

Part S1. The detailed architectures of several models.

Emb

$$G_p(\vec{n}_i, \vec{n}_i) = G_p(n_i \cdot W_{n_1}, W_{n_2} \cdot \vec{n}_i),$$

$$G_l(\vec{n}_i, \vec{n}_i) = G_l(n_i \cdot W_{n_3}, W_{n_4} \cdot \vec{n}_i),$$

where $W_{n_1} \in R^{s_1 \times s_h}, W_{n_2} \in R^{v_1 \times v_h}, W_{n_3} \in R^{s_2 \times s_h}, W_{n_4} \in R^{v_2 \times v_h}$ are embedded matrixes for pocket graph and ligand graph. Then the ligand-pocket graph is combined through the concatenation operation.

$$G_{pl}(\vec{n}_i, \vec{n}_i) = (G_p(\vec{n}_i, \vec{n}_i) || G_l(\vec{n}_i, \vec{n}_i)).$$

GVP

$$\vec{n}', \vec{n}' = GVP(n, \vec{n}),$$

$$n' = f(W_h \cdot \vec{n} || n),$$

$$\vec{n}' = f(\|W_\mu \cdot W_h \cdot \vec{n}\|_2 \odot W_\mu \cdot W_h \cdot \vec{n}).$$

where W_μ, W_h are parameterized matrixes and \odot is the row-wise multiplication operator.

GeomEncoder

GeomEncoder is the double-layer of **GeomMessage**, combined with several MLP or Linear to map the features for the identical dimension.

GeomMessage

$$(n'_i, \vec{n}'_i) = f(n_i \cdot W_{m_1}, W_{m_2} \cdot \vec{n}_i),$$

$$(e'_{ij}, \vec{e}'_{ij}) = f(e_{ij} \cdot W_{m_3}, W_{m_3} \cdot \vec{e}_{ij}),$$

$$m_i = n_i \odot (e'_{ij} \cdot W_{m_5}),$$

$$\vec{m}_i = (e'_{ij} \cdot W_{m_6}) \odot (W_{m_7} \cdot \vec{n}_i) + (n'_i \cdot W_{m_8}) \odot (W_{m_9} \cdot \vec{e}_{ij}),$$

$$(m'_i, \vec{m}'_i) = GVP(m_i, \vec{m}_i).$$

where f is the activate function, m_i is the aggregated message of node i .

Update

$$\left(\vec{n}_i', \vec{n}_i\right) = f\left(n_i \cdot W_{u_1}, W_{u_2} \cdot \vec{n}_i\right),$$

$$n'' = n' + m'$$

$$\vec{n}'' = \vec{n}' + \vec{m}'$$

Part S2. Additional results of retrospective studies on three well-studied targets.

Table S1. The mean binding energies and drug-like properties^a for the top5 molecules across three well-studied targets.

	Active	Random	GraphBP	Pocket2Mo 1	ResGen	FLAG	FragGen
AKT1/TOP5							
Vina Score (↓)	-11.940	-10.340	-9.059	-9.600	-12.62	-6.159	-11.540
QED (↑)	0.264	0.506	0.601	0.745	0.456	0.326	0.327
SA (↑)	0.266	0.314	0.454	0.264	0.422	0.580	0.692
Lipinski (↑)	4.200	4.800	5.000	5.000	5.000	5.000	4.400
LogP	2.475	1.208	1.417	0.967	1.748	0.376	1.124
CDK2/TOP5							
Vina Score (↓)	-11.540	-9.940	-9.860	-9.480	-12.800	-8.660	-11.139
QED (↑)	0.370	0.618	0.464	0.702	0.329	0.483	0.761
SA (↑)	0.742	0.766	0.430	0.508	0.480	0.620	0.726
Lipinski (↑)	4.400	5.000	4.750	5.000	4.200	5.000	5.000
LogP	-0.787	0.484	1.413	0.285	1.768	-0.032	1.857
JAK2/Top5							
Vina Score (↓)	-11.800	-10.280	-10.020	-10.640	-12.000	-8.440	-12.020
QED (↑)	0.544	0.527	0.634	0.443	0.355	0.497	0.323
SA (↑)	0.700	0.816	0.488	0.551	0.382	0.642	0.738
Lipinski (↑)	4.600	5.000	5.000	4.400	3.800	5.000	4.200
LogP	0.908	0.085	1.194	-0.573	0.483	0.712	2.467

