Article

Molecular Modeling Studies of *N*-phenylpyrimidine-4-amine Derivatives for Inhibiting FMS-like Tyrosine Kinase-3

Suparna Ghosh [1], **Seketoulie Keretsu** [1] **and Seung Joo Cho** [1,2,*]

[1] Department of Biomedical Sciences, College of Medicine, Chosun University, Gwangju 501-759, Korea; s.ghosh@chosun.kr (S.G.); keretsu@chosun.kr (S.K.)
[2] Department of Cellular and Molecular Medicine, College of Medicine, Chosun University, Gwangju 501-759, Korea
[*] Correspondence: chosj@chosun.ac.kr; Tel.: +82-62-230-7482 or +82-11-479-1010

Abstract: Overexpression and frequent mutations in FMS-like tyrosine kinase-3 (FLT3) are considered risk factors for severe acute myeloid leukemia (AML). Hyperactive FLT3 induces premature activation of multiple intracellular signaling pathways, resulting in cell proliferation and anti-apoptosis. We conducted the computational modeling studies of 40 pyrimidine-4,6-diamine-based compounds by integrating docking, molecular dynamics, and three-dimensional structure–activity relationship (3D-QSAR). Molecular docking showed that K644, C694, F691, E692, N701, D829, and F830 are critical residues for the binding of ligands at the hydrophobic active site. Molecular dynamics (MD), together with Molecular Mechanics Poison–Boltzmann/Generalized Born Surface Area, i.e., MM-PB(GB)SA, and linear interaction energy (LIE) estimation, provided critical information on the stability and binding affinity of the selected docked compounds. The MD study suggested that the mutation in the gatekeeper residue F691 exhibited a lower binding affinity to the ligand. Although, the mutation in D835 in the activation loop did not exhibit any significant change in the binding energy to the most active compound. We developed the ligand-based comparative molecular field analysis (CoMFA) and comparative molecular similarity index analysis (CoMSIA) models. CoMFA ($q^2 = 0.802$, $r^2 = 0.983$, and $Q_{F3}^2 = 0.698$) and CoMSIA ($q^2 = 0.725$, $r^2 = 0.965$ and $Q_{F3}^2 = 0.668$) established the structure–activity relationship (SAR) and showed a reasonable external predictive power. The contour maps from the CoMFA and CoMSIA models could explain valuable information about the favorable and unfavorable positions for chemical group substitution, which can increase or decrease the inhibitory activity of the compounds. In addition, we designed 30 novel compounds, and their predicted pIC_{50} values were assessed with the CoMSIA model, followed by the assessment of their physicochemical properties, bioavailability, and free energy calculation. The overall outcome could provide valuable information for designing and synthesizing more potent FLT3 inhibitors.

Keywords: FMS-like tyrosine kinase-3; acute myeloid leukemia; MD simulation; binding free energy; 3D-QSAR; CoMFA; CoMSIA

1. Introduction

FLT3 belongs to the type III tyrosine kinase receptor together with KIT, FMS, and platelet-derived growth factor receptor (PDGFR), which are involved in the differentiation, proliferation, and survival of hematopoietic progenitor cells [1]. It is expressed by the stromal cells of bone marrow, placenta, and CD34+ cells, as well as the myeloid and B lymphoid cell lineage [2]. The FLT3 structure comprises five extracellular domains similar to immunoglobulin (Ig) at the N-terminal, followed by a single transmembrane (TM) domain, a cytoplasmic juxtamembrane domain (JMD), and a tyrosine kinase domain (TKD) separated by a kinase insert (KI). An intracellular domain was located at the C-terminal end, as shown in Figure 1a [3–5]. FL is the endogenous ligand of FLT3, which is also expressed in bone marrow stromal cells, exists in a soluble form, or is bound to the membrane. Predominantly, FLT3 is found in an unbound form as an unphosphorylated monomer

coordinated by an inactive kinase domain. The JM domain interacts with the KD to block the ATP binding to the active site. After binding to the ligand FL in the extracellular Ig domains, FLT3 undergoes dimerization and exposes its dimerized domain [6–8]. This event triggers the activation of tyrosine kinase, followed by the phosphorylation of the FLT3 intracellular domain at its various sites. The phosphorylation of FLT3 propagates multiple intracellular signaling pathways, which are essential for phospholipid metabolism, transcription, differentiation, proliferation, cell survival, and apoptosis [9].
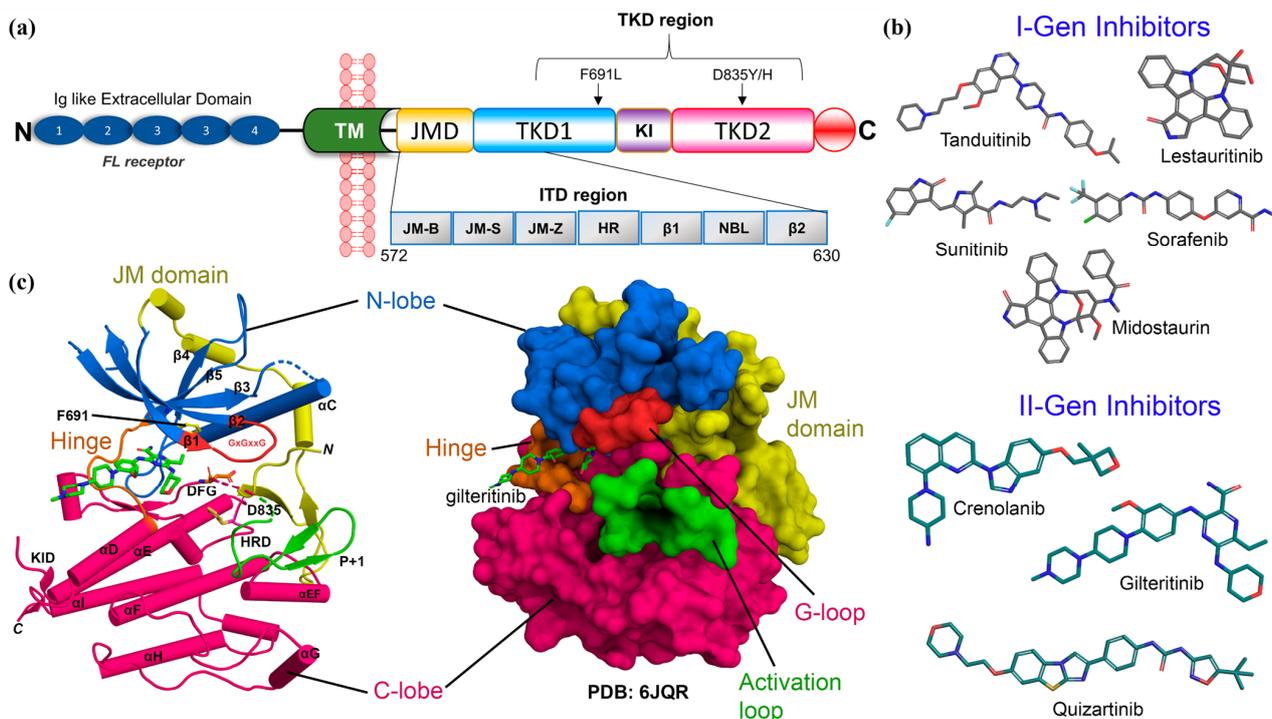

Figure 1. Domain map, inhibitors, and X-ray structure of FLT3. (**a**) Full-length FLT3 contains five immunoglobulins (Ig) like domains at the N-terminal, a transmembrane (TM) domain, and a juxtamembrane domain (JMD), followed by the tyrosine kinase domain (TKD) with a kinase insert (KI) and a C-terminal intracellular domain. F691L and D835Y/H are the two widely reported mutations in the TKD region, reported for 5% of the total AML cases. (**b**) FDA-approved first generation (I-Gen) and second generation (II-Gen) drugs as FLT3 inhibitors in grey and deep teal C atom color scheme. (**c**) X-ray structure of the gilteritinib-bound FLT3 (PDB: 6JQR). The various subregions of FLT3 are shown in cartoon and surface representation.

FLT3 is overexpressed in patients with AML, which is considered an aggressive hematologic malignancy. Active mutations in FLT3 have been reported in ~30% of total AML cases. These mutations can be subdivided into internal tandem duplicates (FLT3-ITD) and point mutations in the tyrosine kinase domain (FLT3-TKD), which are approximately 25% and 5% of the total AML, respectively [10,11]. Therefore, inhibition of FLT3 is an ideal therapeutic choice. Many inhibitors have been subjected to preclinical and clinical trials and have shown promising results. As illustrated in Figure 1b, tyrosine kinase inhibitors such as Tandutinib, Sunitinib, Midostaurin, Lestaurtinib, and Sorafenib were used as first-generation FLT3 inhibitors [12]. Due to the lack of sensitivity and selectivity of first-generation inhibitors, second-generation inhibitors, such as Gilteritinib, Quizartinib, and Crenolanib, have been approved [13,14]. Although, their poor bioavailability and off-target effects often increase drug toxicity in patients, which remains a concern. Patients undergoing AML treatment have often developed resistance to lead compounds through on-target secondary mutations in the kinase domain [13]. Two of the most common mutations have been found in the gatekeeper residue F691L and the activation loop D835Y, as shown in Figure 1c. Mutations in residues I836, D839, and Y842 are also found in the FLT3 kinase domain in patients with AML [14].

Inevitably, the rational development of new FLT3 inhibitors based on existing lead compounds is an ideal choice for achieving therapeutic efficacy in AML. Computer-aided drug design (CADD) has emerged as a promising tool for discovering new FLT3 inhibitors [15]. In a prior study, Bensinger et al. [16] used virtual screening and docking to identify lead compounds that were covalently bound to the DGF-in conformer of FLT3 receptor. These lead compounds were chemically modified to increase cytotoxicity and inhibitory efficacy against wild-type and mutant (D835Y) FLT3, suggesting that they could be a suitable starting point for discovering irreversible inhibitors [16]. In another study by Smith et al. [17], residue D835 plays a critical role in maintaining the DFG-out configuration by acting as an amino terminal capping residue for the αC-helix and serving as an essential space for type II inhibitor binding. However, the mutations in the D835 remain sensitive to type I inhibitors.

In this paper, we performed computational modeling studies such as docking, molecular dynamics (MD), free energy calculation, and three-dimensional structure-activity relationship (3D-QSAR) to a series of 40 pyrimidine-4,6-diamine derivatives, which are reported as type II-like inhibitors of FLT3 by Bharate et al. [18]. In their docking study, the most active compound 13a bound to the hinge loop by forming two H-bond interactions with C694. The compounds exhibited a diverse inhibitory activity range (IC_{50} 13.9 nM-15111 nM) against FLT3. The molecular docking and molecular dynamics revealed the critical interactions with the inhibitors inside the binding pocket. We calculated the MM-PB(GB)SA and LIE binding energy terms to evaluate the protein–ligand binding affinity. We manually induced the F691L and D835Y mutations in FLT3 and performed MD simulations to understand the mutagenic effect in terms of binding affinity to the inhibitor compound. Finally, we conducted CoMFA and CoMSIA studies of the 40 compounds to produce the 3D-QSAR contour map and established the structure–activity relationship. The CoMFA and CoMSIA contour maps described how modification of chemical groups could enhance the inhibitory activity of the compounds.

