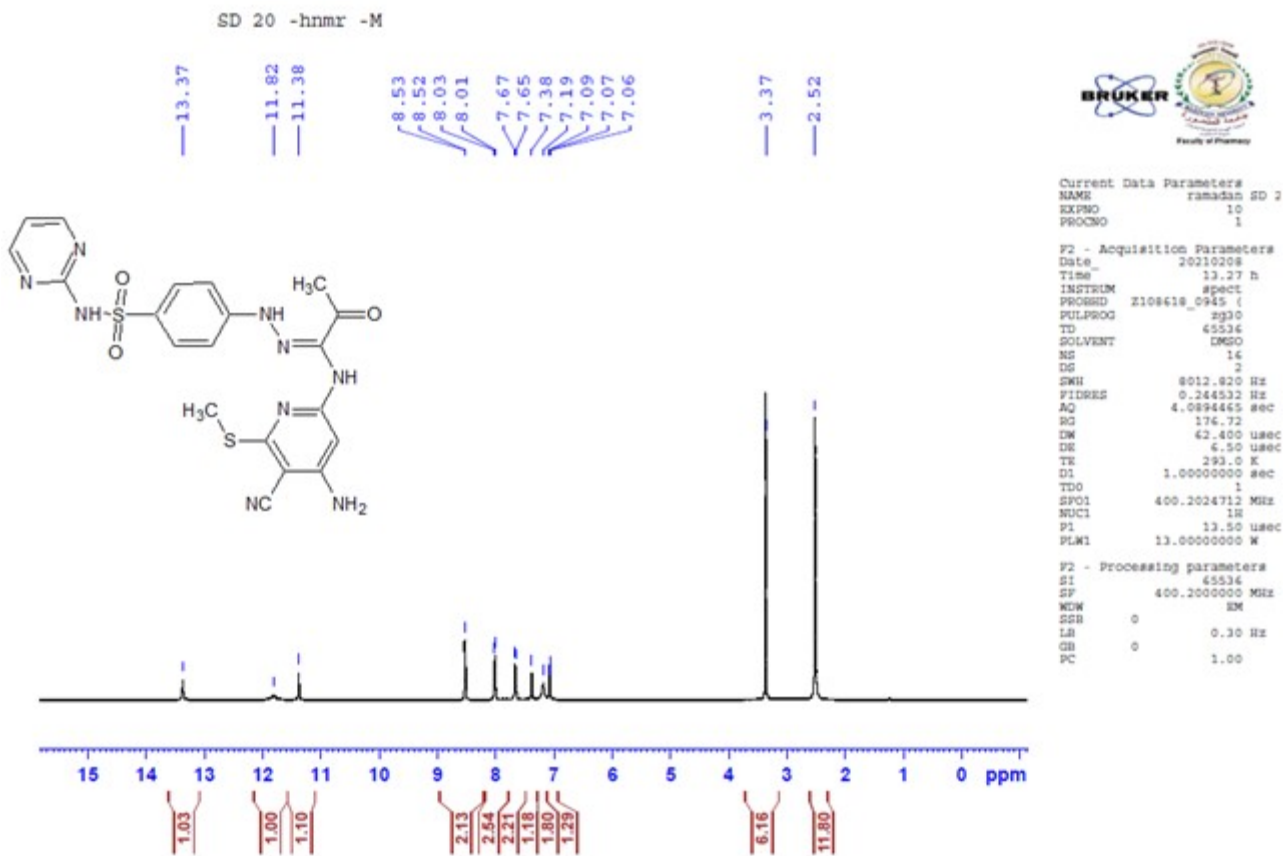
Figure s17: ^{1}H NMR spectrum of compound **14**
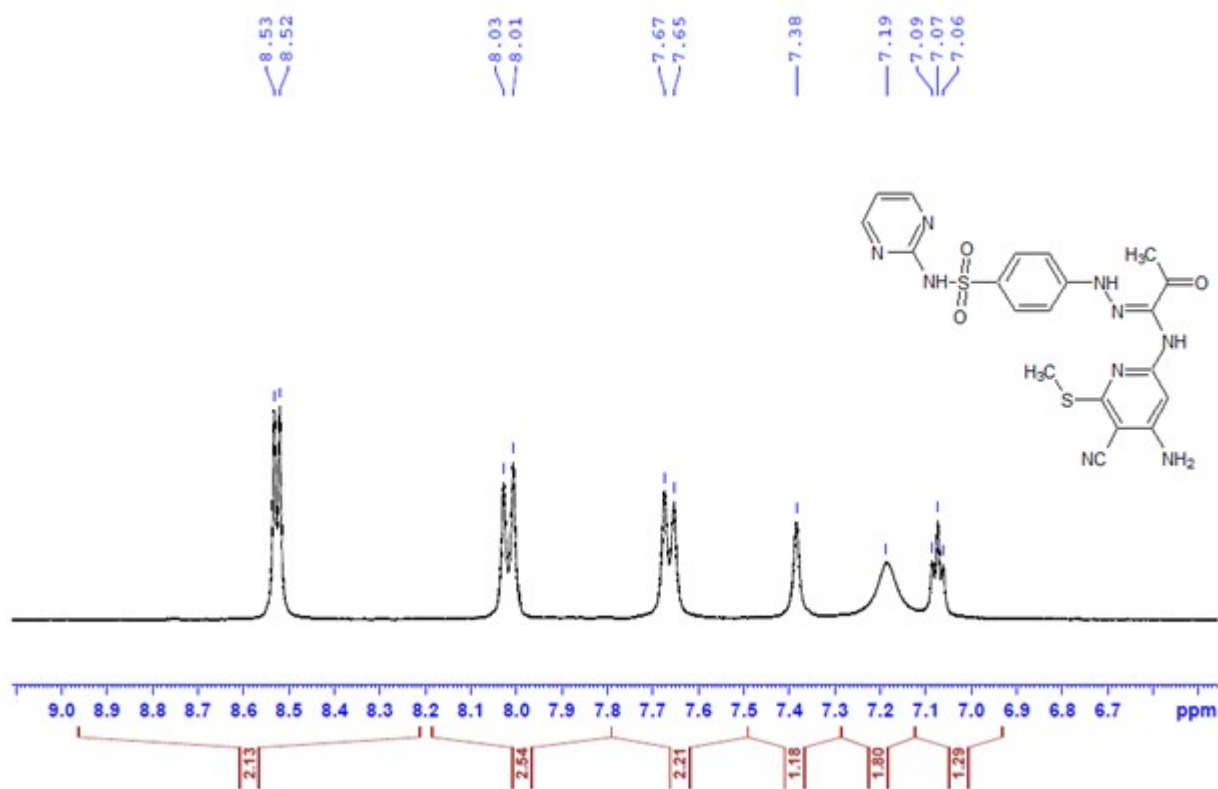
Figure s17: ^{1}H NMR spectrum of compound **14**