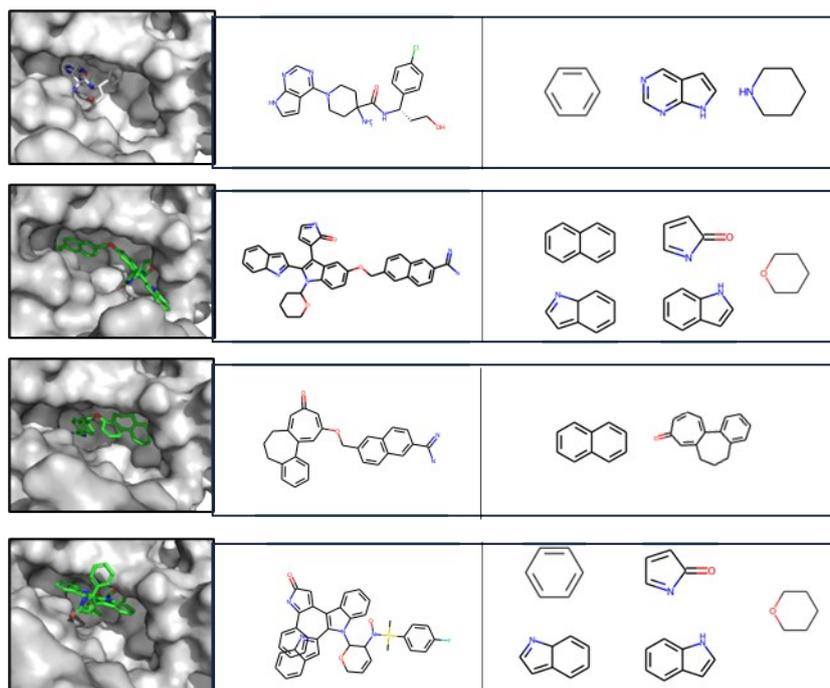


Figure S1. Fragment decomposition of crystal ligand and FragGen's top generated molecules.

Part S3. Ablation study of geometry handling protocols in FragGen.

Table S2. The ablation results of the three geometry handling protocols in FragGen.

	GeomGNN	GeomOPT	Combined
	Top5		
Vina Score (↓)	-10.483	-8.782	-9.926
QED (↑)	0.411	0.561	0.541
SA (↑)	0.579	0.716	0.740
Lipinski (↑)	4.530	4.934	4.871
LogP	0.397	0.389	0.154

