Sains Malaysiana 52(4)(2023): 1203-1215
http://doi.org/10.17576/jsm-2023-5204-13

Molecular Docking Analysis on the Designed Benzimidazole Derivatives as EGFR Inhibitors: Comparison between EGFR Wild-Type ($EGFR_{WT}$) and T790M Mutant

(Analisis Dok Molekul pada Terbitan Benzimidazol Direka sebagai Perencat EGFR: Perbandingan antara Jenis Liar EGFR ($EGFR_{WT}$) dan Mutan T790M)

NURUL AWANI SYAZZIRA JALIL[1] & SHAFIDA ABD HAMID[1,2,*]

[1]Department of Chemistry, Kulliyyah of Science, International Islamic University Malaysia, 25200 Bandar Indera Mahkota, Kuantan, Pahang Darul Makmur, Malaysia
[2]SYNTOF, Kulliyyah of Science, International Islamic University Malaysia, 25200 Bandar Indera Mahkota, Kuantan, Pahang Darul Makmur, Malaysia

Received: 15 August 2022/Accepted: 26 February 2023

ABSTRACT

The non-small cell lung (NSCL) and colorectal cancers are frequently linked with the oncogenic activation of the epidermal growth factor receptor (EGFR), a member of the receptor tyrosine kinase (RTK) family. Current tyrosine kinase inhibitors (TKIs) are susceptible to drug resistance mutations and induce cytotoxicity effects on normal EGFRs. The isosteric nature of benzimidazole with purine renders its great potential to imitate the binding mode of the purine-based ATP and prevents its contact with the EGFR active sites. Here, we report the molecular docking of 50 designed benzimidazole derivatives, as well as Gefitinib and ATP, to analyse and compare their binding modes at $EGFR_{wt}$ and T790M active sites. The design of the ligands is based on our previous study, in which we proposed to evaluate keto- and amino-benzimidazoles, attached to a double bond linker and a phenyl group having electron donating and electron withdrawing groups attached at various positions. Docking simulations showed that keto-benzimidazoles dominated the top ten highest binding affinities in both EGFR-ligand complexes. The presence of sulfonyl substituents contributed to more stable complexes compared to others with binding energies of -8.1 (**7c**) and -7.8 (**11c**) kcal/mol in $EGFR_{wt}$, and -8.3 (**7d**) and -8.4 (**1c**) kcal/mol for T790M mutant. The substituent effects on the benzimidazole contributed not only to the hydrogen bonding and hydrophobic interaction, but also to the often-disregarded Van der Waals forces that are responsible for shape complementarity of the benzimidazoles with the EGFR binding pocket.

Keywords: Benzimidazole; EGFR; molecular docking; tyrosine kinase inhibitor; T790M

ABSTRAK

Sel paru-paru bukan kecil (NSCL) dan kanser kolorektum sering dikaitkan dengan pengaktifan onkogenik reseptor faktor pertumbuhan epidermis (EGFR), ahli keluarga reseptor tirosin kinase (RTK). Perencat tirosin kinase (TKI) masa kini terdedah kepada mutasi rintangan dadah dan mendorong kesan kesitotoksikan pada EGFR normal. Sifat isosterik benzimidazol dengan purin menyebabkan ia berpotensi tinggi untuk meniru mod pengikatan ATP berasaskan purin dan menghalang sentuhannya dengan tapak aktif EGFR. Di sini, kami melaporkan dok molekul bagi 50 terbitan benzimidazol yang direka berserta Gefitinib dan ATP untuk menganalisis dan membandingkan mod pengikatan mereka di tapak aktif $EGFR_{wt}$ dan T790M. Reka bentuk ligan adalah berdasarkan kajian terdahulu kami dan kami mencadangkan untuk menilai benzimidazol-keto dan amino- yang terikat pada penghubung ikatan berganda dan kumpulan fenil yang mempunyai kumpulan pendermaan elektron dan penarikan elektron yang dilampirkan pada pelbagai kedudukan. Simulasi dok mendedahkan bahawa benzimidazol-keto menguasai sepuluh pertalian pengikatan tertinggi dalam kedua-dua kompleks ligan EGFR. Kehadiran penukarganti sulfonil menyumbang kompleks yang lebih stabil berbanding yang lain; dengan tenaga pengikatan -8.1 (**7c**) dan -7.8 (**11c**) kcal/mol dalam $EGFR_{wt}$ dan -8.3 (**7d**) dan -8.4 (**1c**) kcal/mol untuk mutan T790M. Kesan penukarganti pada benzimidazol menyumbang bukan sahaja kepada ikatan hidrogen dan interaksi hidrofobik, tetapi juga kepada daya Van der Waals yang sering diabaikan, yang bertanggungjawab untuk saling melengkapi bentuk benzimidazol dengan poket pengikat EGFR.

Kata kunci: Benzimidazol; dok molekul; EGFR; perencat tirosin kinase; T790M

Introduction

Cancer is caused by unregulated cell proliferation forming either a benign or malignant tumour. A malignant tumour can continuously spread through the bloodstream, infects other tissues and disrupts bodily functions, causing fatality (Mathur et al. 2015; Nussbaumer et al. 2011; Rebucci & Michiels 2013). The Global Cancer Statistics reported 19.3 million new cancer cases in 2020 and estimated a 47% rise by 2040 (Sung et al. 2021). The occurrence of non-small cell lung cancer (NSCL) and colorectal cancer is frequently linked with the oncogenic activation of EGFR. The metastasis can be curbed by targeting and inhibiting the oncogenic epidermal growth factor receptor (EGFR) residing on the cellular membrane (Troiano et al. 2016). EGFR is one of a family of four receptor tyrosine kinases (RTKs) in humans, the others being ErbB2/HER2, ErbB3/HER3, and ErbB4/HER4, which are responsible for initiating complex biological functions such as cell growth, motility, proliferation, and apoptosis by passing intercellular signals. Conformational changes caused by the binding of ligands such as growth factors and hormones to the extracellular domain of RTKs induce their dimerisation. The dimer is only activated once the phosphate group from ATP is transferred to the tyrosine on the C-terminal of the RTK domain. The phosphorylated tyrosine then engages the neighbouring relay proteins, which would give out signals for specific cellular responses (Amelia et al. 2022).

Inhibitors of these receptors, involving the blocking of the EGFR binding site from interacting with agonists have been among the most successful examples of targeted cancer therapies. The action triggers the metastasis stage of malignant tumour cells by silencing the overexpressed or mutated EGFR. There are currently three TKIs generations clinically practiced, all of which bear an azaheterocyclic moiety. However, EGFR-target therapy drugs are continuously being developed to overcome the drugs' limitations in terms of toxicity, non-selectivity, and resistance to mutations.