A)
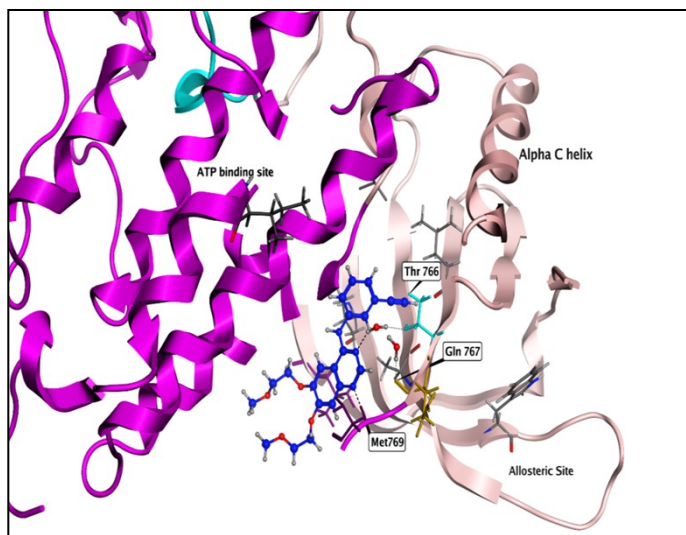
B)
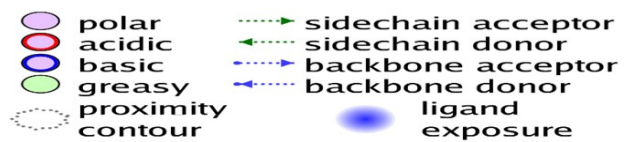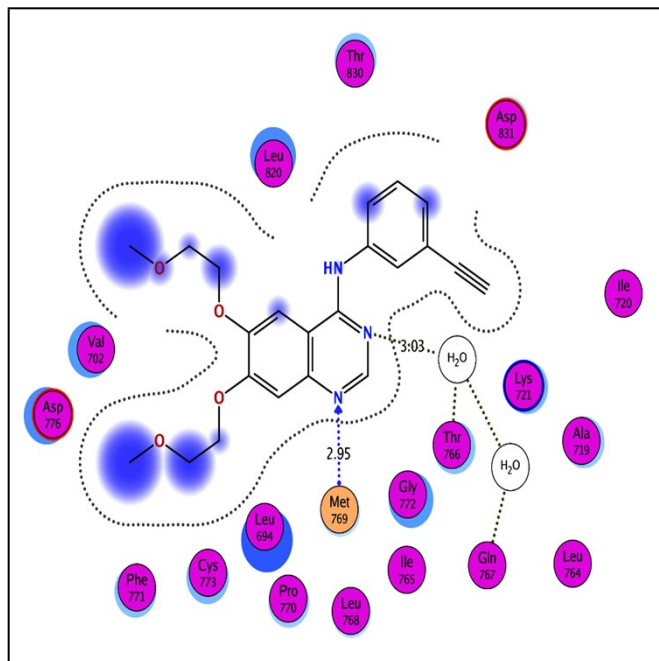
Figure 1S. 3D (**A**) and 2D images (**B**) of the co-crystallized erlotinib (blue) within the wild EGFR kinase (**PDB code: 4HJO**).
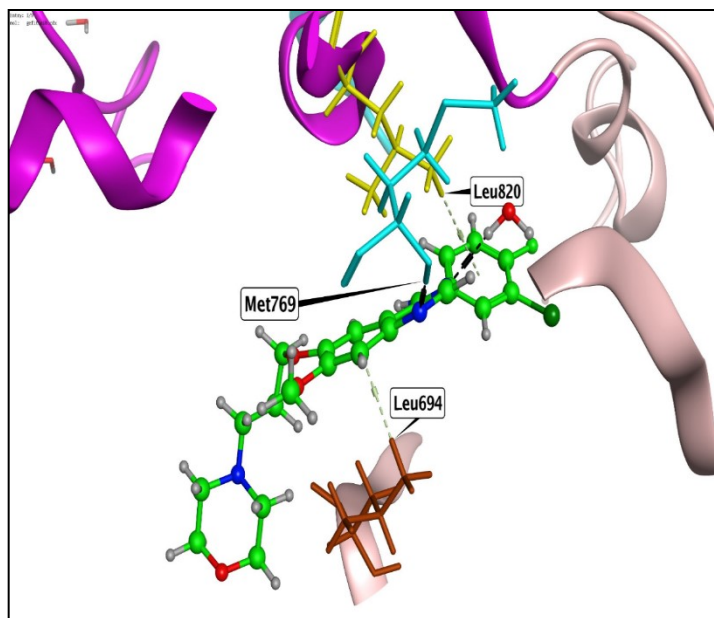
A)
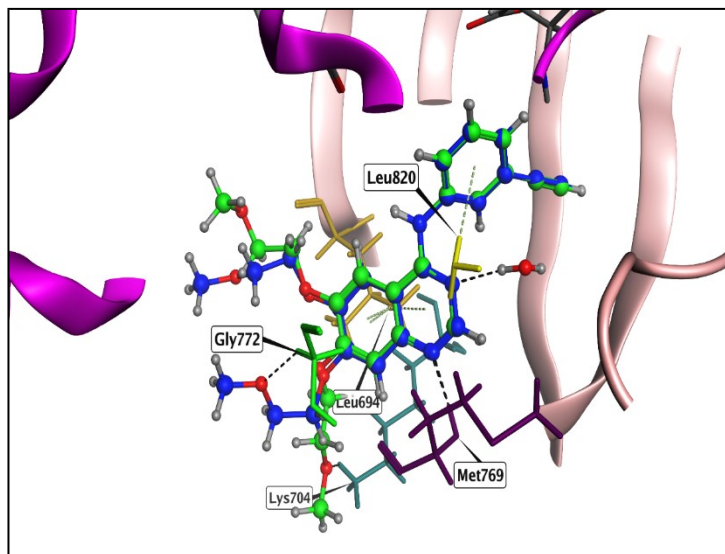