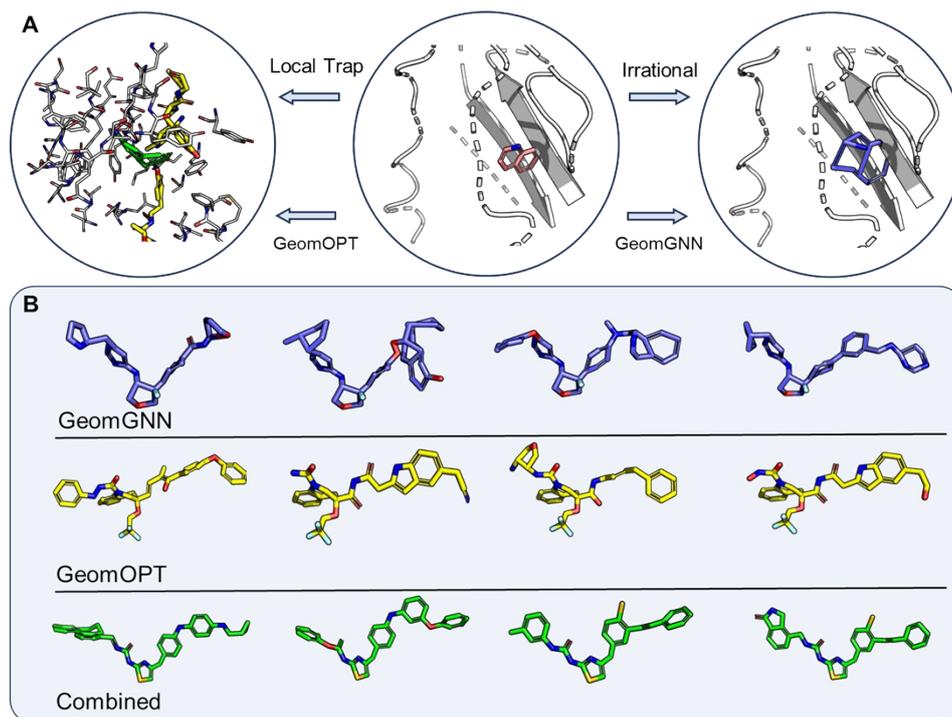
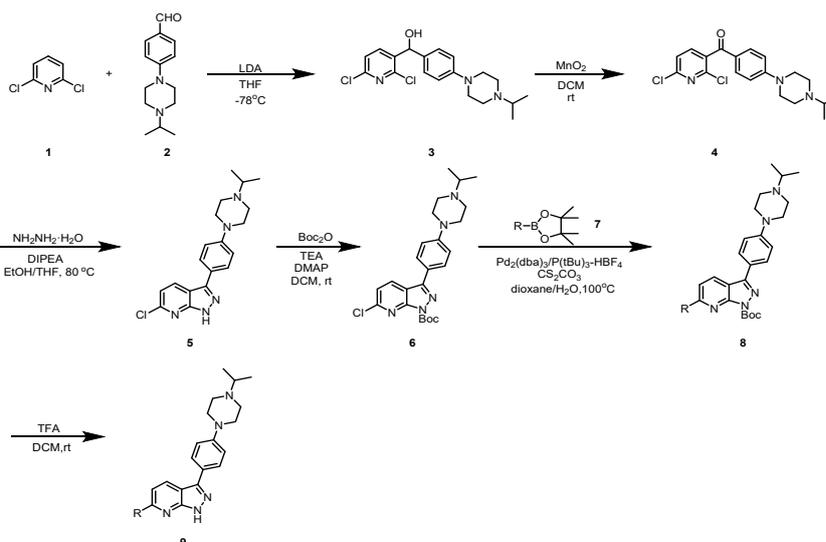


Figure S2. A). Challenges of the GeomOPT and GeomGNN protocols. B). Generated samples of three implemented protocols within FragGen.

Part S4. Synthesis routes and molecular characterization of validated compounds

Type 1



Synthesis of compound 9 (Darma-2)

The typical synthesis of compound 3

To a stirred solution of diisopropylamine (12 mmol) in dry THF (10 mL) at -78 °C under argon was added n-BuLi (1.6 M, 11 mmol) dropwise. The mixture was stirred at -78 °C for 30 min and then a solution of 2,6-dichloropyridine 1 (10 mmol) in THF (10 mL) was added slowly to the reaction system. After stirring at -78 °C for another 2 h, aldehyde 2 (12 mmol) dissolved in THF (5 mL) was added dropwise and the reaction was slowly warmed up to room temperature, aqueous NH₄Cl was added to quench the reaction and the mixture was extracted with EtOAc. The organic layer was washed with brine, dried over anhydrous Na₂SO₄ and concentrated under vacuum to obtain the residue, which was purified by silica gel column chromatography to afford compound 3 (85-92% yield).

The typical synthesis of compound 4

A solution of compound 3 (10 mmol) and MnO₂ (50 mmol) in DCM (20 mL) was stirred at room temperature overnight. The reaction mixture was filtered through a pad of celite and the filtrate was concentrated under reduced pressure. Purification by flash column chromatography afforded compound 4 (88-95 % yield).