Benzimidazole is non-foreign in pharmacology as its derivatives are frequently associated with various biological activities. The scaffold is a structural isostere of indole and purine and thus its derivatives are expected to exhibit good/favourable affinity with various types of receptors. Benzimidazole derivatives exhibit anticancer activity via a few mechanisms, such as through topoisomerase I and II inhibition, DNA intercalation, PARP-poly inhibition, dihydrofolate reductase (DHFR) and aromatase inhibition (Cevik et al. 2022; Karadavi et al. 2020; Mostafa, Gomaa & Elmorsy 2019). For example, the derivatives have been employed in two known cancer drugs, which are Veliparp and Nocodazole, although both drugs act via poly(ADP ribose) polymerase (PARP) inhibition and microtubule function disruption, respectively (Shrivastava et al. 2017). Benzimidazole interacts with Met793 of EGFR via a similar binding mode to quinazoline, owing to the nitrogen atoms on the nucleus acting as hydrogen bond acceptors (Abdullah, Ali & Hamid 2021). Issues such as cytotoxicity, non-selectivity, and unforeseen EGFR gene mutations further demand a more potent chemotherapeutic agent.

Given that benzimidazole is structurally similar to quinazoline, the scaffold for first- and second-generation drugs, the compound could be potentially served as an effective nucleus for future EGFR antagonists. However, there are still lots of studies that need to be carried out before the benzimidazole-based drug can be clinically approved for EGFR inhibition applications. Majority of the studies on structural insights of benzimidazole derivatives showed ambiguous findings on the type and position of the groups attached to the main scaffold, as well as the superiority of electron withdrawing and donating groups incorporation in improving the EGFR binding activity.

Overexpression of the EGFR wild-type ($EGFR_{wt}$) occurs more frequently than in the EGFR mutants (Thomas & Weihua 2019). The EGFR (Figure 1) consists of 28 exons and its mutations are grouped into three classes of nucleotide changes; i) short deletions of residues between Glu746 and Ser752 encoded by exon 19 (Class I), ii) substitution between exons 18 to 21 (Class II), and iii) duplications or insertions in exon 20 (Class III) [6]. These mutations cause the ligands to activate the tyrosine, exerting aberrant intercellular signals that trigger malignant tumour growth (Zandi et al. 2007). The current tyrosine kinase inhibitors (TKIs) exhibited only initial clinical responses before the malignant tumour develops resistance towards the inhibition, and the efficacy of first and second generation TKIs, namely Gefitinib and Erlotinib, are often disrupted by the so-called T790M 'gatekeeper' mutation. The mutation at the EGFR tyrosine kinase domain has been described to be caused by the substitution of threonine with a bulky methionine, which results in steric interference for the TKIs to engage with the pocket. Further studies showed the mutation also increases drug resistance by increasing the ATP binding affinity and thus weakens the binding affinity of the drugs (Vyse & Huang 2019).

The third generation TKI such as Osimertinib addresses the T790M mutation using the flexibility of its structure to pass through the entrance and anchors onto the binding site (Lee at al. 2019). Despite being able to sensitize for the T790M EGFR, Osimertinib is inactive against novel mutation on EGFR L792 and C797S (Grabe et al. 2018; Jiang et al. 2018; Leonetti et al. 2019; Song & Yang 2022). This leads to the development of the fourth generation TKIs such as EAI045 to inhibit the kinase activity of double mutant T790M/C797S. EAI045 however must be used in combination with another anti-EGFR agent, namely Cetuximab (Thomas & Weihua 2019). The resistance mechanism becomes more challenging when multiple mutations are involved.

STRUCTURE-ACTIVITY RELATIONSHIP BETWEEN BENZIMIDAZOLE AND EGFR

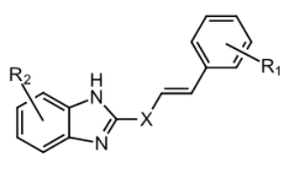
The biological activities exhibited by benzimidazoles prompted continuous studies on their structure-activity relationship to aid drug discovery. The activities are typically elicited when the substituents are grafted at 1,2,4 and/or 5 positions (Figure 3) (Bansal & Silakari 2012), although most studies concentrated on the attachment of various types of groups at C2, including alkyl chain, amide, carbonyl, benzene and heteroatomic ring. In many cases, these groups are also attached to substituted or unsubstituted phenyl group (Akhtar et al. 2017; Celik et al. 2019; Lelais et al. 2016; Srour et al. 2020).

Srour et al. (2020) designed benzimidazoles attached to thiazole and hydrazinylbenzene groups at C2 and found that the EGFR inhibitory potency decreases with a decrease in polarity of the molecules. In particular, derivatives containing *p*-nitrophenyl substituent showed significant EGFR activity. The importance of phenyl group linked to C2 of benzimidazole derivatives is also notable in many other studies, in which better EGFR inhibitory activity was exhibited by the phenyl group compared to the aliphatic chain (Akhtar et al. 2017; Celik et al. 2019). Furthermore, electron-rich substitution at the *para*-position of a phenyl substituent was also found to increase the EGFR inhibitory activity (Akhtar et al. 2018; Labib et al. 2018). Several studies also relate the solubility of the compounds depends on the electronic property of the attached functional group. Furthermore, in many cases, the studies also keep the benzene side of benzimidazole unsubstituted while the few who did the opposite reported good EGFR inhibition. Therefore, insight into the relationship of these components and their binding activity on EGFR is imperative.

MATERIALS AND METHODS

PROTEIN AND LIGAND PREPARATION

The crystal structures for $EGFR_{wt}$ (PDB ID: 3VJO) and T790M mutant (PDB ID: 2JIT) were extracted from RSCB Protein Data Bank http://www.rcsb.org. The docking was carried out on monomeric form of the crystal structure. The AMP-PNP complexed with crystal structure was dismissed to leave a clean docking canvas. The PDB files of the protein and ligand were prepared using AutoDock Tools version 1.5.7 (ADT; Scripps Research Institute, La Jolla, San Diego, USA). The three-dimensional protein target was stripped off all water molecules. Polar hydrogens were integrated into the proteins to allow establishment of hydrogen bonding during docking. Kollman and Gasteiger charges were also added to reconstruct the molecular electrostatic potential. Kollman charges were computed based on quantum mechanics whereas Gasteiger charges were generated based on electronegativity equilibration. Fifty benzimidazole analogues were designed and sketched using Chemsketch following previous published work (Abdullah, Ali & Hamid 2021) and constructed as in Figure 1. The SMILES notations were then converted into 3D structures and auto-optimized based on Universal Force Field (UFF) using Avogadro. The SDF file of ATP and Gefitinib were extracted from PubChem (https://pubchem.ncbi.nlm.nih.gov) and prepared using Autodock Tools.