2. Results and Discussion

2.1. Molecular Docking Analysis

As we conducted the MD simulation and 3D-QSAR study based on the docking pose, the verification of docking reliability is an important step. There were no compounds from the dataset available with the co-crystallized FLT3 form in the PBD database. Thus, we selected the *N*-phynylpyrimidine-4-amine substructure of the docked compounds and compared it with the existing FLT3 conjugated crystal ligand FF-10101 (PDB ID: 5X02) [19], and Quizartinib (PDB ID: 4XUF) [20], and AWO with C-kit (PDB ID: 6ITT) [21] by overlapping them in the ligRMSD server provided in Supplementary Table S1. Since FLT3 and C-kit both are members of the tyrosine kinase family, we opted for the AWO bound C-kit to compare the ligand interactions. The FF-10101 formed the critical H-bond interaction with residue C694 by the amine(-NH_2) group of *N*-phenylpyrazine-2-amine. For the ligand Quizartinib, π–π interactions were found between its aniline ring and residue F691 and residue F830. We found the critical H-bond interaction between C694 and -NH_2 of the *N*-(Pyrimidin-4-yl)thiazol-2-amine moiety, which anchored the ligand to the hinge loop. Similar interactions were observed with the docked compounds M01, M03, and M17. Compounds M20, M24, and M34 were anchored to the hinge loop by forming H-bond interactions between the residue C694 and the *N*-phenylpyrimidine-4-amine-NH_2 group and the π–π stacking with phenylalanine (F691). The final docked pose was selected from the lowest binding energy cluster, which also shared the lowest RMSD with the crystal substructure according to the ECIDALs norms [22,23]. The final RMSD values of the docked compounds from the crystal ligands were arranged and are shown in Supplementary Table S1. The RMSD values suggest the overall docking reliability of the selected compounds. The BEs from the docking study were estimated to be −11.31 kcal/mol, −11.68 kcal/mol, −9.88 kcal/mol, −10.54 kcal/mol, −9.68 kcal/mol, and −10.09 kcal/mol for compounds M01, M03, M17, M20, M24, and M34, respectively.

The 2D protein–ligand interaction is illustrated in Figure 2a. The surrounding residues within 3.5 Å are shown using a color scheme, which is based on the chemical properties of the amino acids. Figure 2b depicts the H-bond, π–π, and π–anion interactions between M01 and residues inside the ATP pocket of FLT3. With residue C694, two H-bond interactions were found, while a third H-bond interaction was observed with catalytic lysine K644. The gatekeeper residue E692 formed the π–anion interaction with the pyrimidine ring. Another gatekeeper residue F691 accomplished the π–π stacking with the phenyl ring of M01. Other residues, such as L616, M665, V675, L856, L576, and I801, participated in hydrophobic interaction inside the hydrophobic cleft between the N lobe and the C lobe. The docked complexes of compounds M03, M17, M20, M24, and M34, which also comply with the ECIDALs norms during the docking process, are shown in Supplementary Figure S2 by 2D interaction diagrams. The best performing protein–ligand single-docked coordinate of each complex was taken for further MD simulation study.
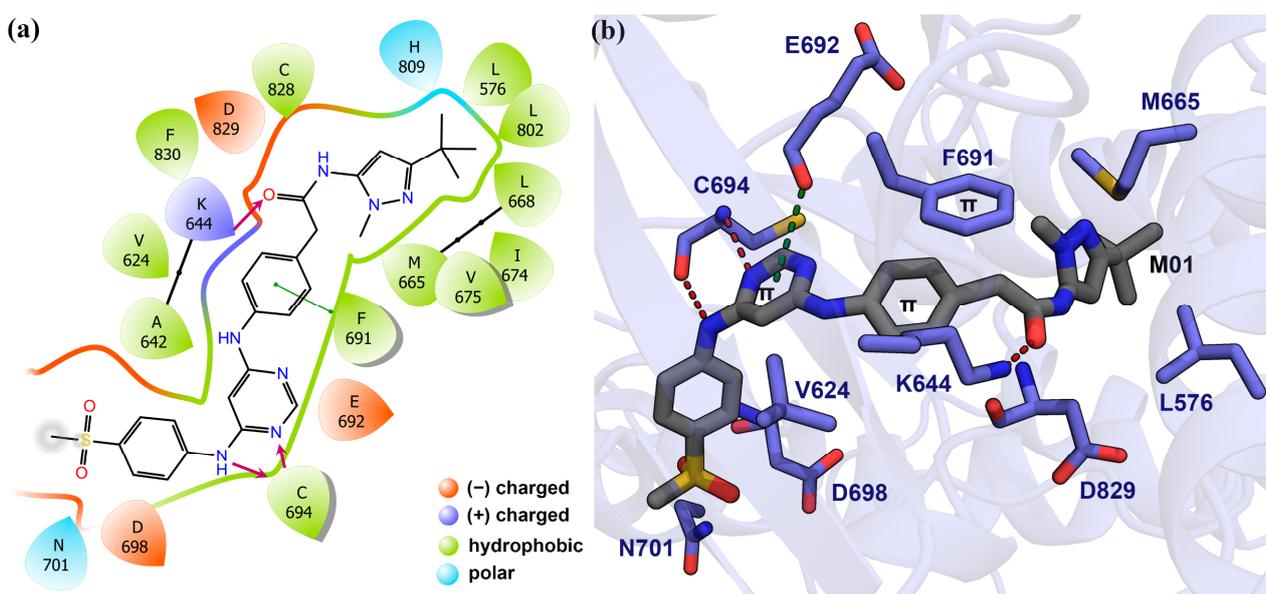

Figure 2. Molecular docking analysis of compound M01 and FLT3. (**a**) Compound M01 surrounded by the active site residues. The residues are illustrated based on their chemical properties in 2D representation. H-bonds are shown by magenta arrows. (**b**) Docking interaction of M01 with FLT3 residues. The H-bond interactions are shown by red dashed lines. The π–anion interaction is shown by green dashed lines.

2.2. MD Analysis

The MD simulations were conducted to evaluate the overall stability of the protein–ligand complexes. During the 100 ns MD run, the protein–ligand complexes achieved convergence within the first 10 ns. The RMSDs of proteins were found in the range of 1.0–4.0 Å, and the RMSDs of ligands were found in the range of 0.5–4 Å, as shown in Figure 3. We observed a low RMSD for compounds M01, M03, and M20, while compounds M17, M24, and M34 had a higher RMSD. In particular, the RMSD of the compounds M24 and M34 reached 4 Å during the simulation. Since the two H-bond interactions with residues K644 and C694 played a critical role in the binding of the ligand inside the ATP pocket, we compared these two interacting H-bond distances for 100 ns in Figure 4. For compound M34, the H-bond distance with K644 was found to be unstable and a major fluctuation was observed, in contrast to the other compounds. Overall, this specific H-bond distance was found to be within the range of 2–3.8 Å in Figure 4a. When comparing the H-bond distance between the ligands and N atom of residue C694, compounds M17 and M34 showed a slightly higher H-bond distance to ~3.7 Å in Figure 4b. However, this H-bond interaction was intact for the remaining compounds. Following that, we compared the average MD structure of protein–ligand complexes from the last 1 ns to their final docked position in

Figure 5. Figure 5a,b,e shows that *N*-phenylpyrimidine-4-amine substructures of M01, M03, and M24 closely retained their docking pose and molecular interactions until the end of the MD run. For M17, we observed that the dimethylamine with the pyrimidine ring was displaced from its initial docked position, which might increase the critical distance for H-bond formation as shown in Figure 5c. Compound M20 formed the H-bond interactions to the C694 with its N atoms of the pyridine ring, while another H-bond interaction remained intact with K644 (Figure 5d). The piperazine ring in M34, on the other hand, shifted forward from its initial docking position inside the ATP pocket (Figure 5f), which could be a reason for the extension of the distance from residue C694. We also observed the rotational and translational displacement of the piperidine ring and the pyrrolidine ring of compounds M03 and M20 by comparing their respective MD trajectories and docking positions.

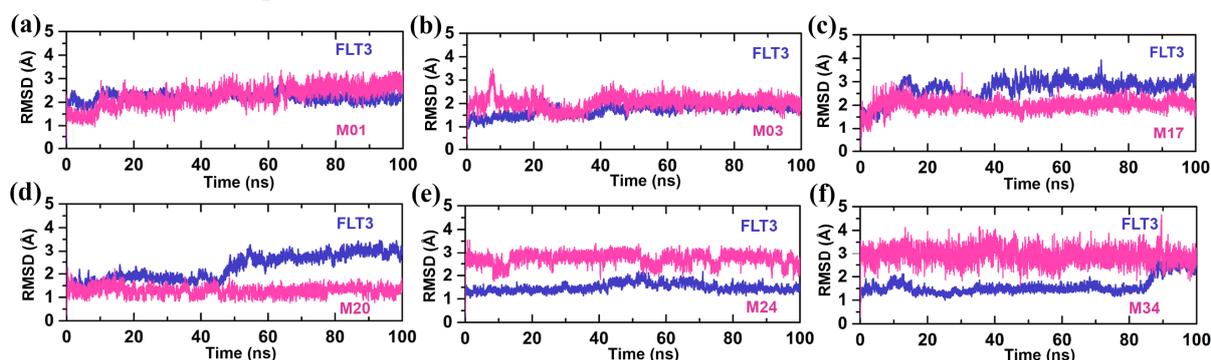**Figure 3.** (**a**) FLT3-M01, (**b**) FLT3-M03, (**c**) FLT3-M17, (**d**) FLT3-M20, (**e**) FLT3-M24, and (**f**) FLT3-M34 are the RMSD graphs of protein–ligand complexes. The protein and ligand RMSDs are shown in slate and pink colors, respectively.
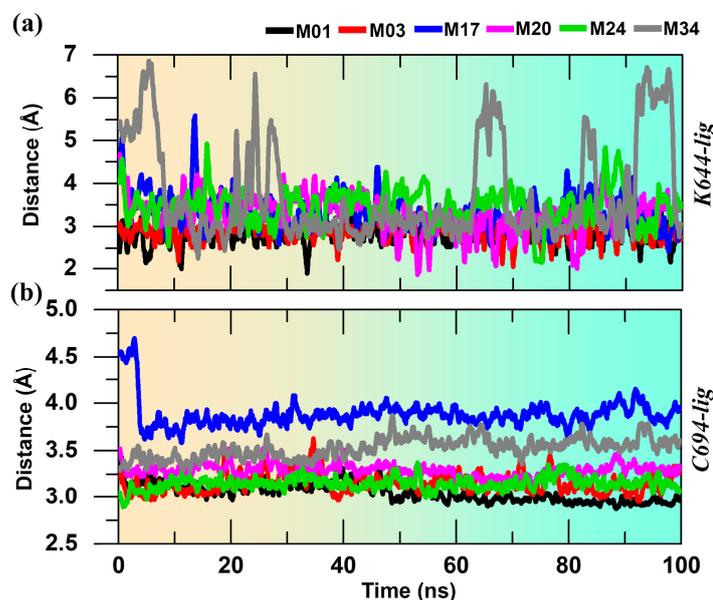