B)
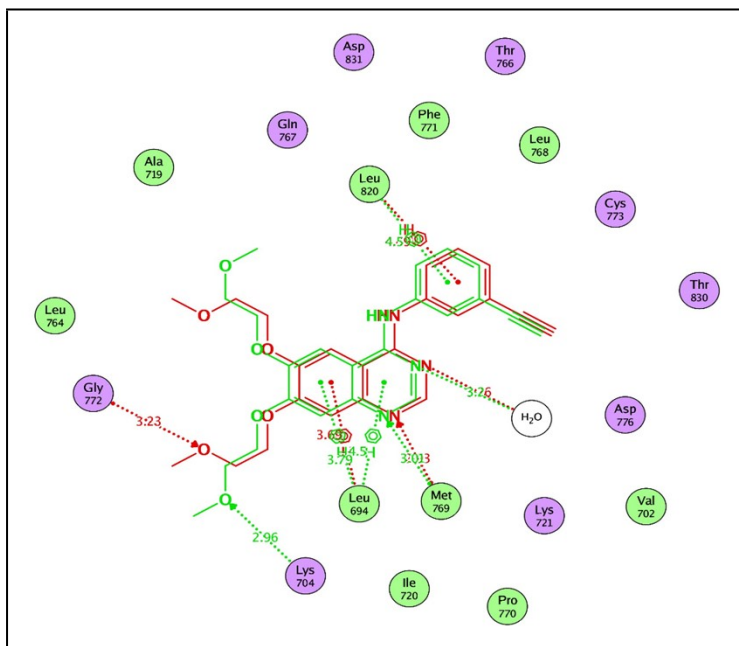
Figure 2S. 3D (**A**) and 2D images (**B**) of the superimposition of the co-crystallized conformers (blue) over re-docked conformers (green) of erlotinib within the wild EGFR kinase **(PDB code: 4HJO)**.
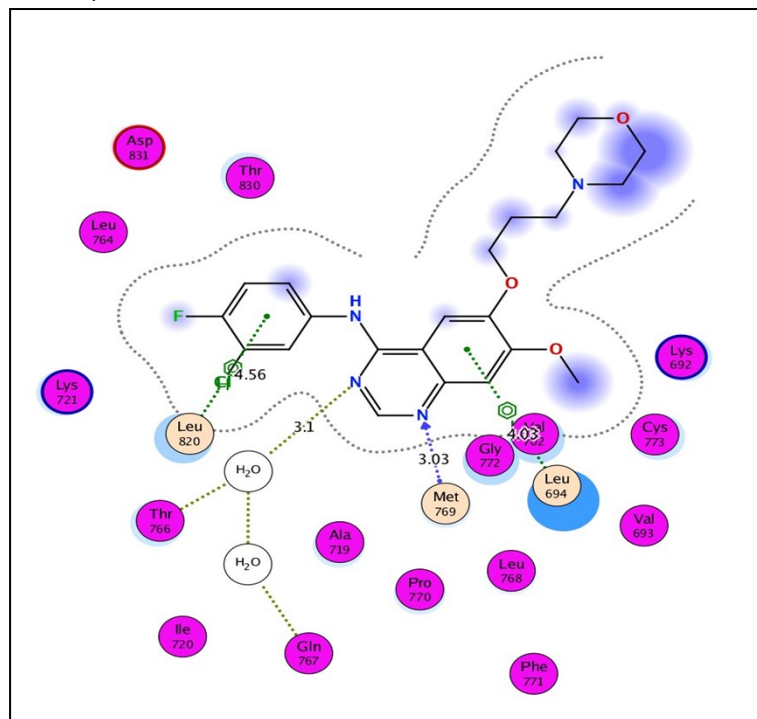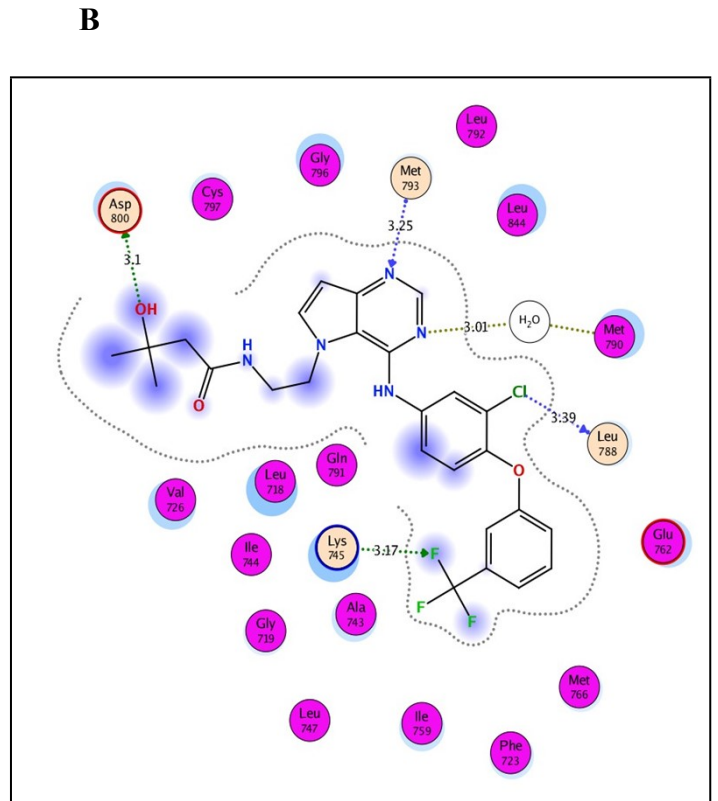
A