X = NH / CO

R_1 / R_2 = OCH_3 , NO_2 , SO_3F , SO_3H

FIGURE 1. The structure of the designed benzimidazole derivatives

MOLECULAR DOCKING

AutoDock Vina version 1.2.3 software was used to perform the molecular docking simulations (The Scripps Research Institute, La Jolla, San Diego, USA). Discovery Studio Biovia 2021 (Dassault Systèmes, San Diego, California, USA) was employed to visualize and modify receptor and ligand structures. All 50 benzimidazole analogues, ATP and Gefitinib were docked at the predicted binding site. We have also superimposed our docking results (3VJO with gefitinib) with 4WKQ (crystal structural of EGFR of kinase domain with gefitinib). The protein superimpose returns an RMSD value of 0.473, which is acceptably low, deeming the docking method as trustable (López-Camacho et al. 2016). The residues making up the binding and active site for human EGFR (P00533), extracted based on the UniProtKB database were found to be Lys745, Asp837, and Asp855 (Yun et al. 2008). The protein grid box, which is the active site for docking was set up using AutoDock Tools to enclose the aforementioned residues. The grid box dimensions for T790M mutant at 0.375 Å are 36 × 44 × 48 and centered at -14.508 × 28.058 × 28.036 while for $EGFR_{wt}$ at 0.375 Å are 40 × 60 × 34 and centered at 51.377 × 3.559 × -25.806. The docking simulations were then performed using AutoDock Vina, where the docking scores (in kcal/mol) were generated, and the binding poses ranked from the lowest to the highest according to the order of binding affinity. The energy range and exhaustiveness for the docking procedure were fixed at 4 and 32, respectively. The protein-ligand interactions were visualized with Biovia Discovery Studio Visualizer and the hydrogen and hydrophobic interactions within the protein-ligand complex were identified.

RESULTS AND DISCUSSION

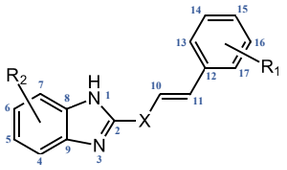
The main approach in the development of EGFR inhibitors is still based on the currently used drugs, employing both bioactivity and bioactivity analyses. Because of the diverse types of groups attached to different positions of the scaffold, structural insights of benzimidazole derivatives showed conflicting findings, especially on the superiority of electron withdrawing and donating groups in improving EGFR binding activity. A closer look at the literature on the docking studies of benzimidazole derivatives with the EGFR protein showed that hydrogen bonds are mainly centralised at the middle part of the structures having an amino or/and carbonyl group, and hydrophobic interaction at the phenyl group of the benzimidazole ring (Abdullah, Ali & Hamid 2021). Thus, we proposed to evaluate benzimidazoles containing an amino or a carbonyl group attached to C2, attached to a double bond linker and a phenyl group. In this study, we designed 50 keto- and amino-benzimidazole derivatives having electron withdrawing groups; nitro, sulfonyl fluoride, sulfonate and electron donating group; methoxy, grafted onto the *ortho-*, *meta-* and *para-* position of the phenyl moiety and on C4, C5 and C7 of the benzimidazole nucleus (Table 1). The structures were optimised and docked against the wild-type EGFR ($EGFR_{wt}$; PDB ID: 3VJO) and EGFR mutant (T790M; PDB ID: 2JIT).

The EGFR-benzimidazole contact analysis shows the analogues interact with the $EGFR_{wt}$ at the ATP and allosteric binding sites, while the binding pocket targeted within T790M mutant is only the ATP binding site (Figure 2). The binding affinity of the compounds with $EGFR_{wt}$ and T790M mutant is shown in Table 2.

The benzimidazole derivatives exhibited better binding affinity with T790M mutant than $EGFR_{wt}$. Eight derivatives with the lowest binding free energy with $EGFR_{wt}$ (-7.8 to -8.1 kcal/mol) occupied the allosteric binding site, suggesting the potential of these ligands as a Type II inhibitor. This observation is attributed mainly to the hydrogen bond interaction of Asp855 within the Asp-Phe-Gly (DFG) motif of the activation loop with the oxygen atoms of carbonyl linker, sulfonyl fluoride, sulfonate and nitro moieties. The DFG motif is responsible for directing Mg^{2+} into the catalytic site to interact with the phosphate groups in ATP (Amelia et al. 2022). The oxygen atom of sulfonyl fluoride and methoxy group also act as Michael acceptors and approach the Cys797 residue in the hinge region of T790M mutant for hydrogen bond formation (**7a**, **7c**, and **11c**). Inhibitors such as HKI-272 and AST1306 attribute their irreversibility to the covalent bond formations with Cys797, hence similar binding mode could potentially render the benzimidazole analogues irreversible and able to silence the tumour cells permanently (Xie et al. 2011; Yun et al. 2008).

The N1 of the imidazole ring becomes a hydrogen donor for Met793, Leu777 and principally Thr854, while the N3 of the ring frequently interacts with Lys745 and Leu703 through hydrogen bonding. On the other hand, benzimidazole nucleus and substituted phenyl hydrophobically interact with Val726, Phe723, Ala743, Leu844, Met790 and Leu718 of both receptors (Table 2). These interactions indicate the frequent contact of benzimidazole analogues with the important residues for ATP binding, namely Lys745, Leu718, Val726, Ala743 and Leu844. Blocking these residues from ATP could

TABLE 1. The designed benzimidazole derivatives



Ligand	X	R_1/R_2 position	R_1	R_2	Ligand	X	R_1/R_2 position	R_1	R_2
1a			$-OCH_3$	H	**7a**			$-OCH_3$	H
1b		C15 (*para*)	$-NO_2$	H	**7b**		C15 (*para*)	$-NO_2$	H
1c			$-SO_2F$	H	**7c**			$-SO_2F$	H
1d			$-SO_3H$	H	**7d**			$-SO_3H$	H
2a			$-OCH_3$	H	**8a**			$-OCH_3$	H
2b		C14 (*meta*)	$-NO_2$	H	**8b**		C14 (*meta*)	$-NO_2$	H
2c			$-SO_2F$	H	**8c**			$-SO_2F$	H
2d			$-SO_3H$	H	**8d**			$-SO_3H$	H
3a			$-OCH_3$	H	**9a**			$-OCH_3$	H
3b		C13 (*ortho*)	$-NO_2$	H	**9b**		C13 (*ortho*)	$-NO_2$	H
3c			$-SO_2F$	H	**9c**			$-SO_2F$	H
3d			$-SO_3H$	H	**9d**			$-SO_3H$	H
4a	NH		H	$-OCH_3$	**10a**			H	$-OCH_3$
4b		C5	H	$-NO_2$	**10b**	CO	C5	H	$-NO_2$
4c			H	$-SO_2F$	**10c**			H	$-SO_2F$
4d			H	$-SO_3H$	**10d**			H	$-SO_3H$
5a			H	$-OCH_3$	**11a**			H	$-OCH_3$
5b		C4	H	$-NO_2$	**11b**		C4	H	$-NO_2$
5c			H	$-SO_2F$	**11c**			H	$-SO_2F$
5d			H	$-SO_3H$	**11d**			H	$-SO_3H$
6a			H	$-OCH_3$	**12a**			H	$-OCH_3$
6b		C7	H	$-NO_2$	**12b**		C7	H	$-NO_2$
6c			H	$-SO_2F$	**12c**			H	$-SO_2F$
6d			H	$-SO_3H$	**12d**			H	$-SO_3H$
13		-	-	-	**14**		-	-	-