Figure 4. H-bond interaction distance analysis from (**a**) residue K644 and (**b**) C694 to the ligands during the production simulation.
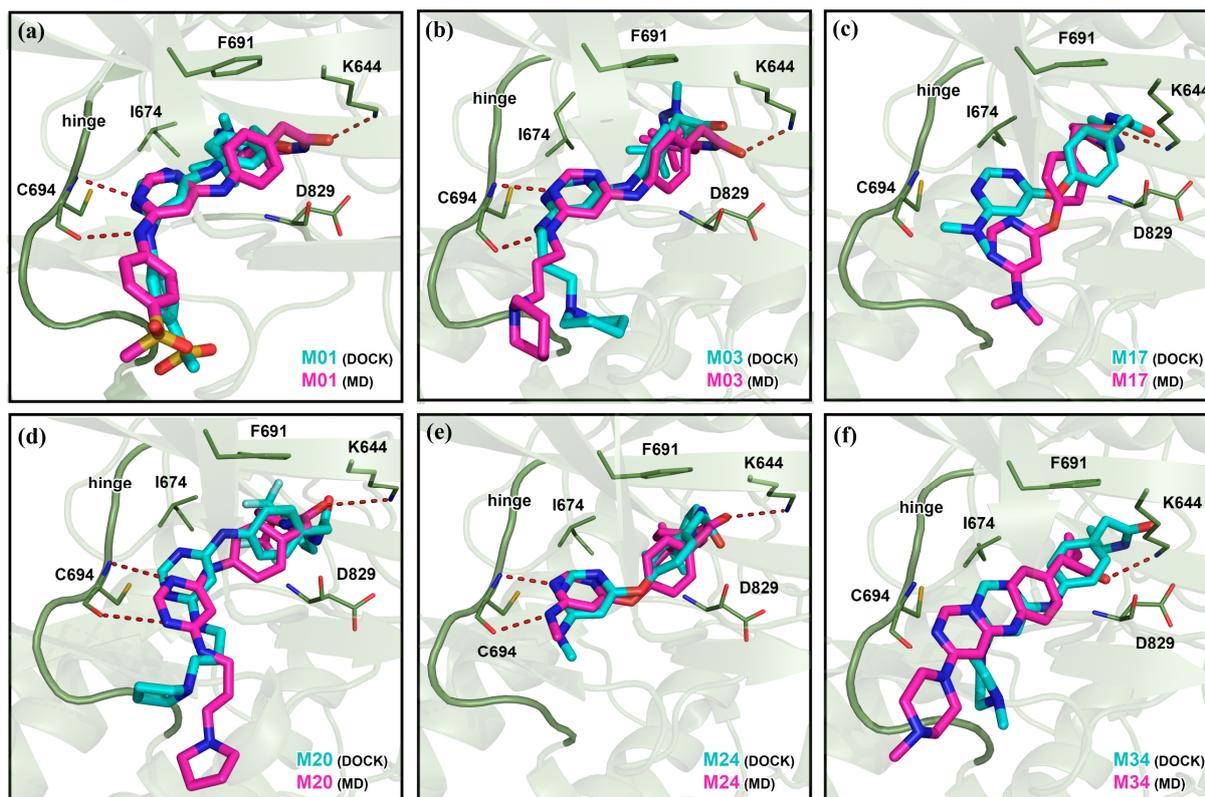
Figure 5. Average MD pose of the ligand at the active site of FLT3. The last 1 ns average MD poses of (**a**) M01, (**b**) M03, (**c**) M17, (**d**) M20, (**e**) M24, and (**f**) M34 inside the binding pocket are shown in pink color. The corresponding docked pose is shown in cyan color. The H-bond interactions are shown by red dashed lines.

2.3. MM-PB(GB)SA and LIE Estimation

We implemented the MM-PB(GB)SA and LIE estimation to investigate the binding affinity of FLT3 to selected ligands. To calculate the MM-PB(GB)SA and LIE binding energy, we used the final 2 ns or 200 frames of each protein–ligand MD trajectory. The final ΔTOTAL BEs were found to be −62.80, −60.27, −47.32, −60.68, −59.56, and −49.31 kcal/mol for compounds M01, M03, M17, M20, M24, and M34, respectively. The final ΔG_{bind} from the LIE estimation was found to be −18.39, −13.43, −12.56, −21.37, −9.75, and −8.14 kcal/mol for compounds M01, M03, M17, M20, M24, and M34, respectively. The detailed MM-PB(GB)SA energy and LIE terms are shown in Table 1. We estimated the per-residue BE decomposition by selecting residues within 4.0 Å from the ligand atoms. The common residues with positive or negative BE decomposition are compared in Table 2. We found that residues L616, V624, A642, K644, M665, I674, V675, F691, Y693, C694, G697, L818, C828, D829, and F830 were the key interacting common residues, contributing the critical BE to the ligands shown in Figure 6a. We acquired the substantially equilibrated 80th ns wild-type (WT) FLT3-M01 complex and manually mutated D835Y, F691L, and both (D835Y and F691L) at the same time. The newly prepared three different FLT3-M01 complexes were subjected to standard 20 ns MD simulations, followed by MM-PB(GB)SA and LIE calculations from the last 2 ns trajectory. The ΔTOTAL binding energy of MM-PB(GB)SA was found to be −63.55, −62.03, and −62.79 kcal/mol for the D835Y, F691L, and both mutants of FLT3 variants to compound M01, respectively. However, we were unable to observe any significant changes in ΔTOTAL BE terms with M01. During the calculation process, we restricted the analysis of per-residue MM-PB(GB)SA decomposition within the 4.0 Å distance from the ligand atoms. The per-residue BE decomposition between the F830 of FLT3 mutants and compound M01 was not generated, suggesting that the

distance between M01 and F830 increased to more than 4.0 Å, whereas the appearance of BE decompositions from residue K614 signify its proximity to compound M01.

Table 1. MM-PB(GB)SA and LIE energy estimation of the FLT3 inhibitor complexes.

Complexes	MM-PB(GB)SA Binding Energy Terms in kcal/mol							LIE (kcal/mol)
	VDWAALS	E_{EL}	$E_{PB/GB}$	E_{SURF}	ΔG_{gas}	ΔG_{solv}	ΔTOTAL	ΔG_{bind}
FLT3-M01	−71.17	−51.49	68.69	−8.82	−122.66	59.86	**−62.80**	**−18.39**
FLT3-M03	−72.23	−275.18	295.32	−8.72	−347.42	286.60	**−60.27**	**−13.43**
FLT3-M17	−53.77	−21.27	34.20	−6.48	−75.04	27.72	**−47.32**	**−12.56**
FLT3-M20	−69.48	−222.37	240.20	−9.02	−291.85	231.17	**−60.68**	**−19.35**
FLT3-M24	−62.38	−28.98	39.64	−7.83	−91.36	31.80	**−59.56**	**−9.75**
FLT3-M34	−56.90	−248.23	262.59	−6.76	−305.14	255.82	**−49.31**	**−8.14**
FLT3$_{(D835Y)}$-M01	−70.94	−51.33	67.31	−8.58	−122.28	58.72	**−63.55**	**−18.39**
FLT3$_{(F691L)}$-M01	−69.99	−53.10	69.92	−8.86	−123.09	61.05	**−62.03**	**−17.95**
FLT3$_{(D835Y,F691L)}$-M01	−70.43	−50.96	67.16	−8.56	−121.39	58.60	**−62.79**	**−17.95**

VDWAALS: van der Waals contribution from MM; E_{EL}: electrostatic energy as calculated by the MM force field; $E_{PB/GB}$: electrostatic contribution to the solvation free energy; E_{SURF}: non-polar solvation free energy; ΔG_{gas}: ΔG in gas phase; ΔG_{solv}: ΔG in solvation state; **ΔTOTAL**: total binding free energy from MM-PB(GB)SA; ΔG_{bind}: final LIE binding energy; **LIE**: linear interaction energy.

Table 2. Per-residue MM-PB(GB)SA ΔTOTAL binding free energy decomposition in kcal/mol.

Residues	Compounds								
	M01	M03	M17	M20	M24	M34	M01 (F691L)	M01 (D835Y)	M01 (F691L, D835Y)
K614	NA	NA	NA	NA	NA	NA	−1.92	−0.66	−0.93
L616	−2.53	−1.90	−2.20	−2.20	−1.13	−2.44	−2.35	−2.38	−2.63
V624	NA	−1.61	−0.60	−1.50	−1.54	NA	−1.68	−1.60	−1.48
A642	−1.25	−1.26	NA	−1.04	−1.00	−1.05	−1.24	−1.30	−1.38
K644	−2.15	−2.26	0.36	−1.43	−0.75	−0.38	−1.84	−1.24	−2.24
M665	−1.48	−1.63	−0.96	−0.99	−1.39	−0.46	−1.46	−1.40	−1.42
I674	−1.03	−1.26	NA	−0.98	NA	NA	−0.92	−1.29	−0.91
V675	−1.81	NA	NA	−0.98	−1.64	−0.93	−1.69	−1.67	−1.62
F691	−2.44	−2.80	−3.63	−2.54	−2.76	−2.15	−1.40 *	−2.63	−1.09 *
Y693	−1.99	−2.04	−0.68	−1.73	−2.04	−1.38	−1.77	−1.98	−1.98
C694	−2.82	−2.24	−0.13	−0.74	−2.41	−0.77	−2.77	−2.72	−2.81
G697	−1.39	−0.77	−1.47	−1.29	−0.39	−1.72	−1.33	−1.41	−1.40
L818	−1.76	−1.67	−1.89	−1.67	−1.56	−1.70	−1.78	−1.68	−1.64
C828	−2.81	−2.86	NA	NA	−4.17	−0.96	−2.79	−3.27	−2.42
D829	−1.22	−1.41	−2.21	−1.12	−0.94	−2.08	−1.19	−1.29	−1.73
F830	−0.78	−2.35	−1.63	−1.71	−1.17	−1.61	NA	NA	NA

NA: Distance of the residues that are more than 4 Å from the compounds or contributed negligible binding energy to the ligand; (*): BE decomposition from the mutated residue.
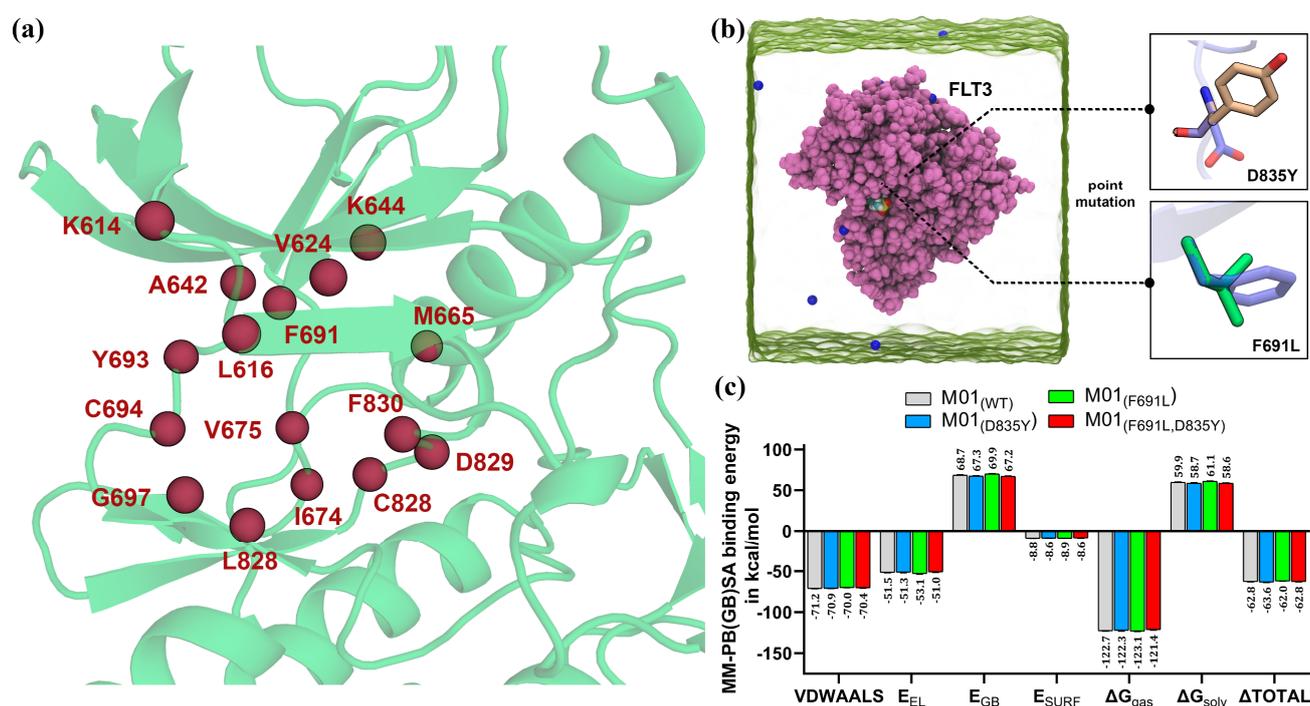
Figure 6. MM-PB(GB)SA binding free energy comparison between M01 and FLT3 variants. (**a**) Common residues that contributed to the ΔTOTAL BE to the ligands are shown in the sphere representation. (**b**) Point mutations were manually performed at D835 and F691 in Pymol to compare the MM-PB(GB)SA binding energy. (**c**) Comparison of the MM-PB(GB)SA binding energy terms between WT, D835Y, F691Y, and both (F691L, D835Y) mutant FLT3 and M01.