The typical synthesis of compound 5

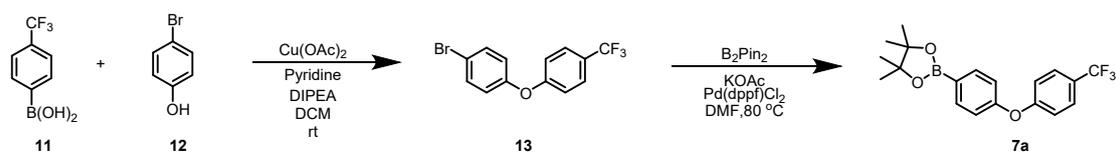
To a solution of compound 4 (10 mmol) in EtOH/THF (v/v = 20 mL/5 mL) was added DIPEA (10 mmol) at 0 °C and subsequently, hydrazine monohydrate (98 %, 11 mmol) was added dropwise. The mixture was stirred at 0 °C for 1 h and was warmed up to room temperature for 1 h. Finally, the mixture was heated at 70 °C overnight. The reaction mixture was concentrated and water was added to the mixture and extracted with EtOAc. The organic layer was washed with water and brine, dried over Na₂SO₄ and concentrated under reduced pressure. Purification by flash column chromatography afforded compound 5 (83-90 % yield).

The typical synthesis of compound 6

To a solution of compound 5 (10 mmol) and DMAP (1 mmol) in dry DCM (15 mL)

under argon was added Et₃N (30 mmol) and a solution of Boc₂O (15 mmol) in DCM (5 mL). Then the mixture was heated to reflux overnight. The reaction mixture was diluted with DCM and washed with water and brine. The organic layer was dried over Na₂SO₄ and concentrated under reduced pressure. Purification by flash column chromatography afforded compound **6** (87-94 % yield).

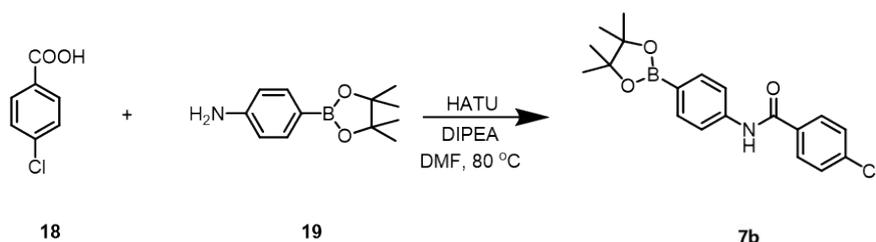
The typical synthesis of phenylboronic acid **7**



A two-neck flask equipped with a magnetic stir bar was charged with **11** (15 mmol), **12** (10 mmol), Cu(OAc)₂ (1 mmol), Pyridine (5 mmol), DIPEA (5 mmol), and then purged with argon three times. DCM (20 mL) was added as solvent and the mixture was stirred at ambient temperature for 3 hours. After completion, water was added to quench the reaction and the mixture was extracted with EtOAc. The organic layer was washed with brine, dried over anhydrous Na₂SO₄ and concentrated under vacuum to obtain the residue, which was purified by silica gel column chromatography to afford **13** as white solid (97 % yield).

A solution of compound **13** (5 mmol), bis(pinacolato)diboron (6.5 mmol), Pd(dppf)Cl₂ (0.25 mmol) and KOAc (15 mmol) in dioxane (10 mL) under argon was heated at 80 °C overnight. The mixture was filtered through a pad of celite and the residue was washed with EtOAc. The filtrate was concentrated under reduced pressure and purified by flash column chromatography to afford compound **7a** as white solid (94 % yield).