B)
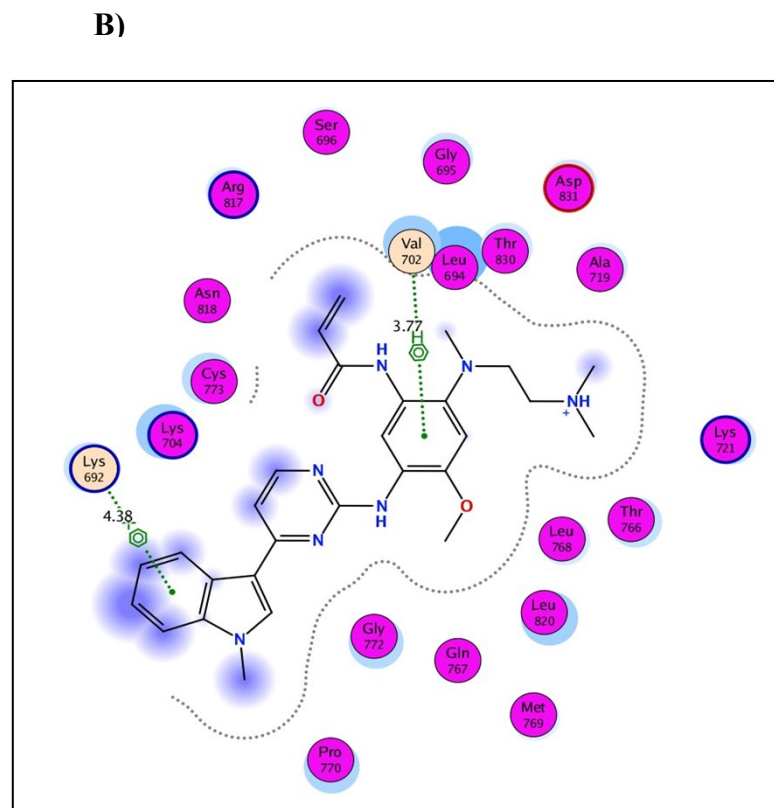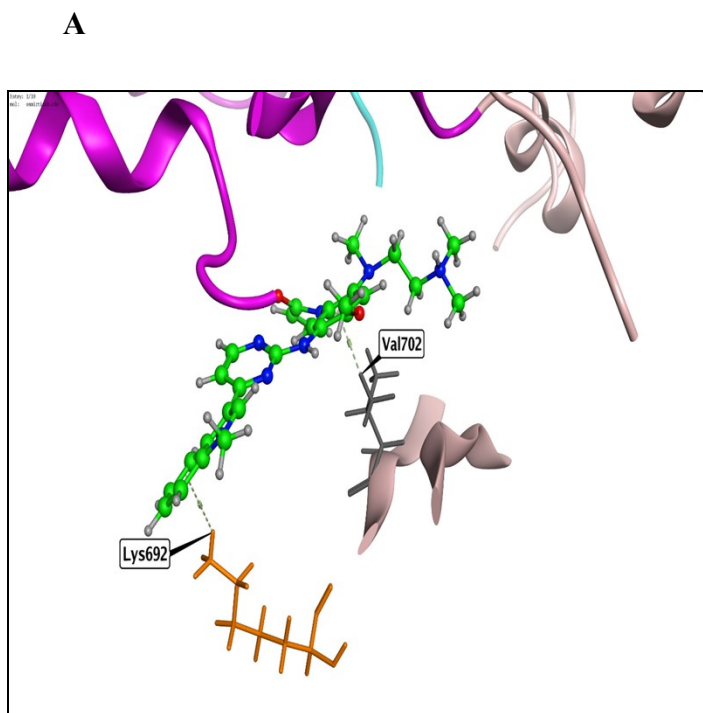
Figure 3S. 3D (**A**) and 2D images (**B**) of osimertinib (green sticks) within the wild EGFR kinase **(PDB code: 4HJO)**.