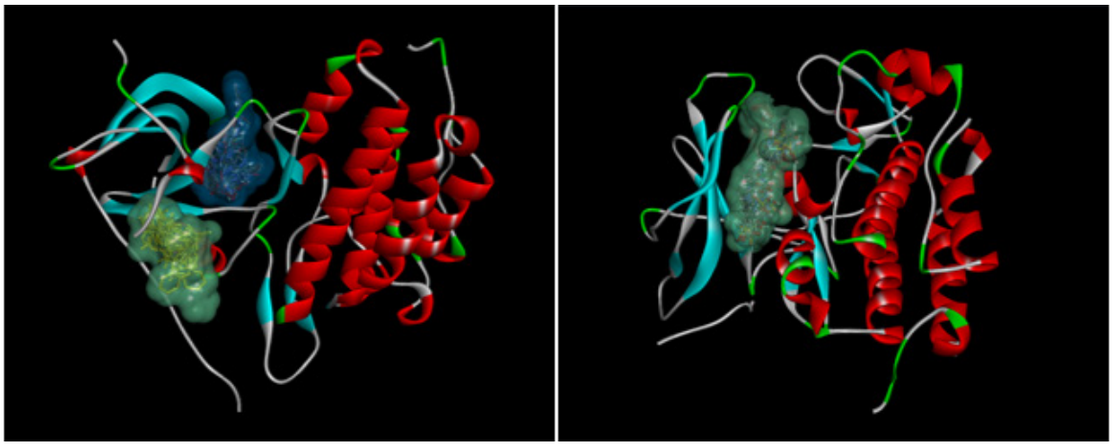


FIGURE 2. ATP (*green*) and allosteric (*blue*) binding sites for benzimidazole analogues in $EGFR_{wt}$ (*left*) and T790M mutant (*right*)

potentially suppress signal transduction pathways and cellular responses. The nitrogen atoms of benzimidazoles with carbonyl linker (**7a**, **7b**, **7c**, **7d**, **8a**, **8c** and **8d)** consistently form hydrogen bonds with Thr854 and Lys745 of $EGFR_{wt}$. The benzimidazole nucleus also forms π-anion interaction with Asp855 while the distal phenyl form π-interaction with Val726 and Ala743.

When the substituents are grafted onto the distal phenyl, the differences in the bond distance gradually declined across the ligand body towards the benzimidazole nucleus (Figure 3). This finding may be due to localised interaction, that is when there is more than one aromatic ring in a system, the substituent effect is prominent on the nearest vertex and not wholly on the other rings (Wheeler 2012). The results also showed that benzimidazoles with carbonyl linkers primarily dominate the top ten highest binding affinities in both EGFR-ligand complexes (90% for $EGFR_{wt}$ and 80% for T790M mutant).

TABLE 2. $EGFR_{wt}$-ligand and T790M-ligand contact analysis from highest to lowest binding affinity (kcal/mol)

$EGFR_{wt}$					T790M				
Ligand	Binding affinity (kcal/mol)	No of H bonds	H bond residues	Hydrophobic interacting residues	Ligand	Binding affinity (kcal/mol)	No of H bonds	H bond residues	Hydrophobic interacting residues
7c	-8.1	4	Lys745, Asp855 Thr854, Met793	Ala743 Val726	**1c**	-8.4	4	Asp837 Asp855 Thr854 Asn842	Lys875 Leu858,Val726
11c	-8.0	3	Lys745, Asp855	Leu718 Leu844, Val726, Ala743, Lys745	**7d**	-8.3	1	Asn842	Ala743, Met790 Leu844, Phe723
12c	-8.0	2	Asp855	Lys745 Val726, Leu718, Ala743, Leu844	**8c**	-8.3	4	Asp837 Asp855 Asn842 Thr854	Lys875, Leu858 Val726, Lys745
12d	-8.0	3	Lys745, Asp855	Leu718, Leu844 Val726, Ala743 Lys745	**10c**	-8.2	2	Lys745	Lys875, Ala722 Phe723, Val726

11d	-7.9	4	Lys745, Asp855 Glu762	Leu844, Leu718 Val726, Lys745 Ala743	**11c**	-8.2	2	Asn842 Cys797	Phe723, Leu844 Ala743
5d	-7.8	4	Asp855, Lys745	Leu718, Leu844 Ala743, Val726 Lys745	**4c**	-8.2	0	-	Ala743, Leu844 Phe723
6d	-7.8	4	Asp855, Lys745	Leu718, Leu844 Val726, Ala743	**7c**	-8.2	1	Cys797	Phe723, Lys745 Val726, Ala743 Leu844
8c	-7.8	4	Met793, Lys745 Thr854	Ala743, Val726	**10d**	-8.1	3	Arg836 Lys860	Leu858, Ile759 Ala755
9c	-7.8	3	Tyr1016 Ile1018, Leu777	Leu778, Ile1018	**12b**	-8.1	3	Thr854 Asn842 Asp855	Ala743, Leu844 Phe723
9d	-7.8	5	Arg776, Leu703	Arg705, Leu778 Arg776	**12d**	-8.1	2	Gln791 Asn842	Leu844, Ala743 Val726, Lys745
10c	-7.7	4	Leu703, Ile1018 Arg776	-	**5d**	-8.1	3	Arg841 Asp855	Val726, Phe723 Leu858, Lys875
11b	-7.7	3	Asp855, Lys745	Leu718, Val726 Ala743, Lys745 Leu788	**6b**	-8.1	0	-	Pro877, Lys875 Leu858, Lys745 Val726
12b	-7.7	2	Thr854, Asp855	Lys745, Val726 Ala743, Leu718	**7b**	-8.1	2	Asn842 Lys745	Ala743, Leu844 Met790, Phe723
5c	-7.7	3	Asp855, Lys745	Leu718, Leu844 Ala743, Val726 Lys745	**8d**	-8.1	0	-	Phe723,Val726 Leu718
7d	-7.7	4	Lys745, Asp855 Thr854, Met793	Val726, Ala743	**9b**	-8.1	2	Thr854 Asn842	Phe723, Lys745 Val726, Ala743 Leu844
8d	-7.7	3	Lys745, Thr854 Met793	Ala743, Val726	**10b**	-8.0	1	Lys745	Phe723, Val726 Leu718, Ala743 Leu844
9b	-7.7	4	Arg776, Leu703	Ile1018	**11d**	-8.0	2	Asn842 Asp855	Leu844 ,Ala743 Phe723
2c	-7.6	5	Val769, Arg776 Ala767, Ile1018	Ile1018, Ala702 Gln701	**12c**	-8.0	1	Asn842	Leu844, Phe723
8b	-7.6	2	Asp855, Met793	Lys745, Val726 Ala743, Thr790	**1d**	-8.0	1	Asp855	Ala743, Leu844 Val726, Phe723
10b	-7.5	2	Met793, Thr654	Leu844, Ala743 Val726	**2c**	-8.0	0	-	Leu844, Ala743 Val726, Phe723
10d	-7.5	2	Thr854, Met793	Val726, Leu844 Ala743	**2d**	-8.0	0	-	Ala743, Leu844 Met790, Val726 Phe723