However, rather lower free energy values were observed in the VDWAALS and E_{EL} energy terms with the FLT3$_{(F691L)}$-M01 complex, as shown in Figure 6b,c. Minor energy differences could occur due to the substitution of the bulky hydrophobic residue F691 with a non-bulky leucine residue, as F691 is the key gatekeeper residue and had a strong π–π interaction with the phenyl ring of M01. The LIE values were found to be −18.39, −17.75, and −17.95 kcal/mol for F691L, and both mutant variants of FLT3 to compound M01, respectively. Residue D835Y mutation in the activation loop was 14.1 Å away from the active site. This mutation in the DFG-out FLT3 did not influence the final BE to the type II-like M01 in the estimation of MM-PB(GB)SA or LIE. Therefore, we assumed that the F691L mutation could partially affect selectivity and ligand interaction by changing the hydrophobic property of the nucleotide-binding pocket. This could also decrease the relative competitiveness of M01 against ATP and induce drug resistance [24,25]. In contrast, the D835Y mutation in the activation loop is strongly associated with the resistance mechanism of the active conformation (DFG-in) tyrosine kinase inhibitors (TKI). A pathway of the allosteric network has been proposed between the activation loop and the DFG motif [26], which plays an important role in stabilizing the adenosine triphosphate (ATP) molecule by chelating the Mg^{2+} ions at the catalytic site. The D835Y mutation could trigger an alteration of the allosteric mechanism in FLT3 by forming a dead mutant of kinases or disfavoring inhibitor binding by increasing entropy, ultimately influencing the entry of the ligand into the ATP-binding pocket [27].

2.4. Dataset Building, 3D-QSAR Model Development, and Model Validation

Coherence selection of the training set and test set compounds is a vital step toward the development of a 3D-QSAR statistical model. The chemical structures and their respective pIC$_{50}$ values, which were well spanned over three log units, are illustrated in Supplementary Table S2. Initially, we classified the compounds into high-, medium-, and low-activity groups. Next, we chose compounds at random from each group to form

the test set while maintaining structural diversity. We developed two sets of 3D-QSAR models, and their statistical validation is summarized in Table 3. The final 1 ns MD average structure of compound M01 was regarded as a 3D bioactive conformer and selected as a template molecule, as described here [28]. The rest of the compounds were modeled based on the template molecule and aligned with the common skeleton of *N*-methylpyrimidine-4-amine. The maximum common substructure (MCS) functionality was used to find the common core. The alignment of the dataset compounds inside the hydrophobic cleft and their unique chemical core are shown in Figure 7a,b. The first model was built by taking 40 compounds altogether. We strictly adhered to the statistical parameters listed in the threshold values column, as well as those described in the Methodology Section when validating each model.

Table 3. Statistics of the CoMFA and CoMSIA models.

Statistical Parameters	3D-QSAR (All Compounds)		3D-QSAR (Training Set Compounds)					Threshold Values
	CoMFA	CoMSIA (SEHA)	CoMFA	CoMSIA (SHD)	CoMSIA (SEHA)	CoMSIA (SEHD)	CoMSIA (SEHAD)	
q^2	0.735	0.725	0.802	0.730	0.726	0.725	0.721	>0.5
ONC	6	5	6	5	5	5	5	
SEP	0.502	0.503	0.452	0.517	0.521	0.522	0.525	
r^2	0.956	0.912	0.983	0.960	0.962	0.965	0.956	>0.6
SEE	0.204	0.284	0.134	0.199	0.194	0.186	0.209	<<1
F-value	119.97	70.84	216.62	114.54	121.34	131.90	104.16	>100
χ^2	0.052	0.078	0.012	0.028	0.027	0.023	0.028	<0.5
RMSE	0.219	0.265	0.119	0.181	0.176	0.169	0.189	<0.3
MAE	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	≈ 0
RSS	1.873	2.744	0.412	0.953	0.904	0.834	1.046	
k	NA	NA	1.033	1.022	1.032	1.033	1.037	$0.85 \leq k \leq 1.15$
k'	NA	NA	0.962	0.969	0.962	0.961	0.958	$0.85 \leq k' \leq 1.15$
$\lvert r_0^2 - r'_0^2 \rvert$	NA	NA	0.109	0.307	0.250	0.224	0.192	<0.3
$(r^2 - r_0^2)/r^2$	NA	NA	0.041	0.002	0.012	0.006	0.117	<0.1
r_m^2	NA	NA	0.724	0.603	0.635	0.649	0.664	>0.5
$\overline{r_m}^2$	NA	NA	0.694	0.449	0.492	0.511	0.623	>0.5
Δr_m^2	NA	NA	0.059	0.307	0.286	0.274	0.083	
Q_{F3}^2	NA	NA	0.698	0.604	0.656	0.668	0.660	>0.6
Q_{ccc}^2	NA	NA	0.821	0.746	0.778	0.787	0.787	
S (%)	52.9	20.2	54.4	28.9	18.5	20.5	16.2	
E (%)	47.1	29.9	45.6	NA	28.8	29.9	24.5	
H%	NA	37.0	NA	47.3	33.9	32.8	28.3	
A%	NA	13.0	NA	NA	18.8	NA	16.0	
D%	NA	NA	NA	23.8	NA	16.8	15.1	

q^2: squared cross-validated correlation coefficient; **ONC**: optimal number of components; **SEP**: standard error of prediction; r^2: squared correlation coefficient; **SEE:** standard error of estimation; **F-value**: F-test value; **RMSE**: root-mean-squared error; **MAE**: mean absolute error; **RSS**: residual sum of error; k: slope of the predicted vs. observed activity at zero intercept; k': slope of the observed vs. predicted activity at zero intercept; r_0^2: squared correlation coefficient between predicted and observed activity; r'_0^2: squared correlation coefficient between predicted and observed activity; Q_{F3}^2: Q^2 metrics for external test set validation; Q_{ccc}^2: concordance correlation coefficient; **S**: steric; **E**: electrostatic; **H**: hydrophobic; **A**: H-bond acceptor; **D**: H-bond donor; **NA**: Not applicable.
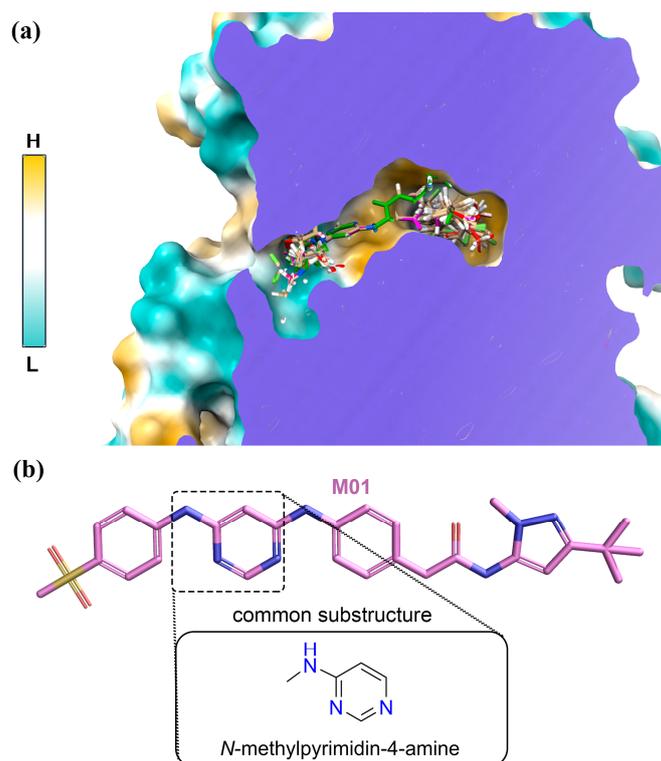
Figure 7. Common substructure-based molecular alignment. (**a**) Alignment of the compounds is shown inside the hydrophobic binding pocket of FLT3 by Z-plane clipping. Cyan to gold-yellow color bar indicate low (L) to high (H) hydrophobicity. (**b**) Common substructure of the compounds in the dataset.

In the first 3D-QSAR model, we obtained the best q^2 and r^2 values of 0.735 and 0.956, respectively, at an ONC of 6 in the CoMFA scheme. The SEHA produced the best statistical CoMSIA model with q^2 and r^2 values of 0.725 and 0.912, respectively at an ONC of 5 among the different combinations of descriptor fields (Supplementary Table S3). The q^2 and r^2 values of both CoMFA and CoMSIA appeared to be greater than the accepted threshold value ($q^2 > 0.5$ and $r^2 > 0.6$), which indicated a good agreement for the internal validation of both models. The χ^2 and RMSE values were found to be 0.052 and 0.219, respectively. In CoMSIA, the χ^2 and RMSE values were found to be 0.078 and 0.265, respectively. The χ^2 and RMSE values are within the acceptable threshold value, which signifies the fitness and accuracy of the models. However, any QSAR model would be uncertain without being validated externally. Therefore, the second model was developed using the training set of 30 compounds. The remaining 10 compounds were used as a test set to analyze the external predictivity. In the second 3D-QSAR model, the training set compounds generated the best q^2 and r^2 values of 0.802 and 0.960, respectively, at an ONC of 6 in the CoMFA scheme. The χ^2 ($\chi^2 = 0.012$) and RMSE (RMSE = 0.119) values were determined within the acceptable range of <0.3 and <0.5 [29], indicating for a good internal validation and model's fitness. For external validation, we calculated multiple statistical parameters described in these studies [30–32]. The k and k' were found to be 1.033 and 0.962, and the $\mid r_0^2 - r'_0^2 \mid$, $(r^2 - r_0^2)/r^2$, r_m^2, and Δr_m^2 were predicted to be 0.109, 0.041, 0.724, and 0.059, respectively. In addition, we acquired the Q_{F3}^2 and Q_{ccc}^2 metrics determination to evaluate the predictive ability of the CoMFA model. The Q_{F3}^2 value was found to be 0.698, which greater than the acceptable range (>0.6), and a high Q_{ccc}^2 value of 0.821 signifies the true predictive power of the model. To generate the best possible CoMSIA model, we applied five molecular descriptor fields—steric (S), electrostatic (E), hydrophobic (H), H-bond acceptor (H), and H-bond donor (D)—in different combinations (Supplementary Table S4). The actual and