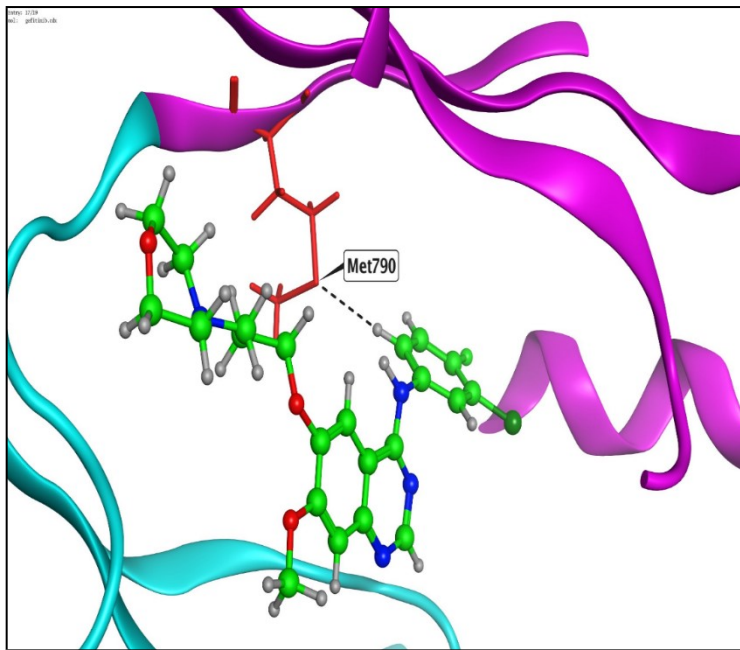**A)**

B)

Figure 4S. 3D (**A**) and 2D images **(B)** of gefitinib (green sticks) within the wild EGFR kinase **(PDB code: 4HJO)**.
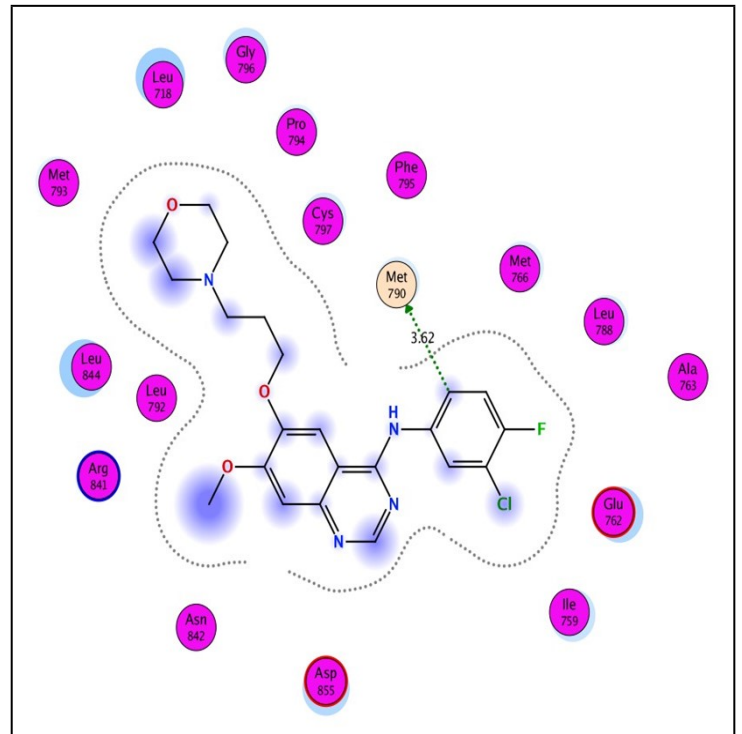
A