1c	-7.5	4	Glu762, Lys745 Met793, Thr854	Ala743, Val726	**3c**	-8.0	0	-	Leu844, Val726 Ala743, Phe723 Lys745
2d	-7.5	3	Arg776, Leu703 Leu777	Arg705, Leu778	**6c**	-8.0	1	Asp855	Lys745, Leu858 Pro877, Lys875 Val726
3b	-7.5	4	Ile1018, Arg776 Leu703	Ala702, Leu778	**8b**	-8.0	2	Asp855 Thr854	Phe723, Val726 Leu844, Ala743
3c	-7.5	3	Arg776, Leu777 Leu703	Leu778	**9c**	-8.0	2	Thr854	Phe723, Lys745 Val726, Ala743 Leu844
3d	-7.5	4	Arg776, Leu777 Leu703	Arg705	**9d**	-8.0	0	-	Phe723, Lys745 Val726, Ala743 Leu844, Leu718
7b	-7.5	4	Lys745, Asp855 Thr854, Met793	Ala743, Val726	**11b**	-7.9	1	Asn842 Asp855	Leu844, Ala743 Phe723
5b	-7.4	1	Asp855	Leu718, Val726 Val743, Ala743	**1b**	-7.9	0	-	Ala743, Leu844 Met790, Val726 Phe723
6b	-7.4	1	Lys745	Ala743, Leu718 Leu844, Lys745	**3d**	-7.9	2	Asp855 Thr854	Val726, Phe723 Lys745, Leu844 Ala743, Leu718
1d	-7.3	3	Lys745, Thr854 Glu762	Val726, Ala743	**5c**	-7.9	1	Asp855	Leu844, Ala743 Met790, Phe723
4b	-7.3	2	Thr790, Met793	Val726, Lys745 Ala743, Cys775 Leu844	**3b**	-7.8	3	Asp855 Thr854 Asn842	Phe723, Leu844 Ala743
6c	-7.3	2	Leu777, Leu703	Leu778	**4d**	-7.8	0	-	Phe723, Leu844 Ala743
ATP	-7.3	13	Thr790, Thr854 Asp855, Lys745 Arg841, Asn842 Met793	Leu844	**6d**	-7.8	1	Asn842	Leu844, Ala743 Phe723
[a]GEF	-7.3	1	Lys745	Ala743, Val726 Leu844, Cys797	**7a**	-7.8	1	Cys797	Lys745, Val726 Leu844, Leu718 Cys797
2b	-7.2	3	Met793, Thr854	Leu718, Val726 Ala743	**10a**	-7.7	1	Lys745	Val726, Phe723 Leu718, Leu844
4c	-7.2	3	Arg705, Leu703 Ala767	Ala702, Leu703 Arg705	**14**	-7.7	0	-	Lys745, Val726 Leu718, Leu792 Leu844
4d	-7.2	5	Ile1018 Tyr1016 Arg776, Ala767 Leu703	Gln701, Ala702	**11a**	-7.7	1	Asn842	Ala743, Leu844 Met790, Leu792 Phe723

9a	-7.2	3	Arg776	Ile1018	**12a**	-7.7	1	Asn842	Ala743, Leu844 Met790, Leu792 Leu718, Phe723
10a	-7.1	1	Met793	Val726, Ala743 Leu844, Leu788 Met766	**3a**	-7.7	0	-	Leu844, Ala743 Met790, Val726 Phe723
12a	-7.1	2	Arg776, Leu777	Leu1017 Leu778, Ile1018 Arg705, Leu703	**13**	-7.6	0	-	Leu844, Ala743 Met790, Val726 Phe723
11a	-7.0	2	Arg776 Tyr1016	Ile1018, Leu703 Arg705	**2b**	-7.6	1	Lys745	Phe723, Val726 Ala743, Leu844
13	-7.0	1	Thr854	Leu788, Thr790 Ala743, Leu718 Val726	**5b**	-7.6	0	-	Lys745, Val726 Leu792, Leu718 Leu844
14	-7.0	0	-	Leu788, Ala743 Thr790, Lys745	**8a**	-7.6	0	-	Phe723, Lys745 Val726, Leu844 Ala743, Leu718 Cys797
5a	-7.0	4	Ile1018, Leu778 Leu777, Leu703	Leu1017 Leu778	[b]GEF	-7.6	1	Lys745	Val726, Lys745 Phe723, Leu858
7a	-7.0	3	Lys745, Asp855 Thr854	Met793, Leu844 Val726, Ala743	**1a**	-7.5	1	Asp855	Lys875, Pro877 Leu858, Lys745 Val726
8a	-7.0	4	Lys745, Asp855 Thr854	Val726, Ala743 Met793, Leu844	**4a**	-7.5	2	Lys745, Thr854	Leu792, Leu844 Phe723
1b	-6.8	3	Lys745, Thr854 Glu762	Val726, Ala743	**4b**	-7.5	1	Lys745	Leu844, Ala743 Phe723, Val726 Lys745
6a	-6.8	1	Thr790	Lys745, Val726 Ala743, Leu718 Leu844	**9a**	-7.5	0	-	Phe723, Lys745 Val726, Leu844 Ala743, Leu718
1a	-6.7	1	Thr854	Lys745, Val726 Thr790, Ala743 Leu718	**6a**	-7.4	1	Asp855	Leu718, Leu792 Leu844, Met790 Ala743, Phe723
2a	-6.5	4	Thr790, Met793 Pro794	Ala743, Leu844 Met766 Val726, Leu718	**5a**	-7.3	1	Asp855	Phe723 Met790, Leu844 Leu792, Leu718
3a	-6.5	2	Leu777, Leu703	Ala702, Arg705 Leu778	**2a**	-7.2	0	-	Phe723, Val726 Met790, Leu844 Ala743
4a	-6.5	0	-	Leu718, Leu844 Ala743, Val726	ATP	-7.2	2	Asn842 Gln791	Leu844, Met790 Ala743