predicted activities with their residuals are tabulated in Supplementary Table S5. The combination of SHD produced the best statistical q^2 values (q^2 = 0.730) at an ONC of 5. However, the best q^2 value does not necessarily indicate a good predictivity. Instead, the best CoMSIA model was selected based on the highest Q^2_{F3} metrics value. We found that the SEHD combination produced q^2 and r^2 values of 0.725 and 0.965, respectively, at an ONC of 5. The model had the Q^2_{F3} value of 0.665, the highest among all. Despite this, the model successfully passed all other statistical parameters. The *k* and *k'* were found to be 1.033 and 0.961, and the $|r_0^2 - r'^2_0|$, $(r^2 - r_0^2)/r^2$, r_m^2, and Δr_m^2 were predicted to be 0.109, 0.041, 0.724, and 0.059, respectively. Thereafter, the SEHD combination was selected as the final CoMSIA model. Overall, the second QSAR model, including CoMFA and CoMSIA (SEHD), produced the highest statistically significant values; therefore, we selected these models for contour map generation and SAR analysis. The PLS regression graph of CoMFA and CoMSIA is shown in Figure 8a,c, which described that our 3D-QSAR model could adequately predict the pIC$_{50}$ values of the test set compounds. The contribution of the steric and electrostatic field was found to be 54.4% and 45.6% in the CoMFA model, respectively, while the contributions of the steric, electrostatic, hydrophobic, and H-bond donor were found to be 20.5%, 29.9%, 32.8%, and 16.8%, respectively, in the CoMSIA model. Both models established good agreement between experimental activity and predicted activity by showing acceptable criteria ($q^2 > 0.6$, $r^2 > 0.8$) while having a satisfactory predictive performance ($Q^2_{F3} > 0.5$).
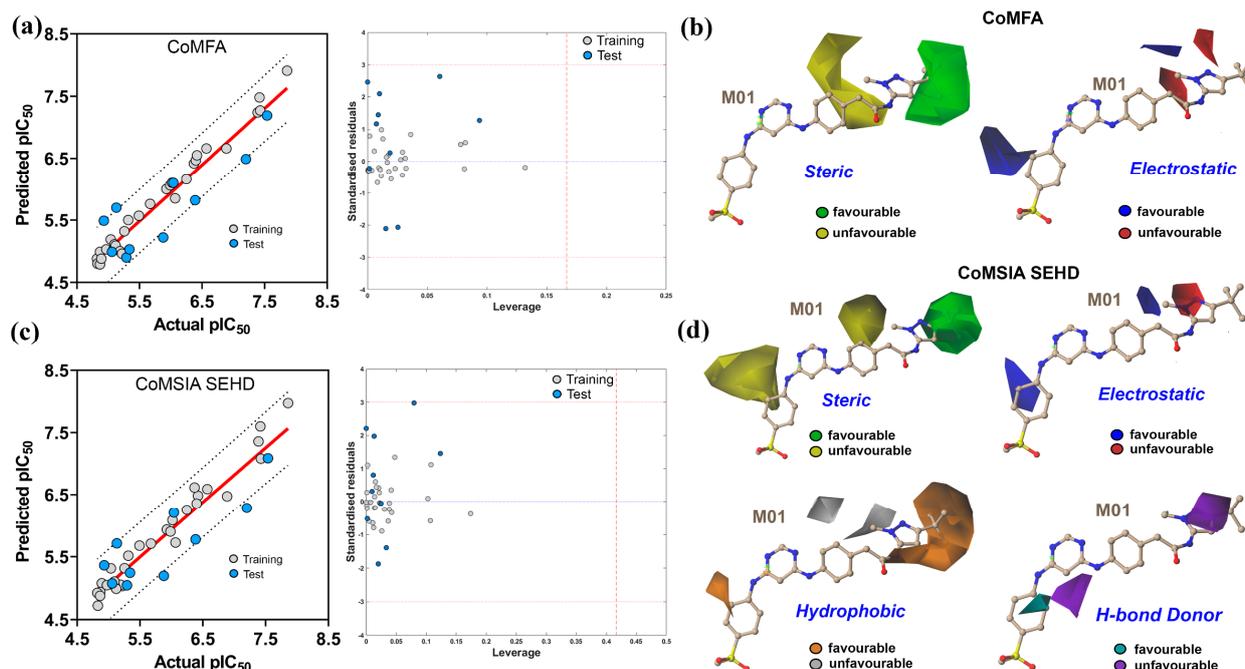

Figure 8. PLS scatter plot, Williams plot, and CoMFA-CoMSIA contour maps analysis. (**a**) Scatter plot of actual vs. predicted pIC$_{50}$ values from the CoMFA and the corresponding Williams plot for AD analysis. (**b**) The green and yellow contours from CoMFA indicate the favorable and unfavorable substitution for the steric groups. The blue and red contours indicate the favorable and unfavorable positions for the electropositive groups. (**c**) Scatter plot of actual vs. predicted pIC$_{50}$ values from the CoMSIA and the corresponding William's plot for AD analysis. (**d**) From CoMSIA, green and yellow contours signify the favorable and unfavorable positions for the steric group, whereas the blue and red contours show the favorable and unfavorable substitutions for electropositive groups. The orange and gray contours show the favorable and unfavorable positions for the hydrophobic groups. The cyan and purple colors signify the favorable and unfavorable positions for the H-bond donor groups.

The progressive scrambling method was applied to evaluate the additional stability of the CoMFA and CoMSIA models, as shown in Table 4. We ran 100 independent scrambles with minimum and maximum bin sizes of 2 and 10. At component number 6, scrambling

Q^2 and cSDEP were measured to be 0.502 and 0.713, respectively. The $dq^2/dr^2_{yy'}$ did not exceed the value 1.2 ($dq^2/dr^2_{yy'}$ = 1.198) in the CoMFA. In the CoMSIA, the scrambling Q^2, cSDEP, and $dq^2/dr^2_{yy'}$ were found to be 0.518, 0.709, and 0.982, respectively. Progressive scrambling helps to identify the optimal number of components of the model. In addition, it detects the model's sensitivity over a small perturbation when applied to the data.

Table 4. Progressive scrambling results from the CoMFA and CoMSIA(SEHD) training set compounds.

Components	CoMFA			CoMSIA (SEHD)		
	Q^2	cSDEP	$dq^2/dr^2_{yy'}$	Q^2	cSDEP	$dq^2/dr^2_{yy'}$
1	0.191	0.828	0.175	0.230	0.807	0.306
2	0.358	0.750	0.731	0.429	0.707	0.996
3	0.477	0.689	0.855	0.491	0.662	1.052
4	0.489	0.696	1.480	0.550	0.652	1.221
5	0.479	0.711	1.821	**0.518**	**0.770**	**0.982**
6	**0.502**	**0.713**	**1.198**	0.410	0.709	1.513
7	0.518	0.718	1.663	0.400	0.796	2.285

The QSAR models are generated by a limited number of compounds and can predict inhibitory activity for an unknown chemical that has a very similar chemical constitution. To assess the reliability of the CoMFA and CoMSIA models and specify the outliers, we used the applicability domain analysis by distance-based Williams plot as described in previous work [33]. The CoMFA and CoMSIA PLS plots, and the corresponding Williams plots, portrayed the standardized residuals of the training set and test set compounds against their leverage values. Compounds with a high leverage (h_i) value greater than the warning leverage (h^*) can be detected as outliers and have a substantial impact on the fitness of the model. We used the *Applicability Domain toolbox*, a MATLAB package developed by Sahigara et al. [34] to perform the AD analysis. In our study, all the compounds fell within the ±3 standardized residuals and the estimated warning leverages in CoMFA (h^* = 0.16) and CoMSIA (h^* = 0.42), confirming the overall predictive reliability of the models.

2.5. CoMFA and CoMSIA Contour Map Analysis

We generated 3D contour maps around the most active compound M01 in the SYBYL-X2.1 suit to elucidate the structure–activity relationships. The information collected from the contour map could be useful to improve the inhibitory potency of small molecules by altering the chemical groups. Figure 8b displayed the contour maps of the CoMFA analysis. The green and yellow contours around the pyrazole ring represent the favorable and unfavorable substitutions for the bulky and steric chemical groups. Specifically, the location of the green contour arrived near the residues M665 and V675, with which M01 formed the hydrophobic interaction in the docking study, and these residues contributed a BE of −1.48 and −1.81 kcal/mol, respectively. Compounds M01–M06, M11, and M38–M40 had steric groups in their R_2 position at the green contour, which could explain why its inhibitory activity is higher than that of compounds M31–M35.

In contrast, compounds M07, M08, and M10 had the steric phenyl ring facing the yellow contour at the R_2 position, which could decrease their inhibitory potency due to the probability of the steric hindrance effect. In Figure 8b, the blue and red contours show the favorable positions for the electropositive and electronegative groups. Compounds M01–M05, M38–M39, and M40 had the N atom in their pyrazole ring in the R_2 position and exhibited a higher inhibitory activity. Besides, a blue contour near the red contours at the R_2 position was found to show an unfavorable position for an electronegative group. Compounds M12–M16, M36, and M37 bearing the oxygen atom on their isoxazole ring and compounds M23, M27, M29, and M30 bearing the electronegative fluorine atom in the position R_2 exhibited a lower inhibitory activity than M01–M05. For the same reason, compound M40 had less inhibitory potency than compound M01. Another blue contour was present near the phenyl ring at the R_1 position, showing the favorable position of an elec-

tropositive chemical group. For compounds M15, M22, M24, and M35, an electropositive group was absent and showed lower inhibitory activity than compound M16.

Figure 8d illustrates the contour maps from the CoMSIA model analysis. We found that the steric and electrostatic descriptors yielded contour maps similar to those produced by CoMFA model. As a result, they are not discussed further. It also provided additional information on the hydrophobic and H-bond donor descriptors fields, making it preferred for SAR study. The favorable and unfavorable substitutions for hydrophobic chemical groups are shown by an orange-gray color scheme. A large orange contour at position R_2 around the pyrazole ring indicated that a hydrophobic chemical group would be beneficial in increasing inhibitory activity. In compounds M18–M21, M25, and M28, the methyl groups were replaced by fluorine atoms, and the presence of the methylsulfonyl group in M22 may affect the inhibitory activity of these compounds.

Similarly, the favorable and unfavorable positions for the H-bond donor groups are shown by cyan and purple contours. For this reason, compound M36 had an N atom in the R_2 position toward the purple contour and exhibited less inhibitory activity. From the analysis of the CoMFA and CoMSIA contour maps, we summarized an ideal SAR scheme in Figure 9. A positive charge group and an H-bond donor group on the methyl phenyl sulfone ring at position R_1 could enhance the potency of M01. At that position, residue C694 formed two H-bond interactions with the M01 and contributed -2.82 kcal/mol MM-PB(GB)SA ΔTOTAL binding free energy. As a result, a hydrophobic group next to this residue would most likely hinder the formation of an H-bond, lowering the binding free energy decomposition.
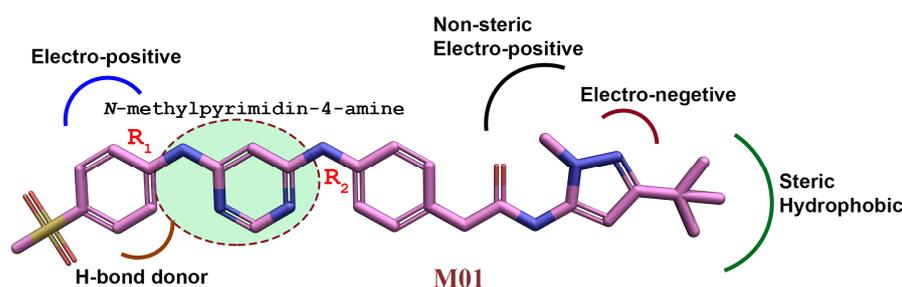

Figure 9. The three-dimensional structure–activity relationship by taking compound M01 as a reference.

Besides, the substitution of a smaller electron-donating group and the addition of a bulky hydrophobic group on the methylpyrazole ring in R_2 could improve the bioactivity of the compounds. However, the docking and MD study suggested that the steric substitution at the R_2 position was not infinite. A larger hydrophobic substitution could cause steric clashes with the hydrophobic residues in the αC-helix. Thereby, a smaller chemical group with steric and hydrophobic properties would be the preferred alteration strategy for an efficient interaction.