B
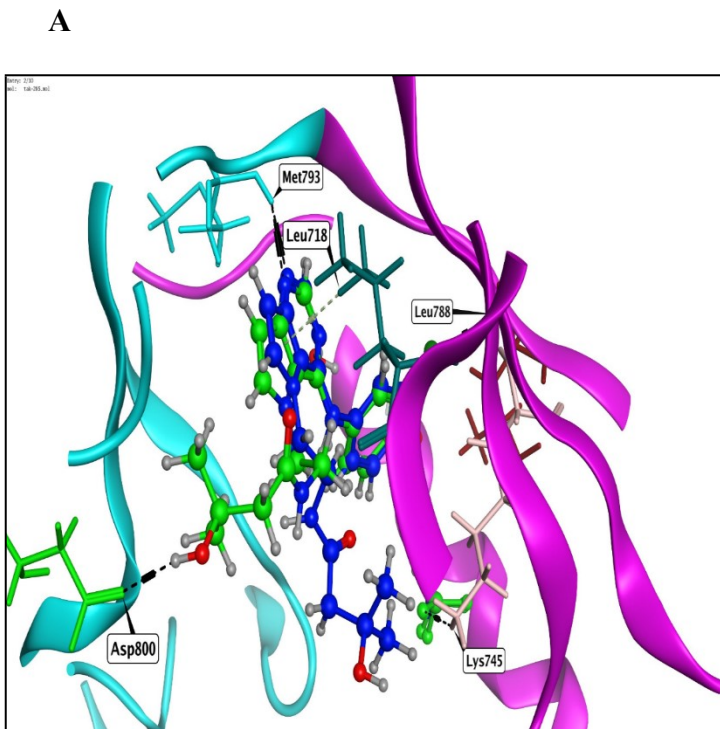
Figure 5S. 3D image of the co-crystallized conformers of TAK-285 (blue) over the re-docked conformers (green) (**A**) and 2D image of the re-docked conformer (green) of TAK-285 within the mutant EGFR^{T790M} kinase **(PDB code:3W2O) (B)**.
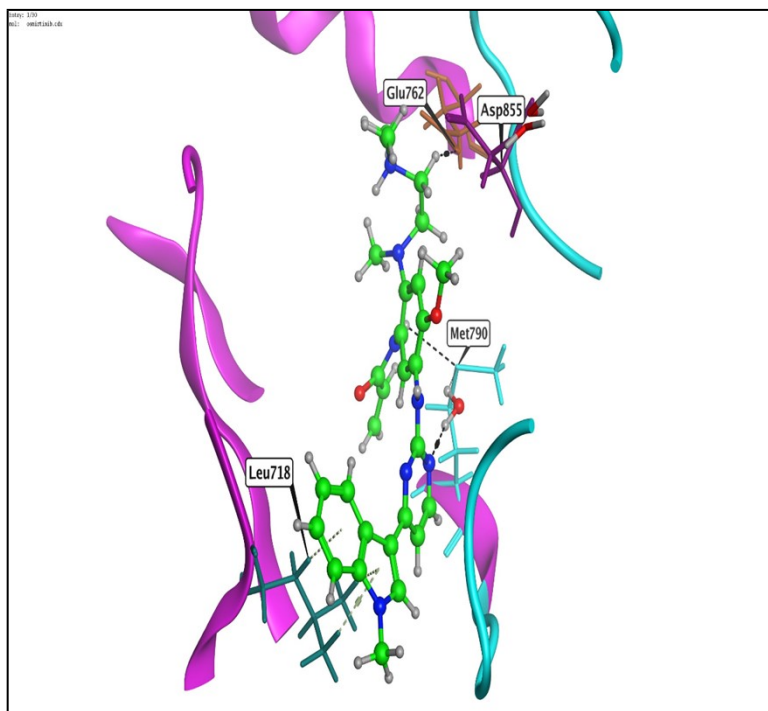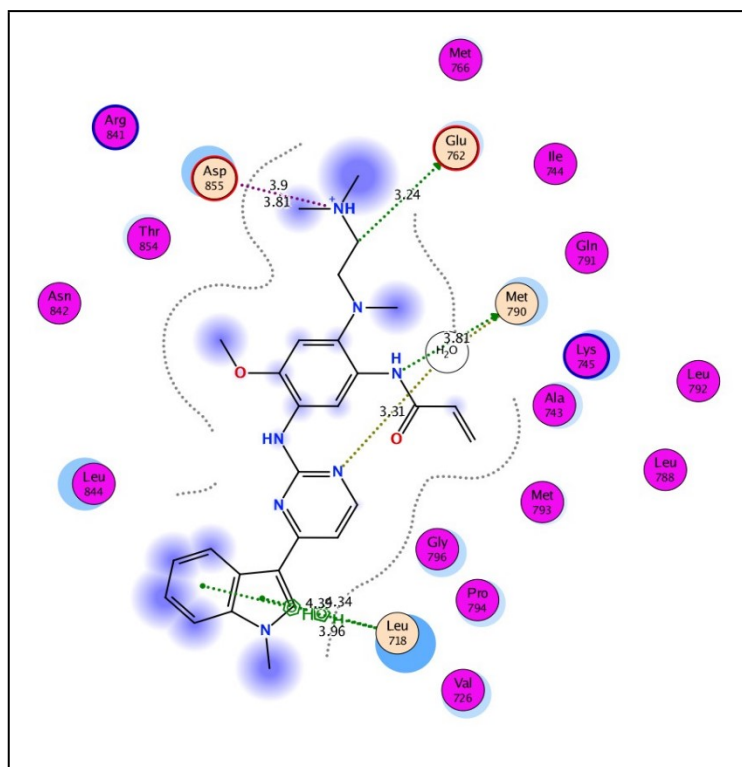
A)

B)

Figure 6S. 3D (**A**) and 2D images (**B**) of gefitinib (green sticks) within the mutant EGFR^{T790M} kinase (**PDB code: 3W2O**).