[a]GEF = Gefitinib

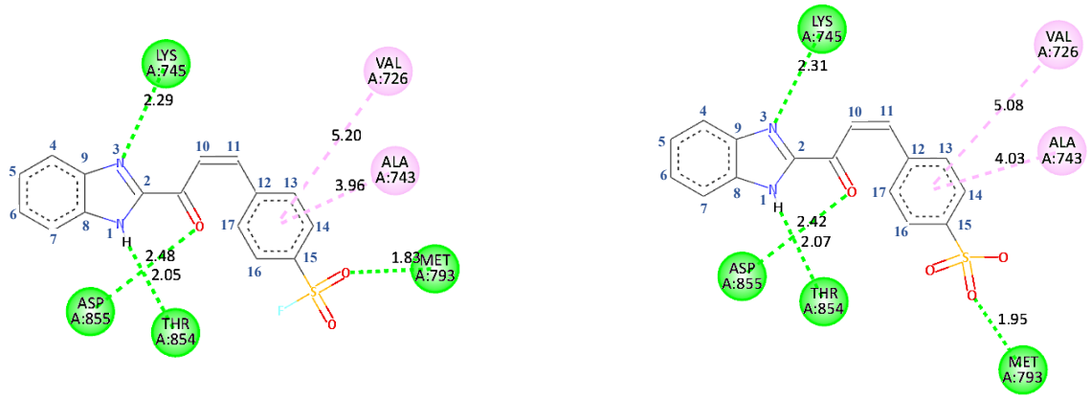


FIGURE 3. Interactions of (a) $EGFR_{wt}$-**7c** (left) and (b) $EGFR_{wt}$-**7d** (right)

The binding energies of unsubstituted amino- (**13**) and keto-benzimidazole (**14**) are -7.0 kcal/mol for $EGFR_{wt}$ and -7.6 and -7.7 kcal/mol, respectively, for T790M mutant. The presence of sulfonyl substituents significantly gave good protein-ligand complex binding affinities as shown by the lower binding free energies (more stable complexes) compared to others; -8.1 (**7c**) and -7.8 (**11c**) kcal/mol in $EGFR_{wt}$ complex, and -8.3 (**7d**) and -8.4 (**1c**) kcal/mole for T790M mutant. These observations may be attributed to the strong electron withdrawing effect of both moieties. Sulfonyl fluoride is expected to reduce the electron density of its phenyl more than sulfonate as fluorine is more electronegative than oxygen, and thus attenuates the π-alkyl interactions between the aromatic ring of benzimidazole, and Val726 and Ala743 of the protein (Liégeois et al. 2014; Neel et al. 2017). In this study, however, the fluorine effect is not prominent enough as the binding free energies for both sulfonyl fluoride and sulfonate are quite similar. Nitro substituent, being slightly weaker electron withdrawing than sulfonyl fluoride and sulfonate showed lower binding affinity when docked against both EGFR receptors. Methoxy is the only substituent exerting electron donating effect among the substituents and consistently gave the highest binding free energies in the range of -6.5 to -7.2 kcal/mol in $EGFR_{wt}$ and -7.2 to -7.7 kcal/mol for T790M mutant. Hence, we postulated that electron withdrawing groups contribute to the protein-ligand interaction better than electron donating groups. In addition, the binding affinity also correlates with the strength of the electron withdrawing group ($-SO_2F$ = $-SO_3H$ > $-NO_2$ > H > $-OCH_3$). Interestingly, no correlation was observed between the binding affinity of EGFR-ligand complex with the C4, C5 and C7 position of benzimidazole and the *para*, *ortho*- and *meta*- position of the substituted phenyl.

COMPARISON OF DOCKING RESULT FROM BENZIMIDAZOLE, ATP AND GEFITINIB

Despite having 13 hydrogen bonds with distances ranging from 2.07 to 3.05 Å, the ATP-$EGFR_{wt}$ complex gave a binding energy of only -7.3 kcal/mol, while benzimidazole analogues docked against the same protein showed a maximum of -8.1 kcal/mol binding free energy with only four hydrogen bonds at 1.83 to 2.48 Å distance range. Similarly, the significantly different binding energy is also emulated in the ATP-T790M mutant complex when compared with the benzimidazole derivatives against the T790M mutant. This deviation signifies that binding affinity does not necessarily correlate with the number of hydrogen bonds. Instead, the lower binding free energy of the benzimidazole analogues could be attributed to the surface shape complementarity that is governed by the Van der Waals interactions of the docked ligands (Tsai et al. 2001). Although hydrogen bond interactions played a major role in stabilising the ligands at the binding interface, the complementary shape between proteins and their ligands is also critical for predicting protein-ligand binding affinities as it governs the attraction and repulsion between non-bonding atoms (Hsu, Chen & Yang 2008).

Unlike ATP, the benzimidazole derivatives mostly adopt a conformation that complements the shape of the binding pocket (Figure 4). As torsional degrees of freedom were not assigned to the protein and ligands during the docking procedure, it is assumed that the interactions in this study follow a 'lock-and-key'

model instead of the induced fit model. Hence, the complementary shape of the benzimidazoles would be a better 'key' for the binding pocket compared to ATP.

On the whole, benzimidazole analogues displayed lower binding free energy than Gefitinib. The binding affinity difference between Gefitinib and the benzimidazole analogues scales at 0.5~0.8 kcal/mol in both $EGFR_{wt}$ and T790M mutant. This observation could also be associated to the shape of the ligands within the receptors, as shown in Figure 5. Benzimidazoles were shown to possess a more complementary shape for the binding pocket of $EGFR_{wt}$ and T790M mutant when compared to Gefitinib.

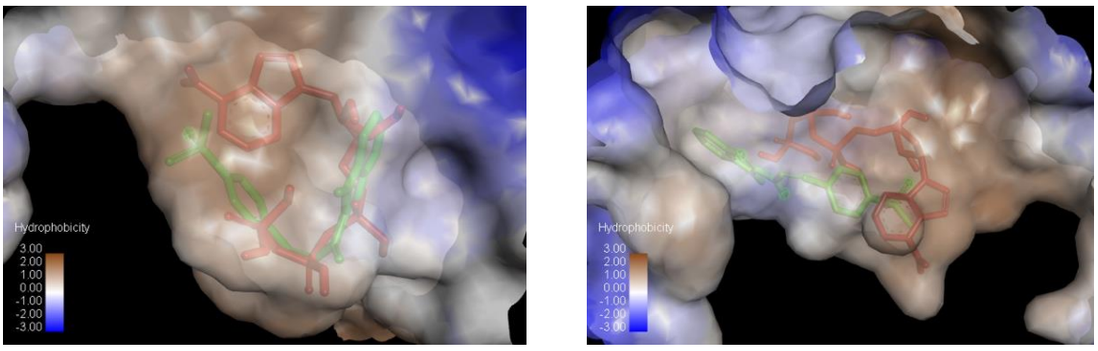


FIGURE 4. Comparison of adopted conformation by benzimidazole analogue (*green*) and ATP (*red*) after being docked against (a) $EGFR_{wt}$ and (b) T790M mutant