The typical synthesis of phenylboronic acid **7b**



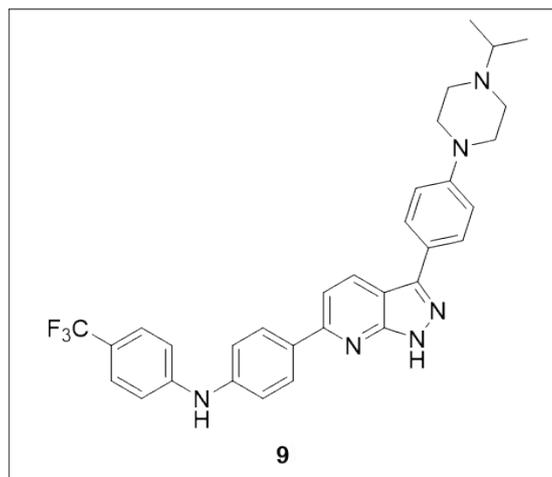
A two-neck flask equipped with a magnetic stir bar was charged with **18** (10 mmol), **19** (11 mmol) and then purged with argon three times. DMF (20 mL) was added as solvent. The reaction was stirred at 80 °C for 3 hours. After completion, water was added to quench the reaction and the mixture was extracted with EtOAc. The organic layer was washed with brine, dried over anhydrous Na₂SO₄ and concentrated under vacuum to obtain the residue, which was purified by silica gel column chromatography to afford the **7b** as white solid (89 % yield).

The typical synthesis of compound 8

A solution of compound **6** (1 mmol), phenylboronic acid **7** (2 mmol), Pd₂(dba)₃ (0.05 mmol), P(*t*-Bu)₃HBF₄ (0.2 mmol) and Cs₂CO₃ (3 mmol) in dioxane/H₂O (v/v = 5 mL/0.5 mL) under argon was heated at 100 °C overnight. The reaction mixture was filtered through a pad of celite and the residue was washed with EtOAc. The filtrate was washed with water and brine, dried over Na₂SO₄ and concentrated under reduced pressure. Purification by flash column chromatography afforded compound **8** (65-80 % yield).