2.6. Designing of New Compounds

In the context of SAR analysis, we designed 30 new compounds using M01 as a template and evaluated their predictive activity in the CoMSIA-SHED model (Supplementary Table S6). Of these, 16 designed compounds were predicted to have higher inhibitory activity (pIC_{50}) than the most active compound M01. We re-docked these compounds against FLT3 and determined their physicochemical property and SA score (Supplementary Table S7). An SA score predicts the difficulty level in synthesizing the unknown chemical compounds. An SA score of 1 indicates the easy synthesis route, while an SA score of 10 signifies the difficult synthesis route. The SA score of the newly proposed compounds ranged from 3 to 5, suggesting that the compounds would be simple to moderately challenging to synthesize. To ensure that the new compounds would be viable lead compounds, we further assessed the ADMET properties (Supplementary Table S8). The absorption value

of more than 70% revealed that all compounds could be well absorbed in the human intestine. The LogBB > −2 and LogPS < −3 signify a better blood–brain barrier and CNS permeability. All the designed compounds were shown to be permeable through the BBB and CNS, while some of the compounds, such as D01, D02, D09–D12, and D17 may have a poor, yet acceptable CNS permeability. The cytochrome P450 enzymes are involved in the metabolism and biotransformation of the drug-like compounds. Except for D22, the other designed compounds are predicted to be substrates and inhibitors of CYP3A4. Excretion can be described as a constant ratio between drug concentration and renal clearance from the human body. The designed compounds yielded the acceptable values of total clearance, which were given in log mL/min/kg unit. Additionally, none of the designed compounds was predicted to be toxic in the AMES test.

The designed compounds with FLT3-bound complexes converged well within the 50 ns of the MD simulation (Supplementary Figure S3); therefore, they were not extended further. We calculated the MM-PB(GB)SA binding free energy of the last 2 ns of each trajectory (Table 5 and Figure 10) and per-residue binding free energy decomposition (Supplementary Table S9). The designed compounds such as D02–D08, D12, D15, and D22 showed a higher binding free energy compared to the most active compound. The last 1 ns average structure of each complex is illustrated in Supplementary Figure S4. The results strongly suggest that SAR-assisted newly designed compounds could have a better binding affinity for FLT3.

Table 5. MM-PB(GB)SA binding energy estimation between FLT3 and newly designed compounds.

Complexes	MM-PB(GB)SA Binding Energy Terms in kcal/mol						
	VDWAALS	E_{EL}	$E_{PB/GB}$	E_{SURF}	ΔG_{gas}	ΔG_{solv}	ΔTOTAL
FLT3-D01	−71.01	−45.14	66.81	−8.84	−116.16	57.96	**−58.19**
FLT3-D02	−70.82	−32.90	49.71	−9.17	−103.73	40.53	**−63.19**
FLT3-D03	−71.38	−55.60	71.34	−9.04	−126.98	62.30	**−64.68**
FLT3-D04	−72.37	−54.24	71.68	−9.30	−126.62	62.38	**−64.24**
FLT3-D05	−76.94	−63.43	79.81	−9.90	−140.38	69.90	**−70.47**
FLT3-D07	−73.59	−48.56	60.95	−9.45	−122.15	51.49	**−70.66**
FLT3-D08	−72.79	−33.30	50.01	−8.72	−106.10	41.29	**−64.81**
FLT3-D09	−74.03	−54.80	78.07	−9.34	−128.84	68.73	**−60.10**
FLT3-D10	−69.67	−49.27	71.01	−9.41	−118.94	61.59	**−57.35**
FLT3-D11	−80.40	−37.05	71.92	−10.24	−117.45	61.68	**−55.77**
FLT3-D12	−76.69	−27.84	46.29	−9.39	−104.53	36.90	**−67.63**
FLT3-D14	−80.42	−28.40	56.25	−10.07	−108.82	46.17	**−62.65**
FLT3-D15	−78.07	−57.04	70.96	−9.45	−135.11	61.50	**−73.61**
FLT3-D17	−66.50	−64.74	81.42	−8.66	−131.24	72.75	**−58.49**
FLT3-D21	−72.32	−17.78	46.93	−9.08	−90.11	37.84	**−52.27**
FLT3-D22	−73.23	−25.85	45.39	−9.44	−79.08	35.94	**−63.14**

VDWAALS: van der Waals contribution from MM; E_{EL}: electrostatic energy as calculated by the MM force field; $E_{PB/GB}$: electrostatic contribution to the solvation free energy; E_{SURF}: non-polar solvation free energy; ΔG_{gas}: ΔG in gas phase; ΔG_{solv}: ΔG in solvation state; ΔTOTAL: total binding free energy from MM-PB(GB)SA.
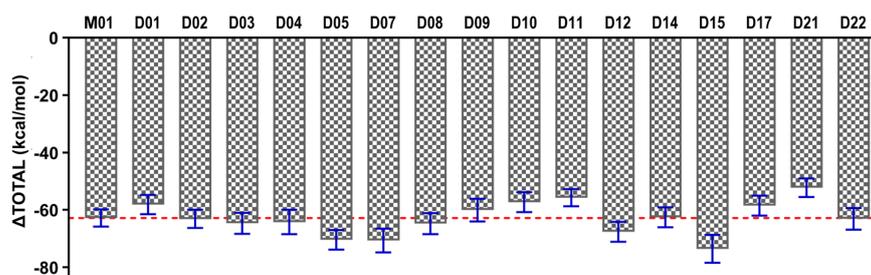

Figure 10. Comparison of the final ΔTOTAL binding energy terms of the newly designed compounds. Compounds D02–D08, D12, D15, and D22 were estimated to have higher binding free energy compare to M01 (red dashed line) and the standard deviation sign is shown in blue.

3. Materials and Methods

3.1. Protein Structure Preparation and Molecular Docking

The FLT3 protein structure with a resolution of 2.20 Å was retrieved from the Protein Databank (PDB: 6JQR) [8]. The water molecules and ligands from the crystallographic solution were removed from the protein crystal. The missing residues were remodeled using Modeler-10.1 (University of San Francisco, San Francisco, CA, USA) in UCSF Chimera-1.14 (RBVI, UCSF, San Francisco, CA, USA), and the final model was endorsed in a DFG-out configuration. The model with the lowest DOPE score was selected and the entire protein structure was verified using the PROCHECK (DOE-MBI service, UCLA, Los Angeles, CA, USA) server (Supplementary Figure S1).

Compound M01 was the most active compound in the dataset and was thus chosen for the docking study, as described in the previous study [35–37]. Compounds M03, M17, M20, M24, and M34, with different subgroups at their positions R_1 and R_2 exhibiting a diverse range of activity, were also included for the docking study. Briefly, the protein was prepared by assigning polar hydrogens and Kollman charges in AutoDockTools (AutoDock 4.2, Scripps Research, La Jolla, CA, USA). The ligand was sketched and minimized, and hydrogens were added in SYBYL-2.1, followed by the addition of Gasteiger charges. The number of rotatable bonds was fixed at six to the ligands to avoid conformational explosion. We assigned the active site according to the X-ray structure of Quizertinib-bound FLT3 (PDB:4XUF). The grid box size was set to 50 × 60 × 50 in the X, Y, and Z directions, respectively, with a grid spacing of 0.375 and a grid center of X = −34, Y = −10, and Z = −25 to compute the grid parameters using AutoGrid. For the conformational search, the Lamarckian Genetic Algorithm (LGA) was employed. Finally, AutoDock-4.2 was used to perform 100 docking runs. This protocol was repeated for the remaining compounds. AutoDockTools, Pymol (Schrodinger, Inc., New York, NY, USA), and Maestro (Schrodinger, Inc., New York, NY, USA) were used to analyze the docking results. All protein–ligand docked complexes were taken for the MD simulation study.

3.2. Molecular Dynamics

The MD simulation was carried out with GROMACS 2019.5 [38] using the Amber14SB [39] force field. The topology and parameters of the small molecules were generated using ACPYPE [40]. The system was prepared by placing the protein–ligand complex in a cubic periodic box. The complex was then wrapped with the TIP3P water model in such a way so that the minimum thickness of the water wall was maintained at 10 Å from the protein atoms. An adequate amount of NA^+ and Cl^- ions were added to neutralize the system.

The steepest descent algorithm was applied for 10,000 steps to minimize the system, followed by 200 ps of constant substance, volume, and temperature (NVT) simulation using the modified Berendsen thermostat (V-rescale) to attain a temperature of 300 K. Next, a 400 ps constant substance, pressure, and temperature (NPT) simulation was executed using the modified Berendsen barostat (V-rescale) to achieve 1 bar of pressure. The backbone and heavy atoms of the ligands were kept restrained during the minimization, NVT, and NPT simulation steps. Finally, the system was subjected to the 100 ns MD simulation run by un-restraining the backbone and heavy atoms. The Particle Mesh Ewald (PME) scheme was employed to maintain the electrostatic interactions, and the SHAKE algorithm was employed to deal with the bond length constraints. The cut-off distance was set at 12 Å to calculate the coulombic and van der Waals (vdW) interactions, respectively. The protocol was followed for the rest of the docked complexes as described in the previous study [41]. RMSD and H-bond distances were calculated by the in-build '*gmx rms*' and '*gmx distance*' functions in Gromacs.

3.3. MM-PB(GB)SA and LIE

The MM-PB(GB)SA is a useful technique for computing the end-state binding free energy between the protein and the ligand. In our study, the binding free energy of MM-

PB(GB)SA was calculated using the gmx_MMPBSA [42] package based on MMPBSA.py [43]. The protein–ligand binding free energy of MM-PB(GB)SA can be expressed by Equation (1),

$$\Delta G_{bind} = \Delta G_{complex} - \Delta G_{protein} - \Delta G_{ligand} \tag{1}$$

$$\Delta G_{bind} = \Delta E_{MM} + \Delta G_{sol} - T\Delta S \tag{2}$$

$$\Delta E_{MM} = (\Delta E_{vdW} + \Delta E_{ele}) \tag{3}$$

$$\Delta G_{sol} = \Delta G_{GB} + \Delta G_{SA} \tag{4}$$

where $\Delta G_{complex}$ stands for the total binding free energy between the protein–ligand complexes. The total free energy of the protein and ligand in the solvent was expressed by $\Delta G_{protein}$ and ΔG_{ligand}. In Equation (2), the ΔE_{MM} stands for the interaction energy between the protein–ligand complex under the gas-phase condition, which was estimated by calculating the van der Waals (ΔE_{vdW}) and electrostatic (ΔE_{ele}) energies (Equation (3)). The ΔG_{sol} stands for the free energy solvation, which was derived by calculating the polar solvation ΔG_{GB} and non-polar solvation ΔG_{SA} energy in Equation (4). The entropy contribution of the system is represented by $T\Delta S$. The entropy ($-T\Delta S$) calculation through nmode or Quasi-harmonic (QH) approximation is a computationally cumbersome process; therefore, the entropy term, ($-T\Delta S$) was not considered in this study.

The final 2 ns trajectories from each system were taken for the MM-PB(GB)SA total binding energy and per-residue binding energy decomposition analysis.

For the end-state LIE calculation, each of the docked ligands was simulated in an unbound state explicitly by keeping the same parameters as described in the MD simulation method. In the LIE method, the binding affinity (ΔG_{bind}) can be written as:

$$\Delta G_{bind} = \alpha \left(\langle V^{vdW}_{lig-surr} \rangle_{bound} - \langle V^{vdW}_{lig-surr} \rangle_{unbound} \right) + \beta \left(\langle V^{ele}_{lig-surr} \rangle_{bound} - \langle V^{ele}_{lig-surr} \rangle_{unbound} \right) \tag{5}$$

The Coulomb interaction ($V^{vdW}_{lig-surr}$) and electrostatic interaction ($V^{ele}_{lig-surr}$) values in unbound form were computed by the '*gmx enemat*' function. These values were supplied in the final LIE calculation using the '*gmx lie*' function. As described in these studies [44–46], we used the scaling factors $\alpha = 0.181$ and $\beta = 0.43$ by assigning the neutral charge to the ligands.

3.4. Dataset Building, Molecular Alignment, and CoMFA-CoMSIA (3D-QSAR) Study

The dataset of 40 pyrimidine-4,6-diamine-based compounds reported to be FLT3 inhibitors was taken for this study. We selected the last 1 ns MD average structure of compound M01 as a template structure. Based on the template structure, the rest of the compounds were sketched, and partial charges were assigned by the Gasteiger method in SYBYL-X2.1 (Tripos, Inc., St. Louis, MO, USA). All compounds were then subjected to minimization by tripos forcefield using the 0.05 kcal/mol convergence criterion with a maximum iteration of 2000 runs. Molecular alignment is an essential step towards the development of the 3D-QSAR model. We used the 'common substructure-based alignment' and 'database alignment' functions available in SYBYL-X2.1 to align compounds over the template structure as described in earlier studies [47,48].