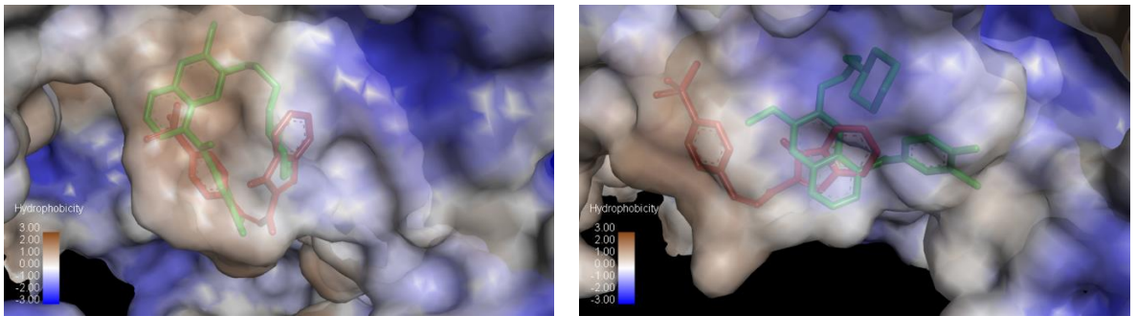


FIGURE 5. Comparison of adopted conformation by benzimidazole analogue (*red*) and Gefitinib (*green*) after being docked against (a) $EGFR_{wt}$ and (b) T790M mutant

CONCLUSIONS

Molecular docking analysis was used to evaluate the binding activity of 50 designed benzimidazole analogues toward $EGFR_{wt}$ and T790M mutant active sites. The analysis showed strong binding affinities of the derivatives against the target receptors, ranging from -4.77 to -8.30 kcal mol^{-1}. Our docking study also showed keto-benzimidazoles exhibiting the top ten

highest binding affinities against both EGFR active sites, ranging from -7.8 to -8.1 kcal/mol (for $EGFR_{wt}$) and -8.1 to -8.4 (in T790M). The top ten compounds also showed better binding affinity compared to Gefitinib and ATP in both active sites. The N1 of the imidazole ring acted as a hydrogen donor for Thr854, Met793 and Leu777, while Lys745 and Leu703 interacted with N3 via hydrogen bonds. The benzimidazole scaffold and substituted phenyl moiety interacted hydrophobically with Val726, Phe723, Ala743, Leu844, Met790 and Leu718 in both receptors. The strong electron withdrawing sulfonyl substituent contributed significantly to the protein-ligand complex binding affinity. However, the complementarity shape of the benzimidazoles with the binding pocket of $EGFR_{wt}$ and T790M mutant also played a pivotal role in regulating the binding free energies. No correlation was seen between the binding free energy with the C4, C5 and C7 positions of benzimidazole and the *para*-, *ortho*- and *meta*- positions of the substituted phenyl. Intermolecular steric clashes were noticed in some complexes, namely **2a**, **3b**, **4d**, **5c**, **5d**, **6b**, **6d**, **8b**, **9d**, **10c**, **10d,** and **12c** for $EGFR_{wt}$-benzimidazole, and **5a**, **5b**, **6a**, **6b,** and **9d** for T790M-benzimidazole. As the flexibility of both protein and ligand were not taken into account, their rigidity causes the ligands' conformation to not fit perfectly into the binding pocket. Employing molecular dynamics could prevent these clashes and provide better binding poses as the software allows some torsional degree of freedom. Although studies on the design of benzimidazoles as EGFR inhibitors have been presented by many authors, the matter is still insufficiently explored. This study could assist chemists in the design and synthesis of this class of benzimidazoles toward finding the next generation of EGFR inhibitors. The compounds can be further validated by *in vitro* and *in vivo* assays.

ACKNOWLEDGEMENTS

We are grateful to the Ministry of Higher Education (MOHE), Malaysia for the Fundamental Research Grant Scheme (FRGS/1/2018/STG01/UIAM/02/1). We would also like to thank Nurasyikin Hamzah for the useful discussion.

REFERENCES

Abdullah, M.N., Ali, Y. & Hamid, S.A. 2021. Insights into the structure and drug design of benzimidazole derivatives targeting the epidermal growth factor receptor (EGFR). *Chemical Biology & Drug Design* 100(6): 921-934.

Akhtar, M.J., Khan, A.A., Ali, Z., Dewangan, R.P., Rafi, M., Hassan, M.Q., Akhtar, M.S., Siddiqui, A.A., Partap, S., Pasha, S. & Yar, M.S. 2018. Synthesis of stable benzimidazole derivatives bearing pyrazole as anticancer and EGFR receptor inhibitors. *Bioorganic Chemistry* 78: 158-169.

Akhtar, M.J., Siddiqui, A.A., Khan, A.A., Ali, Z., Dewangan, R.P., Pasha, S. & Yar, M.S. 2017. Design, synthesis, docking and QSAR study of substituted benzimidazole linked oxadiazole as cytotoxic agents, EGFR and erbB2 receptor inhibitors. *European Journal of Medicinal Chemistry* 126: 853-869.

Amelia, T., Kartasasmita, R.E., Ohwada, T. & Tjahjono, D.H. 2022. Structural insight and development of EGFR tyrosine kinase inhibitors. *Molecules* 27(3): 819.

Bansal, Y. & Silakari, O. 2012. The therapeutic journey of benzimidazoles: A review. *Bioorganic & Medicinal Chemistry* 20(21): 6208-6236.

Celik, I., Ayhan-Kılcıgil, G., Guven, B., Kara, Z., Gurkan-Alp, A.S., Karayel, A. & Onay- Besikci, A. 2019. Design, synthesis and docking studies of benzimidazole derivatives as potential EGFR inhibitors. *European Journal of Medicinal Chemistry* 173: 240-249.

Çevik, U.A., Celik, I., Mella, J., Mellado, M., Özkay, Y. & Kaplancikli, A.Z. 2022. Design, synthesis, and molecular modeling studies of a novel benzimidazole as an aromatase inhibitor. *ACS Omega* 7(18): 16152-16163.

Grabe, T., Lategahn, J. & Rauh, D. 2018. C797S resistance: The undruggable EGFR mutation in non-small cell lung cancer? *ACS Medicinal Chemistry Letters* 9(8): 779-782.

Hsu, K-C., Chen, Y-F. & Yang, J-M. 2008. Binding affinity analysis of protein-ligand complexes. *2008 2nd International Conference on Bioinformatics and Biomedical Engineering.* pp. 167-171.

Jiang, T., Su, C., Ren, S., Cappuzzo, F., Rocco, G., Palmer, J.D., van Zandwijk, N., Blackhall, F., Le, X., Pennell, N.A. & Zhou, C. 2018. A consensus on the role of Osimertinib in non-small cell lung cancer from the AME Lung Cancer Collaborative Group. *Journal of Thoracic Disease* 10(7): 3909-3921.

Karadayi, F.Z., Yaman, M., Kisla, M.M., Keskus, A.G., Konu, O. & Ates-Alagoz, Z. 2020. Design, synthesis and anticancer/antiestrogenic activities of novel indole-benzimidazoles. *Bioorganic Chemistry* 100: 103929.