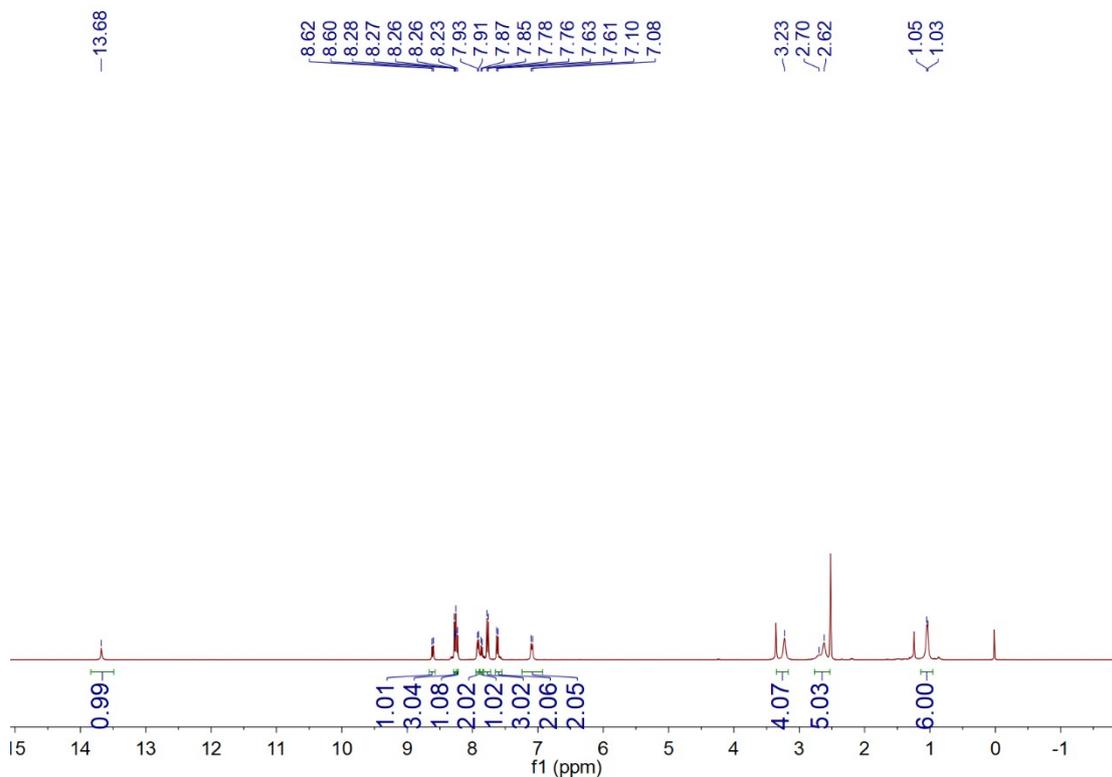
The typical synthesis of compound 9

To a solution of compound **8** (0.5 mmol) in DCM (5 mL) was added TFA (2.5 mmol) and the reaction was stirred at room temperature overnight. The reaction mixture was concentrated, quenched with saturated NaHCO₃ and extracted with EtOAc. The organic layer was washed with water and brine, dried over Na₂SO₄ and concentrated under reduced pressure. Purification by flash column chromatography afforded the desired product **9** (64-77 % yield).

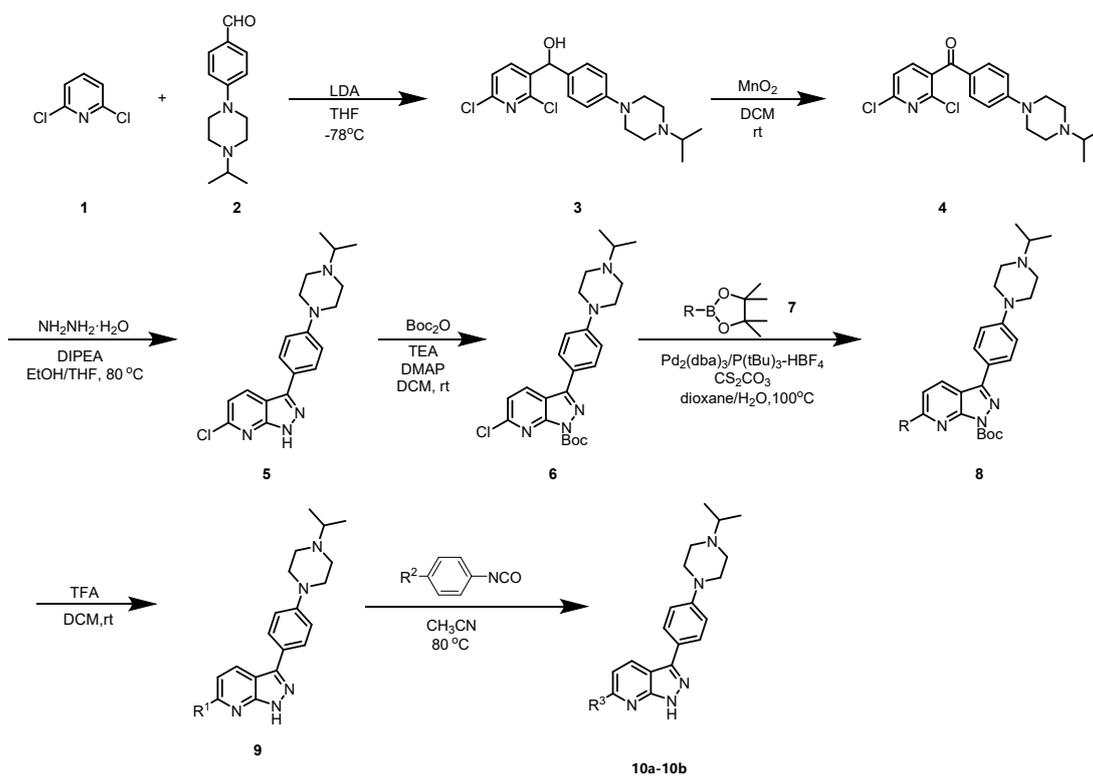


4-(3-(4-(4-isopropylpiperazin-1-yl)phenyl)-1H-pyrazolo[3,4-b]pyridin-6-yl)-N-(4-(trifluoromethyl)phenyl)aniline (9a) – Darma2

Yellow solid. ¹H NMR (400 MHz, DMSO-*d*₆) δ 13.68 (s, 1H), 8.61 (d, *J* = 8.4 Hz, 1H), 8.37 – 8.15 (m, 3H), 8.25 – 8.22 (m, 1H), 7.92 (d, *J* = 8.4 Hz, 2H), 7.86 (d, *J* = 8.4 Hz, 1H), 7.77 (d, *J* = 8.8 Hz, 3H), 7.62 (d, *J* = 8.4 Hz, 2H), 7.09 (d, *J* = 8.4 Hz, 2H), 3.26 – 3.13 (br, 4H), 2.78 – 2.59 (br, 5H), 1.04 (d, *J* = 6.4 Hz, 6H).