For the 3D-QSAR study, the IC_{50} values of the compounds were converted into logarithmic IC_{50} (pIC_{50}) values. The dataset was divided into the training set of 30 compounds to develop the model and a test set of 10 compounds to evaluate the external predictive capability. CoMFA and CoMSIA are the two widely established 3D-QSAR approaches, which were used to establish the structure–activity relationship of the compounds in the dataset. In CoMFA, the steric fields (S) were calculated using the Lennard–Jones potential function, whereas the electrostatic fields (E) were calculated using the Coulombic potential function. The compounds were placed one after another in a spatial grid box with a grid spacing of 2.0 by maintaining an energy tolerance of 30 kcal/mol. The sp^3 carbon atom was

assigned as a probe by setting the van der Waals radii at 1.52 Å, with a net charge of +1.0. The other parameters were accepted by default in SYBYL-X2.1.

In the CoMSIA model, besides the steric and electrostatic fields, three additional descriptors, such as hydrophobic (H), H-bond acceptor (A), and H-bond donor (D) fields, were also adopted. The Gaussian-type functions were used to distinguish the distance between the probe atoms and the molecule's atoms for all grid points. The rest of the parameters were kept similar to the CoMFA parameters. All descriptors were used in different combinations to get the best possible CoMSIA model.

The partial least squares (PLS) method was adopted to analyze the internal validation of CoMFA and CoMSIA models. The leave-one-out (LOO) method was applied to obtain the cross-validation coefficient (q^2) and the optimal number of components (ONC) with a column-filtering value of 2.0 kcal/mol. Subsequently, the non-cross-validated correlation coefficient (r^2), Fisher's statistics (F value), and the standard error of estimation (SEE) were calculated, and finally, the pIC$_{50}$ values were predicted for each compound of the test set. The value of q^2 and r^2 greater than 0.5 and 0.6, respectively, as well as $r^2 - q^2$ not exceeding 0.3 well indicated the internal validity of the QSAR models.

The fitness of the CoMFA and CoMSIA models were assessed by estimating the chi-squared (χ^2) and root-mean-squared error (RMSE) [29]:

$$x^2 = \sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{y_i^2} \tag{6}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-1}} \tag{7}$$

where y_i and $\hat{y}_i$ are the observed and predicted activity, respectively, and n is the number of compounds being measured. The large χ^2 (≥ 0.5) and RMSE (≥ 1.0) values reflect the poor predictive accuracy of the QSAR models.

The external validation of the CoMFA and CoMSIA by test set compounds was a crucial step in determining the true predictive power of any QSAR models. The following criteria proposed by Roy et al. [30], Gramatica et al. [31], and Todeschini et al. [32] were used to externally validate the models. The in-depth analysis process is described in the Supplementary Material under the Methodology Section. The progressive scrambling study was conducted as described in this study [49].

The applicability domain (AD) analysis of the QSAR models was done using the leverage approach as described in earlier studies [34]. The standardized residuals from the activity values of the training set and test set compounds were plotted against their leverage values in the Williams plot. The detailed method was described in the Supplementary Materials. The leverage value of any compound exceeding the warning leverage (h^*), as shown with the red dotted line, was denoted as outliers and influenced the model quality.

3.5. Contour Maps Analysis

The contour maps were generated from CoMFA and CoMSIA to explain the structure–activity relationship of the existing compounds. The most active compound M01 was placed in the center as a reference, and the contour maps were shown as 3D StDev*Coeff to elucidate the field effects of the descriptors. The green and yellow contours signify the favorable and unfavorable substitutions for steric groups [50,51]. The blue and red contours convey the favorable and unfavorable substitutions for an electropositive group in CoMFA and CoMSIA. Similarly, the favorable and unfavorable substitutions for the hydrophobic, H-bond acceptor, and H-bond donor are represented by orange-gray, magenta-slate, and cyan-purple color schemes.

3.6. Designing of the New Compounds

Based on the SAR study, we designed 30 new compounds and predicted their activity by the CoMSIA (SEHD) model. Sixteen compounds out of them were predicted to

be a higher pIC50 value than the most active compound M01. Thereafter, we analyzed physicochemical properties, the synthetic accessibility (SA) score by the SWISSADME [52] webserver, and the absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties by the pkCSM [53] webserver of these compounds. Finally, the designed compounds were subjected to MM-PB(GB)SA binding free energy calculation.

4. Conclusions

The development of FLT3 inhibitors is a promising strategy for achieving the therapeutic goal of the treatment of AML. In this work, we used molecular docking and MD simulations to better understand the crucial interaction and stability of the ligands inside the binding pocket. The protein–ligand affinity was estimated by computing the MM-PB(GB)SA and LIE-binding energy, including the mutant FLT3 complexes. In a previously conducted docking study, residues Q575, L576, Q577, K644, F691, C694, L818, and L832 alongside the DFG motif (D829, F830, and G831) formed the fundamental docking interface for a group of diverse inhibitors. In our study, the combination of docking and MD study exposed several key residues, such as L616, M665, F691, E692, N701, L818, and F830 inside the ATP-binding pocket, which could be responsible for the affinity and selectivity of the type II inhibitors. K644 and C694 are the two important residues that form H-bond interactions. Subsequently, we established the CoMFA ($q^2 = 0.802$, $r^2 = 0.983$) and CoMSIA ($q^2 = 0.725$, $r^2 = 0.965$) models, both of which show a reasonable statistical correlation between actual and predictive activity and internal verification ability. The models also showed satisfactory stability and sensitivity in progressive scrambling analysis. The Q_{F3}^2 metrics from CoMFA and CoMSIA were found to be 0.698 and 0.668, respectively, indicating the external predictive power. The models could predict the activity values of new compounds having a similar scaffold. Following that, we developed the CoMFA and CoMSIA contour maps for the SAR study. Contour maps could provide crucial information regarding how chemical group substitution may improve the inhibitory activity of chemical compounds, which may be explained further by their docking and MD poses. Finally, we designed 30 new compounds, 16 of which had a higher predictive pIC$_{50}$ than compound M01. In addition, free energy calculation of the selected designed compounds revealed a greater binding affinity to the FLT3. Compounds D02–D08, D12, D15, and D22, in particular, might offer potential inhibitory activity against FLT3.

In this study, several effective computational approaches and reliable statistics were used, and they may provide several key mechanistic interpretations at the molecular level. The overall findings of this work may be beneficial in delivering theoretical guidance in the future development and synthesis of *N*-phenylpyrimidine-4-amine-based FLT3 inhibitors.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/ijms222212511/s1.

Author Contributions: S.G., S.K. and S.J.C. conceptualized and designed the project. S.G. and S.K. performed the experiments, data analysis and figure making and wrote the manuscript. S.J.C. superintended the project. All authors discussed the work and remarked on this manuscript. All authors have read and agreed to the published version of the manuscript.

Informed Consent Statement: Not applicable.

Data Availability Statement: The authors declared the availability of all materials and data upon request.

Conflicts of Interest: The authors have no financial competing interest to declare.

Abbreviations

FLT3	FMS-like tyrosine kinase-3
AML	acute myeloid leukemia
MM-PB(GB)SA	Molecular Mechanics Poison–Boltzmann/Generalized Born Surface Area
LIE	linear interaction energy
BE	binding Energy
CoMFA	comparative molecular field analysis
CoMSIA	comparative molecular similarity indices analysis
3D-QSAR	three-dimensional structure–activity relationship

References

1. Wang, Z.; Cai, J.; Cheng, J.; Yang, W.; Zhu, Y.; Li, H.; Lu, T.; Chen, Y.; Lu, S. FLT3 Inhibitors in Acute Myeloid Leukemia: Challenges and Recent Developments in Overcoming Resistance. *J. Med. Chem.* **2021**, *64*, 2878–2900. [CrossRef]
2. Cueto, F.; Sancho, D. The Flt3L/Flt3 Axis in Dendritic Cell Biology and Cancer Immunotherapy. *Cancers* **2021**, *13*, 1525. [CrossRef] [PubMed]
3. Pannecoucke, E.; Raes, L.; Savvides, S.N. Engineering and crystal structure of a monomeric FLT3 ligand variant. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **2021**, *77*, 121–127. [CrossRef] [PubMed]
4. Yen, S.-C.; Chen, L.-C.; Huang, H.-L.; Ngo, S.-T.; Wu, Y.-W.; Lin, T.E.; Sung, T.-Y.; Lien, S.-T.; Tseng, H.-J.; Pan, S.-L.; et al. Investigation of Selected Flavonoid Derivatives as Potent FLT3 Inhibitors for the Potential Treatment of Acute Myeloid Leukemia. *J. Nat. Prod.* **2021**, *84*, 1–10. [CrossRef]
5. Wilson, K.R.; Villadangos, J.A.; Mintern, J.D. Dendritic cell Flt3—regulation, roles and repercussions for immunotherapy. *Immunol. Cell Biol.* **2021**, *99*, 962–971. [CrossRef]
6. Almatani, M.F.; Ali, A.; Onyemaechi, S.; Zhao, Y.; Gutierrez, L.; Vaikari, V.P.; Alachkar, H. Strategies targeting FLT3 beyond the kinase inhibitors. *Pharmacol. Ther.* **2021**, *225*, 107844. [CrossRef] [PubMed]
7. Hasegawa, H.; Taniguchi, H.; Nakamura, Y.; Kato, T.; Fujii, S.; Ebi, H.; Shiozawa, M.; Yuki, S.; Masuishi, T.; Kato, K.; et al. FMS-like tyrosine kinase 3 (FLT3) amplification in patients with metastatic colorectal cancer. *Cancer Sci.* **2021**, *112*, 314. [CrossRef]
8. Kawase, T.; Nakazawa, T.; Eguchi, T.; Tsuzuki, H.; Ueno, Y.; Amano, Y.; Mori, M.; Yoshida, T. Effect of Fms-like tyrosine kinase 3 (FLT3) ligand (FL) on antitumor activity of gilteritinib, a FLT3 inhibitor, in mice xenografted with FL-overexpressing cells. *Oncotarget* **2019**, *10*, 6111. [CrossRef] [PubMed]
9. Roskoski, R., Jr. The role of small molecule Flt3 receptor protein-tyrosine kinase inhibitors in the treatment of Flt3-positive acute myelogenous leukemias. *Pharmacol. Res.* **2020**, *155*, 104725. [CrossRef]
10. Bjelosevic, S.; Gruber, E.; Newbold, A.; Shembrey, C.; Devlin, J.R.; Hogg, S.J.; Kats, L.; Todorovski, I.; Fan, Z.; Abrehart, T.C.; et al. Serine Biosynthesis Is a Metabolic Vulnerability in FLT3-ITD–Driven Acute Myeloid Leukemia. *Cancer Discov.* **2021**, *11*, 1582–1599. [CrossRef]
11. Qiao, X.; Ma, J.; Knight, T.; Su, Y.; Edwards, H.; Polin, L.; Li, J.; Kushner, J.; Dzinic, S.H.; White, K.; et al. The combination of CUDC-907 and gilteritinib shows promising in vitro and in vivo antileukemic activity against FLT3-ITD AML. *Blood Cancer J.* **2021**, *11*, 1–17. [CrossRef]
12. Scholl, S.; Fleischmann, M.; Schnetzke, U.; Heidel, F.H. Molecular Mechanisms of Resistance to FLT3 Inhibitors in Acute Myeloid Leukemia: Ongoing Challenges and Future Treatments. *Cells* **2020**, *9*, 2493. [CrossRef]
13. Mosquera Orgueira, A.; Bao Pérez, L.; Mosquera Torre, A.; Peleteiro Raíndo, A.; Cid López, M.; Díaz Arias, J.Á.; Ferreiro Ferro, R.; Antelo Rodríguez, B.; González Pérez, M.S.; Albors Ferreiro, M.; et al. FLT3 inhibitors in the treatment of Acute Myeloid Leukemia: Current status and future perspectives. *Minerva Med.* **2020**, *111*, 427–442. [CrossRef]
14. Bertoli, S.; Dumas, P.-Y.; Bérard, E.; Largeaud, L.; Bidet, A.; Delabesse, E.; Tavitian, S.; Gadaud, N.; Leguay, T.; Leroy, H.; et al. Outcome of Relapsed or Refractory FLT3-Mutated Acute Myeloid Leukemia before Second-Generation FLT3 Tyrosine Kinase Inhibitors: A Toulouse–Bordeaux DATAML Registry Study. *Cancers* **2020**, *12*, 773. [CrossRef] [PubMed]
15. Bhujbal, S.P.; Keretsu, S.; Cho, S.J. Design of New Therapeutic Agents Targeting FLT3 Receptor Tyrosine Kinase Using Molecular Docking and 3D-QSAR Approach. *Lett. Drug Des. Discov.* **2020**, *17*, 585–596. [CrossRef]
16. Bensinger, D.; Stubba, D.; Cremer, A.; Kohl, V.; Waßmer, T.; Stuckert, J.; Engemann, V.; Stegmaier, K.; Schmitz, K.; Schmidt, B. Virtual Screening Identifies Irreversible FMS-like Tyrosine Kinase 3 Inhibitors with Activity toward Resistance-Conferring Mutations. *J. Med. Chem.* **2019**, *62*, 2428–2446. [CrossRef]
17. Smith, C.C.; Lin, K.; Stecula, A.; Sali, A.; Shah, N.P. FLT3 D835 mutations confer differential resistance to type II FLT3 inhibitors. *Leukemia* **2015**, *29*, 2390–2392. [CrossRef]
18. Bharate, J.B.; McConnell, N.; Naresh, G.; Zhang, L.; Lakkaniga, N.R.; Ding, L.; Shah, N.P.; Frett, B.; Li, H.-Y. Rational Design, Synthesis and Biological Evaluation of Pyrimidine-4,6-diamine derivatives as Type-II inhibitors of FLT3 Selective Against c-KIT. *Sci. Rep.* **2018**, *8*, 3722. [CrossRef] [PubMed]
19. Yamaura, T.; Nakatani, T.; Uda, K.; Ogura, H.; Shin, W.; Kurokawa, N.; Saito, K.; Fujikawa, N.; Date, T.; Takasaki, M.; et al. A novel irreversible FLT3 inhibitor, FF-10101, shows excellent efficacy against AML cells with FLT3 mutations. *Blood* **2018**, *131*, 426–438. [CrossRef]