Lee, Y., Kim, T.M., Kim, D.W., Kim, S., Kim, M., Keam, B., Ku, J.L. & Heo, D.S. 2019. Preclinical modelling of Osimertinib for NSCLC with EGFR Exon 20 insertion mutations. *Journal of Thoracic Disease* 14(9): 1556-1566.

Lelais, G., Epple, R., Marsilje, T.H., Long, Y.O., McNeill, M., Chen, B., Lu, W., Anumolu, J., Badiger, S., Bursulaya, B., DiDonato, M., Fong, R., Juarez, J., Li, J., Manuia, M., Mason, D.E., Gordon, P., Groessl, T., Johnson, K., Jia, Y., Kasibhatla, S., Li, C., Isbell, J., Spraggon, G., Bender, S. &

Michellys, P-Y. 2016. Discovery of (*R*,*E*)-*N*-(7-Chloro-1-(1-[4-(dimethylamino)but-2-enoyl]azepan-3-yl)-1H-benzo[d] imidazol-2-yl)-2-methylisonicotinamide (EGF816), a novel, potent, and WT sparing covalent inhibitor of oncogenic (L858R, ex19del) and resistant (T790M) EGFR mutants for the treatment of EGFR mutant non-small-cell lung cancers. *Journal of Medicinal Chemistry* 59(14): 6671-6689.

Labib, M.B., Philoppes, J.N., Lamie, P.F. & Ahmed, E.R. 2018. Azole-hydrazone derivatives: Design, synthesis, *in vitro* biological evaluation, dual EGFR/HER2 inhibitory activity, cell cycle analysis and molecular docking study as anticancer agents. *Bioorganic Chemistry* 76: 67-80.

Leonetti, A., Sharma, S., Minari, R., Perego, P., Giovannetti, E. & Tiseo, M. 2019. Resistance mechanisms to Osimertinib in EGFR-mutated non-small cell lung cancer. *British Journal of Cancer* 121: 725-737.

Liégeois, J-F., Lespagnard, M., Salas, E.M., Mangin, F., Scuvée-Moreau, J. & Dilly, S. 2014. Enhancing a CH−π interaction to increase the affinity for 5-HT1A receptors. *ACS Medicinal Chemistry Letters* 5(4): 358-362.

López-Camacho, E., García-Godoy, M.J., García-Nieto, J., Nebro, A.J. & Aldana-Montes, J.F. 2016. A new multi-objective approach for molecular docking based on RMSD and binding energy. In *Algorithms for Computational Biology.* Lecture Notes in Computer Science, vol 9702, edited by Botón-Fernández, M., Martín-Vide, C., Santander-Jiménez, S. & Vega-Rodríguez, M. Springer, Cham.

Mathur, G., Nain, S. & Sharma, P.K. 2015. Cancer: An overview. *Academic Journal of Cancer Research* 8(1): 1-9.

Mostafa, A.S., Gomaa, R.M. & Elmorsy, M.A. 2019. Design and synthesis of 2-phenyl benzimidazole derivatives as VEGFR-2 inhibitors with anti-breast cancer activity. *Chemical Biology Drug Design* 93(4): 454-463.

Neel, A.J., Hilton, M.J., Sigman, M.S & Toste, F.D. 2017 Exploiting non-covalent π interactions for catalyst design. *Nature* 543(7647): 637-646.

Nussbaumer, S., Bonnabry, P., Veuthey, J.L. & Fleury-Souverain, S. 2011. Analysis of anticancer drugs: A review. *Talanta* 85(5): 2265-2289.

Rebucci, M. & Michiels, C. 2013. Molecular aspects of cancer cell resistance to chemotherapy. *Biochemical Pharmacology* 85(9): 1219-1226.

Shrivastava, N., Naim, M.J., Alam, M.J., Nawaz, F., Ahmed, S. & Alam, O. 2017. Benzimidazole scaffold as anticancer agent: Synthetic approaches and structure-activity relationship. *Archiv. Der. Pharmazie.* 350: e201700040.

Song, C. & Yang, X. 2022. Osimertinib-centered therapy against uncommon epidermal growth factor receptor-mutated non-small-cell lung cancer-A mini review. *Frontiers in Oncology* 12: 834585.

Srour, A.M., Ahmed, N.S., Abd El-Karim, S.S., Anwar, M.M. & El-Hallouty, S.M. 2020. Design, synthesis, biological evaluation, QSAR analysis and molecular modelling of new thiazol-benzimidazoles as EGFR inhibitors. *Bioorganic Medicinal Chemistry* 28(18): 115657.

Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A. & Bray, F. 2021. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer Journal for Clinicians* 71(3): 209-249.

The UniProt Consortium. 2019. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Research* 47: D506-D515.

Thomas, R. & Weihua, Z. 2019. Rethink of EGFR in cancer with its kinase independent function on board. *Frontiers in Oncology* 9: 800.

Troiani, T., Napolitano, S., Corte, C.M.D., Martini, G., Martinelli, E., Morgillo, F. & Ciardiello, F. 2016. Therapeutic value of EGFR inhibition in CRC and NSCLC: 15 years of clinical evidence. *ESMO Open* 1(5): e000088.

Tsai, C-J., Norel, R., Wolfson, H.J., Maizel, J.V. & Nussinov, R. 2001. Protein-ligand interactions: Energetic contributions and shape complementarity. *Encyclopedia of Life Sciences.* pp. 1-8.

Vyse, S. & Huang, P.H. 2019. Targeting EGFR exon 20 insertion mutations in non-small cell lung cancer. *Signal Transduction and Targeted Therapy* 4: 5.

Wheeler, S.E. 2012. Understanding substituent effects in noncovalent interactions involving aromatic rings. *Accounts of Chemical Research* 46(4): 1029-1038.

Xie, H., Lin, L., Tong, L., Jiang, Y., Zheng, M., Chen, Z., Jiang, X., Zhang, X., Ren, X., Qu, W., Yang, Y., Wan, H., Chen, Y., Zuo, J., Jiang, H., Geng, M. & Ding, J. 2011 AST1306, A novel irreversible inhibitor of the epidermal growth factor receptor 1 and 2, exhibits antitumor activity both *in vitro* and *in vivo*. *PLoS ONE* 6(7): e21487.

Yun, C-H., Mengwasser, K.E., Toms, A.V., Woo, M.S., Greulich, H., Wong, K-K., Meyerson, M. & Eck, M.J. 2008. The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *PNAS* 105(6): 2070-2075.

Zandi, R., Larsen, A.B., Andersen, P., Stockhausen, M-T. & Poulsen, H.S. 2007. Mechanisms for oncogenic activation of the epidermal growth factor receptor. *Cell Signal* 19(10): 2013-2023.

*Corresponding author; email: shafida@iium.edu.my