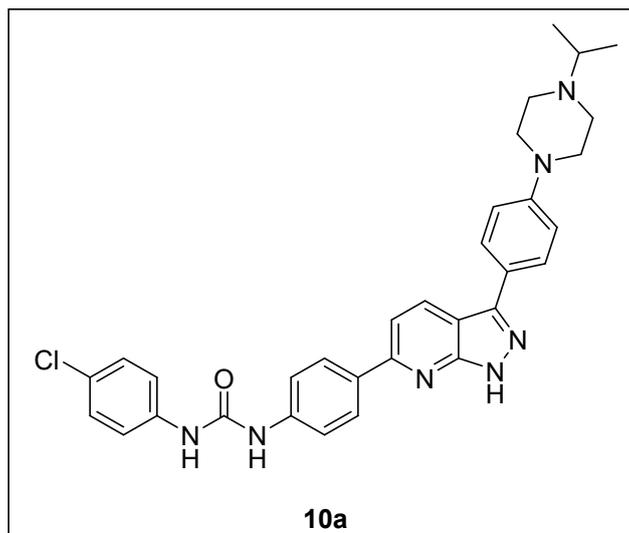
Type 2



Synthesis of Type 2 compounds (10a-Darma1; 10b-Darma3)

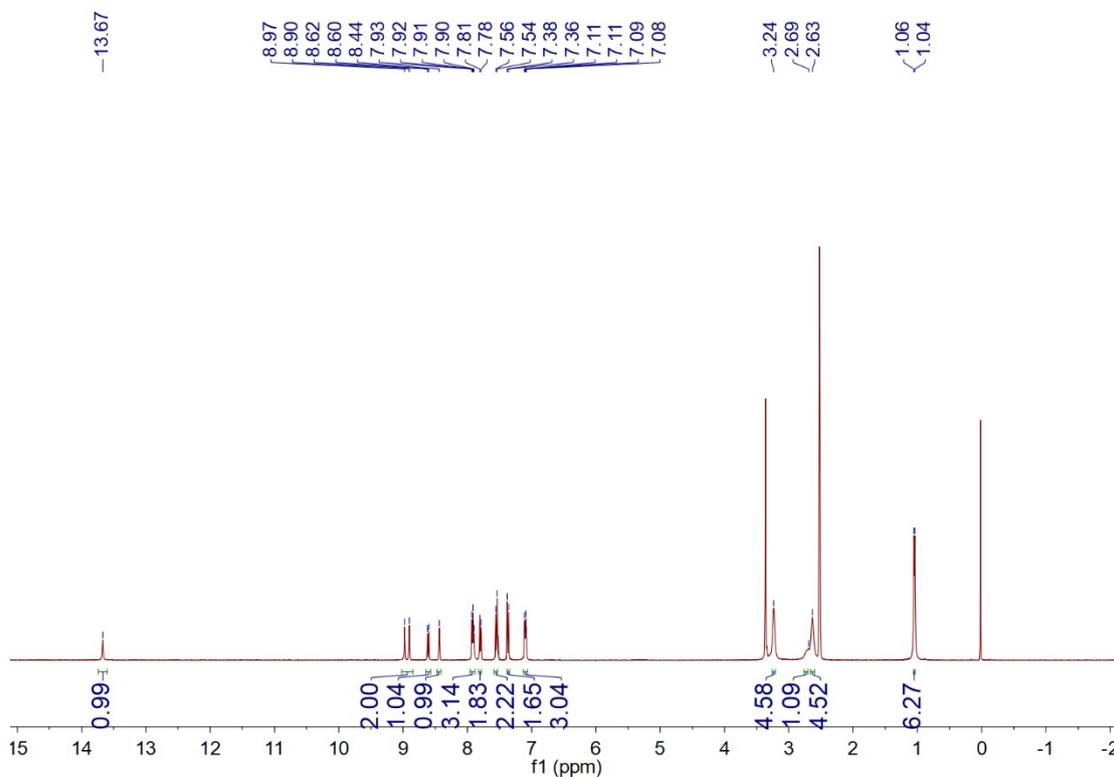
The typical synthesis of compound 10

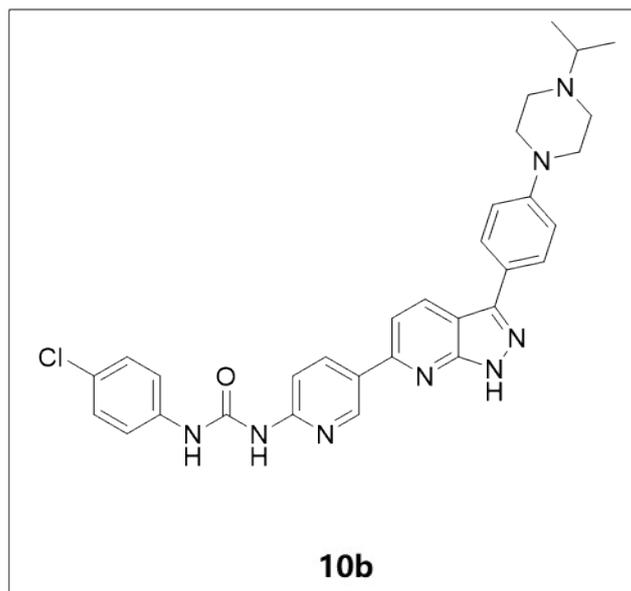
To a solution of compound **9** (0.5 mmol) in CH₃CN (5 mL) was added Isocyanate (0.55 mmol) and the reaction was stirred at room 80 °C overnight. After completion, the mixture was concentrated under vacuum to obtain the residue, which was purified by silica gel column chromatography to afford the **10** (50-57 % yield).



1-(4-chlorophenyl)-3-(4-(3-(4-(4-isopropylpiperazin-1-yl)phenyl)phenyl)-1*H*-pyrazolo[3,4-*b*]pyridin-6-yl)phenyl)urea (10a) – Darma1