A)

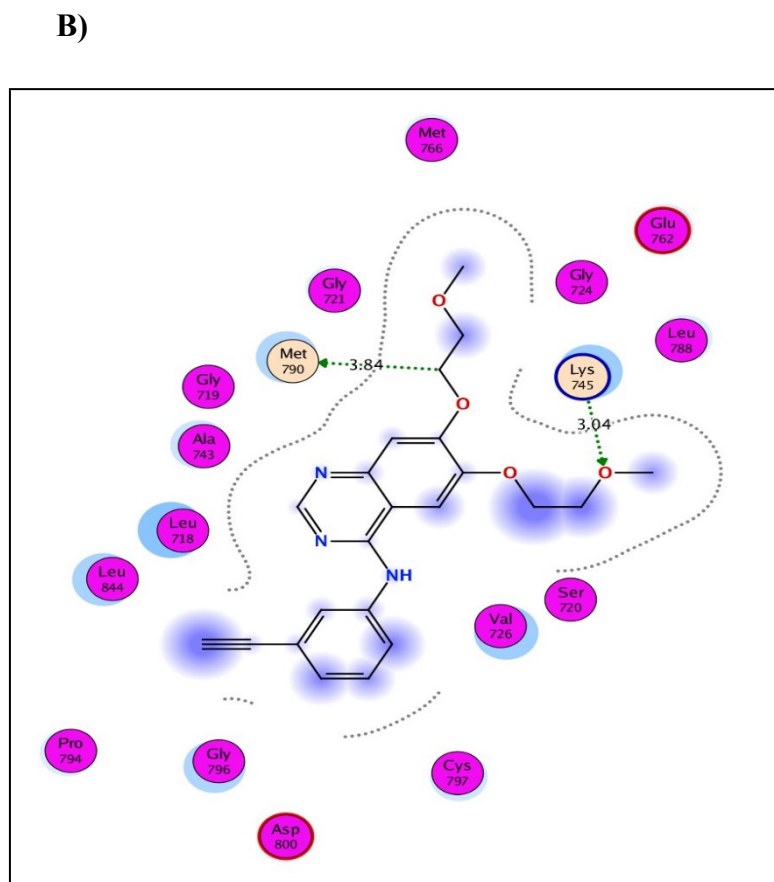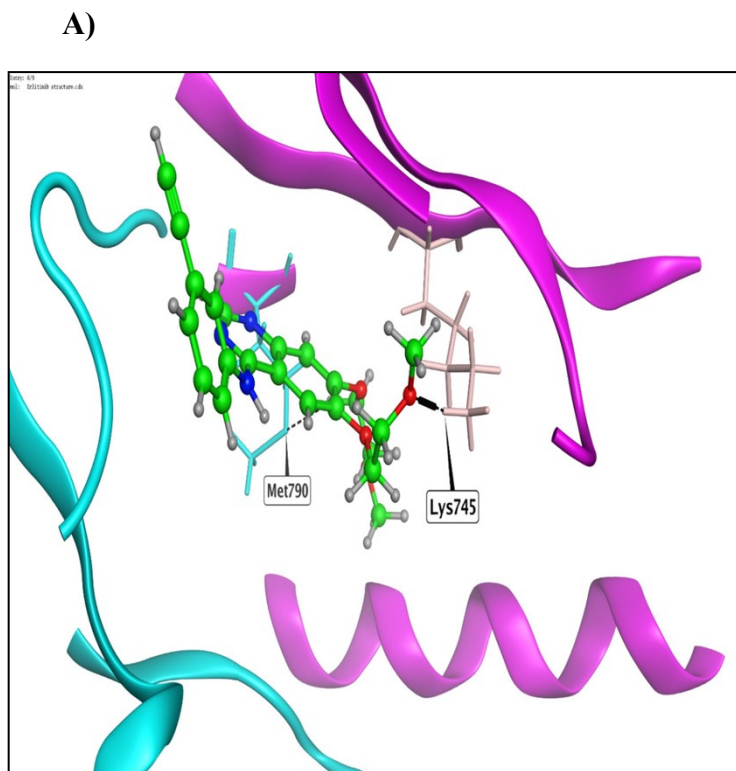**B)**

Figure 7S. 3D (A) and 2D images (B) of osimertinib (green sticks) within the mutant EGFR^{T790M} kinase (PDB code: 3W2O).

A)

B)

Figure 8S. 3D (**A**) and 2D images **(B)** of erlotinib (green sticks) within the mutant EGFR^{T790M} kinase **(PDB code: 3W2O)**.

Table S1. Physicochemical properties, lipophilicity, water-solubility, pharmacokinetics, drug-likeness, medicinal chemistry, and toxicity properties obtained *via* SwissADME and pkCSM websites of compounds **8**, **12**, and 14 the references Erlotinib, and Gefitinib.