20. Wu, T.-S.; Lin, W.-H.; Tsai, H.-J.; Hsueh, C.-C.; Hsu, T.; Wang, P.-C.; Lin, H.-Y.; Peng, Y.-H.; Lu, C.-T.; Lee, L.-C.; et al. Discovery of Conformational Control Inhibitors Switching off the Activated c-KIT and Targeting a Broad Range of Clinically Relevant c-KIT Mutants. *J. Med. Chem.* **2019**, *62*, 3940–3957. [CrossRef] [PubMed]

21. Zorn, J.A.; Wang, Q.; Fujimura, E.; Barros, T.; Kuriyan, J. Crystal Structure of the FLT3 Kinase Domain Bound to the Inhibitor Quizartinib (AC220). *PLoS ONE* **2015**, *10*, e0121177. [CrossRef]

22. Gutiérrez, C.M.; Cáceres-Rojas, D.; Adasme-Carreño, F.; Palomo, I.; Fuentes, E.; Caballero, J. Docking and quantitative structure–activity relationship of bi-cyclic heteroaromatic pyridazinone and pyrazolone derivatives as phosphodiesterase 3A (PDE3A) inhibitors. *PLoS ONE* **2017**, *12*, e0189213. [CrossRef]

23. Caballero, J.; Morales-Bayuelo, A.; Navarro-Retamal, C. Mycobacterium tuberculosis serine/threonine protein kinases: Structural information for the design of their specific ATP-competitive inhibitors. *J. Comput. Mol. Des.* **2018**, *32*, 1315–1336. [CrossRef]

24. Nakaoku, T.; Kohno, T.; Araki, M.; Niho, S.; Chauhan, R.; Knowles, P.P.; Tsuchihara, K.; Matsumoto, S.; Shimada, Y.; Mimaki, S.; et al. A secondary RET mutation in the activation loop conferring resistance to vandetanib. *Nat. Commun.* **2018**, *9*, 1–9. [CrossRef]

25. Mondal, J.; Tiwary, P.; Berne, B.J. How a Kinase Inhibitor Withstands Gatekeeper Residue Mutations. *J. Am. Chem. Soc.* **2016**, *138*, 4608–4615. [CrossRef]

26. Ahuja, L.G.; Aoto, P.; Kornev, A.P.; Veglia, G.; Taylor, S.S. Dynamic allostery-based molecular workings of kinase:peptide complexes. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 15052–15061. [CrossRef]

27. Barouch-Bentov, R.; Sauer, K. Mechanisms of drug resistance in kinases. *Expert Opin. Investig. Drugs* **2011**, *20*, 153–208. [CrossRef]

28. Zhang, L.; Tsai, K.-C.; Du, L.; Fang, H.; Li, M.; Xu, W. How to generate reliable and predictive CoMFA models. *Curr. Med. Chem.* **2011**, *18*, 923–930. [CrossRef]

29. Veerasamy, R.; Rajak, H.; Jain, A.; Sivadasan, S.; Varghese, C.P.; Agrawal, R.K. Validation of QSAR models-strategies and importance. *Int. J. Drug Des. Discov.* **2011**, *3*, 511–519.

30. Roy, K.; Chakraborty, P.; Mitra, I.; Ojha, P.K.; Kar, S.; Das, R.N. Some case studies on application of "rm2" metrics for judging quality of quantitative structure-activity relationship predictions: Emphasis on scaling of response data. *J. Comput. Chem.* **2013**, *34*, 1071–1082. [CrossRef]

31. Gramatica, P.; Sangion, A. A Historical Excursus on the Statistical Validation Parameters for QSAR Models: A Clarification Concerning Metrics and Terminology. *J. Chem. Inf. Model.* **2016**, *56*, 1127–1131. [CrossRef]

32. Todeschini, R.; Ballabio, D.; Grisoni, F. Beware of Unreliable Q2! A Comparative Study of Regression Metrics for Predictivity Assessment of QSAR Models. *J. Chem. Inf. Model.* **2016**, *56*, 1905–1913. [CrossRef] [PubMed]

33. Abdizadeh, R.; Hadizadeh, F.; Abdizadeh, T. QSAR analysis of coumarin-based benzamides as histone deacetylase inhibitors using CoMFA, CoMSIA and HQSAR methods. *J. Mol. Struct.* **2020**, *1199*, 126961. [CrossRef]

34. Sahigara, F.; Ballabio, D.; Todeschini, R.; Consonni, V. Assessing the validity of QSARs for ready biodegradability of chemicals: An applicability domain perspective. *Curr. Comput. Drug Des.* **2014**, *10*, 137–147. [CrossRef]

35. Ghosh, S.; Keretsu, S.; Cho, S.J. 3D-QSAR, Docking and Molecular Dynamics Simulation Study of C-Glycosylflavones as GSK-3β Inhibitors. *J. Chosun Natural Sci.* **2020**, *13*, 170–180.

36. Ghosh, S.; Keretsu, S.; Cho, S.J. Computational Modeling of Novel Phosphoinositol-3-kinase γ Inhibitors Using Molecular Docking, Molecular Dynamics, and 3D-QSAR. *Bull. Korean Chem. Soc.* **2021**, *42*, 1093–1111. [CrossRef]

37. Ghosh, S.; Keretsu, S.; Cho, S.J. Designing of the *N*-ethyl-4-(pyridin-4-yl)benzamide based potent ROCK1 inhibitors using docking, molecular dynamics, and 3D-QSAR. *PeerJ* **2021**, *9*, e11951. [CrossRef]

38. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.E.; Berendsen, H.J.C. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701–1718. [CrossRef]

39. Maier, J.A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K.E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713. [CrossRef]

40. Da Silva, A.W.S.; Vranken, W.F. ACPYPE-Antechamber python parser interface. *BMC Res. Notes* **2012**, *5*, 1–8. [CrossRef]

41. Keretsu, S.; Ghosh, S.; Cho, S. Molecular Modeling Study of c-KIT/PDGFRα Dual Inhibitors for the Treatment of Gastrointestinal Stromal Tumors. *Int. J. Mol. Sci.* **2020**, *21*, 8232. [CrossRef] [PubMed]

42. Tresanco, M.S.V.; Valdes-Tresanco, M.E.; Valiente, P.A.; Frías, E.M. *gmx_MMPBSA (Version v1.4.3)*; Zenodo: 2021. [CrossRef]

43. Miller, B.R., III; McGee, T.D., Jr.; Swails, J.M.; Homeyer, N.; Gohlke, H.; Roitberg, A.E. MMPBSA. py: An efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321. [CrossRef] [PubMed]

44. Rifai, E.A.; van Dijk, M.; Geerke, D.P. Recent Developments in Linear Interaction Energy Based Binding Free Energy Calculations. *Front. Mol. Biosci.* **2020**, *7*, 7. [CrossRef]

45. Ngo, S.T.; Hung, H.M.; Nguyen, M.T. Fast and accurate determination of the relative binding affinities of small compounds to HIV-1 protease using non-equilibrium work. *J. Comput. Chem.* **2016**, *37*, 2734–2742. [CrossRef]

46. Rifai, E.A.; Van Dijk, M.; Vermeulen, N.P.E.; Yanuar, A.; Geerke, D.P. A Comparative Linear Interaction Energy and MM/PBSA Study on SIRT1–Ligand Binding Free Energy Calculation. *J. Chem. Inf. Model.* **2019**, *59*, 4018–4033. [CrossRef]

47. Bang, S.; Cho, S. Comparative Molecular Field Analysis (CoMFA) and Comparative Molecular Similarity Index Analysis (CoMSIA) Study of Mutagen X. *Bull. Korean Chem. Soc.* **2004**, *25*, 1525–1530. [CrossRef]

48. Gadhe, C.G.; Madhavan, T.; Kothandan, G.; Cho, S.J. In Silico Quantitative Structure-Activity Relationship Studies on P-gp Modulators of Tetrahydroisoquinoline-Ethyl-Phenylamine Series. *BMC Struct. Biol.* **2011**, *11*, 5. [CrossRef]

49. Zięba, A.; Laitinen, T.; Patel, J.; Poso, A.; Kaczor, A. Docking-Based 3D-QSAR Studies for 1,3,4-oxadiazol-2-one Derivatives as FAAH Inhibitors. *Int. J. Mol. Sci.* **2021**, *22*, 6108. [CrossRef]

50. Gadhe, C.G.; Kothandan, G.; Cho, S.J. Large variation in electrostatic contours upon addition of steric parameters and the effect of charge calculation schemes in CoMFA on mutagenicity of MX analogues. *Mol. Simul.* **2012**, *38*, 861–871. [CrossRef]

51. San Juan, A.A.; Cho, S.J. 3D-QSAR study of microsomal prostaglandin E_2 synthase (mPGES-1) inhibitors. *J. Mol. Model.* **2007**, *13*, 601–610. [CrossRef] [PubMed]

52. 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*, 1–13. [CrossRef]

53. Pires, D.E.V.; Blundell, T.L.; Ascher, D.B. pkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures. *J. Med. Chem.* **2015**, *58*, 4066–4072. [CrossRef]