Yellow solid. ¹H NMR (400 MHz, DMSO-*d*₆) δ 13.67 (s, 1H), 8.94 (m, 2H), 8.61 (d, *J* = 8.4 Hz, 1H), 8.44 (s, 1H), 7.91 (dd, *J* = 8.8, 5.2 Hz, 3H), 7.80 (d, *J* = 8.8 Hz, 2H), 7.55 (d, *J* = 8.8 Hz, 2H), 7.37 (d, *J* = 8.8 Hz, 2H), 7.10 (dd, *J* = 9.2, 3.2 Hz, 3H), 3.24 (br, 4H), 2.69 (m, 1H), 2.63 (br, 4H), 1.05 (d, *J* = 6.4 Hz, 6H).





1-(4-chlorophenyl)-3-(5-(3-(4-(4-isopropylpiperazin-1-yl)phenyl)-1H-pyrazolo[3,4-*b*]pyridin-6-yl)pyridin-2-yl)urea (10b) – Darma3

Yellow solid. ¹H NMR (400 MHz, Chloroform-*d*) δ 8.64 (d, *J* = 8.4 Hz, 1H), 7.99 – 7.88 (m, 6H), 7.56 (m, 1H), 7.40 (m, 1H), 7.19 – 7.13 (m, 1H), 7.09 (d, *J* = 8.4 Hz, 2H), 6.99 (m, 1H), 3.25 (br, 4H), 2.77 (m, 1H), 2.67 (br, 4H), 1.06 (d, *J* = 6.4 Hz, 6H).