Physicochemical properties	8	12	14	Erlotinib	Gefitinib
Formula	$C_{20}H_{17}N_5O_4S$	$C_{19}H_{13}N_7O_2S_2$	$C_{17}H_{15}N_9O_2S_2$	$C_{22}H_{23}N_3O_4$	$C_{22}H_{24}ClFN_4O_3$
Molecular weight	423.45 g/mol	435.48 g/mol	441.49 g/mol	393.44 g/mol	446.90 g/mol
Num. heavy atoms	30	30	30	29	31
Num. arom. Heavy atoms	18	21	12	16	16
Fraction Csp3	0.05	0.00	0.06	0.27	0.36
Num. rotatable bonds	8	6	6	10	8
Num. H-bond acceptors	7	7	8	6	7
Num. H-bond donors	2	2	4	1	1
Molar refractivity	111.67	114.36	119.04	111.40	121.66
TPSA (topological polar surface area)	138.86 Å²	169.64 Å²	216.04 Å²	74.73 Å²	68.74 Å²
Lipophilicity					
Log $P_{o/w}$ (iLOGP)	2.02	2.43	1.18	3.55	4.04
Log $P_{o/w}$ (XLOGP3)	2.98	4.16	1.76	2.63	4.11
Log $P_{o/w}$ (WLOGP)	3.22	3.93	1.85	3.07	4.32
Log $P_{o/w}$ (MLOGP)	-0.03	0.21	-1.38	2.06	2.55
Log $P_{o/w}$ (SILICOS-IT)	1.67	2.41	0.61	4.77	4.31
Consensus Log $P_{o/w}$	1.97	2.63	0.80	3.22	3.86
Water solubility					
Log S (ESOL)	-4.26	-5.28	-3.59	-4.11	-5.05
Solubility	2.33e-02 mg/ml; 5.51e-05 mol/L	2.27e-03 mg/ml; 5.21e-06 mol/L	1.15e-01 mg/ml; 2.59e-04 mol/L	3.03e-02 mg/ml; 7.71e-05 mol/L	3.95e-03mg/ml; 8.83e-06 mol/L
Class	Moderately soluble	Moderately soluble	Soluble	Moderately soluble	Moderately soluble
Log S (Ali)	-5.56	-7.43	-5.91	-4.56	-5.26
Solubility	1.17e-03 mg/ml; 2.76e -06 mol/L	1.62e-05 mg/ml; 3.71e -08 mol/L	5.38e-04 mg/ml; 1.22e -06 mol/	1.10e-02 mg/ml; 2.78e-05 mol/L	2.46e-03 mg/ml; 5.50e-06 mol/L
Class	Moderately soluble	Poorly soluble	Moderately soluble	Moderately soluble	Moderately soluble
Log S (SILICOS-IT)	-6.99	-7.36	-5.77	-7.26	-7.94
Solubility	4.29e-05 mg/ml; 1.01e-7 mol/L	1.91e-05 mg/ml; 4.40e-08 mol/L	7.57e-04 mg/ml; 1.71e-06 mol/L	2.15e-05 mg/ml; 5.46e-08 mol/L	5.14e-06 mg/ml; 1.15e-08 mol/L

Physicochemical properties	8	12	14	Erlotinib	Gefitinib
Class	Poorly soluble	Poorly soluble	Moderately soluble	Poorly soluble	Poorly soluble
Pharmacokinetics					
GI absorption	Low	Low	Low	High	High
BBB permeant	No	No	No	Yes	Yes
P-gp substrate	No	No	Yes	No	No
CYP1A2 inhibitor	No	No	No	Yes	No
CYP2C19 inhibitor	Yes	Yes	No	Yes	Yes
CYP2C9 inhibitor	Yes	Yes	No	Yes	Yes
CYP2D6 inhibitor	No	No	No	Yes	Yes
CYP3A4 inhibitor	Yes	Yes	No	Yes	Yes
Log Kp (skin permeation)	-6.77 cm/s	-6.00 cm/s	-7.74 cm/s	-6.35 cm/s	-6.11 cm/s
Druglikeness					
Lipinski	Yes; 0 violation	Yes; 0 violation	Yes; 1 violation; NorO>10	Yes; 0 violation	Yes; 0 violation
Ghose	Yes	Yes	Yes	Yes	Yes
Veber	Yes	No; 1 violation; TPSA>140	No; 1 violation; TPSA>140	Yes	Yes
Egan	No; 1 violation; TPSA>131.6	No; 1 violation; TPSA>131.6	No; 1 violation; TPSA>131.6	Yes	Yes
Muegge	Yes	No; 1 violation; TPSA>150	No; 1 violation; TPSA>150	Yes	Yes
Bioavailability score	0.55	0.55	0.55	0.55	0.55
Medicinal chemistry					
PAINS	1 alert: imine-one-A	0 alert	0 alert	0 alert	0 alert
Brenk	2 alerts: beto-keto-anhydride, imine-1	1 alert: imine-one-A	1 alert: imine-one-A	1 alert: triple bond	0 alert
Leadlikeness	No; 2 violations: MW>350, Rotors>7	No; 2 violations: MW>350, XLOG3>3.5	No; 1 violation: MW>350	No; 2 violations: MW>350, Rotors>7	No; 3 violations: MW>350, Rotors>7, XLOG3>3.5
Synthetic accessibility	3.17	3.43	4.25	3.19	3.26
Toxicity					

Physicochemical properties	8	12	14	Erlotinib	Gefitinib
AMES toxicity (Yes/No)	No	No	No	No	No
Max. tolerated dose (human) (log mg/kg/day)	0.347	0.041	0.347	-0.629	-0.304
hERG I inhibitor (Yes/No)	No	No	No	No	No
hERG II inhibitor (Yes/No)	Yes	Yes	No	Yes	Yes
Oral rat acute toxicity (LD50) (mol/kg)	2.364	2.446	2.007	2.676	2.688
Oral rat chronic toxicity (LOAEL) (log mg/kg_bw/day)	1.859	1.245	1.375	0.969	1.491
Hepatotoxicity (Yes/No)	Yes	Yes	Yes	Yes	Yes
Skin sensitization (Yes/No)	No	No	No	No	No
T. pyriformis toxicity (log μg/L)	0.373	0.305	0.292	0.318	0.293
Minnow toxicity (log mM)	1.209	0.999	1.542	-1.725	-1.952

Abbreviation: TPSA, topological polar surface area.