Accepted Manuscript

Discovery, optimization and biological evaluation for novel c-Met kinase inhibitors

Haoliang Yuan, Qiufeng Liu, Li Zhang, Shihe Hu, Tiantian Chen, Huifang Li, Yadong Chen, Yechun Xu, Tao Lu

PII: S0223-5234(17)30976-5

DOI: 10.1016/j.ejmech.2017.11.073

Reference: EJMECH 9948

To appear in: *European Journal of Medicinal Chemistry*

Received Date: 23 September 2017

Revised Date: 23 November 2017

Accepted Date: 26 November 2017

Please cite this article as: H. Yuan, Q. Liu, L. Zhang, S. Hu, T. Chen, H. Li, Y. Chen, Y. Xu, T. Lu, Discovery, optimization and biological evaluation for novel c-Met kinase inhibitors, *European Journal of Medicinal Chemistry* (2017), doi: 10.1016/j.ejmech.2017.11.073.

Discovery, Optimization and Biological Evaluation for Novel c-Met kinase Inhibitors

Haoliang Yuan,[ab#] Qiufeng Liu,[cde#] Li Zhang,[a#] Shihe Hu,[a] Tiantian Chen,[c] Huifang Li,[f] Yadong Chen,[a]

Yechun Xu[c*] and Tao Lu[ab*]

[a] Laboratory of molecular design and drug discovery, School of Science, [b] State Key Laboratory of Natural Medicines, Jiangsu Key Laboratory of Drug Discovery for Metabolic Disease, Institute of Pharmaceutical Research, China Pharmaceutical University, 24 Tongjiaxiang, Nanjing 210009, China.

[c] CAS Key Laboratory of Receptor Research, Drug Discovery and Design Center, Shanghai Institute of Materia Medica, Chinese Academy of Sciences (CAS), Shanghai 201203, China.

[d] School of Life Science and Technology, ShanghaiTech University, Shanghai, 200031, China

[e] University of Chinese Academy of Sciences, Beijing, 100049, China

[f] Vancouver Prostate Centre, University of British Columbia, 2660 Oak Street, Vancouver, British Columbia V6H 3Z6, Canada

# Authors contributed equally to this work.

Corresponding Author

[*]Yechun Xu

Phone:86-21-50801267; Fax: 86-21-50801267; E-mail: ycxu@simm.ac.cn.

[*]Tao Lu

Phone: 86-25-86185086; Fax: 86-25-83302827; E-mail: lutao@cpu.edu.cn.

Notes

The authors declare no competing interest.

ABSTRACT

The c-Met kinase has emerged as an attractive target for developing antitumor agents because of its close relationship with the development of many human cancers, poor clinical outcomes and even drug resistance. A series of novel c-Met kinase inhibitors have been identified with multiple workflow in this work, including virtual screening, X-ray crystallography, biological evaluation and structural optimization. The experimentally determined crystal structure of the best hit compound **HL-11** in c-Met kinase domain was highly consistent with the computational prediction. Comparison of the hit compounds with different c-Met kinase inhibitory activity by molecular dynamics simulations suggested the key protein-ligand interactions for structural optimization. Based on these, structural optimization produced compound **11e** with better c-Met kinase inhibitory activity and improved anti-proliferative activity. These experimental findings proved the reliability and efficiency of our in silico methods. This strategy will facilitate further lead discovery and optimization for novel c-Met kinase inhibitors.

Keywords

c-Met, Virtual Screening, Pharmacophore, X-ray Crystallography, Structural Optimization

1. Introduction

The c-Met kinase is a receptor tyrosine kinase that expressed in endothelial and epithelial cells. Hepatocyte growth factor (HGF) is the natural ligand of c-Met, also known as a scatter factor. The HGF/c-Met signaling pathway plays important roles in normal development, organogenesis, and homeostasis [1]. The activation of c-Met by HGF causes receptor dimerization and autophosphorylation of tyrosine 1234 and 1235. Phosphorylated c-Met further triggers the activation of downstream signaling pathways such as the Ras/MAPK, c-Src, and PI3K/Akt pathways. Meanwhile, c-Met induces an invasive program consisting of cell proliferation, migration, invasion, and survival that is essential for physiological events during normal processes such as morphogenesis, liver regeneration, and wound healing. Aberrant HGF/c-Met signaling through constitutive activation, gene amplification, mutations, and activation of an autocrine loop occurs in virtually all types of solid tumors, and as such is implicated in multiple tumor oncogenic processes, such as mitogenesis, survival, angiogenesis, and invasive growth, especially in the metastatic process [2-5]. Furthermore, overexpression of c-Met and HGF was demonstrated to correlate with poor prognosis or metastatic progression in a number of major human cancers, including lung, prostate, renal, ovarian, gastric, and liver cancers [6]. In this sense, c-Met has recently attracted considerable interest as a therapeutic target for various cancers [7].

The c-Met kinase inhibitors are grouped according to their binding mode in the binding site. Compounds with a U-shaped binding mode in the ATP binding site are Type I c-Met kinase inhibitors. In general, compounds of this type exhibit relatively good cellular selectivity profiles with a limited number of off-target kinase hits, such as Crizotinib and PF-04217903. They compete with ATP to bind in the ATP binding site and mimic the interactions between ATP and the binding site, including the hydrogen bond interactions with the residues in the hinge region (Pro1158, Tyr1159 and Met1160). Interaction with the activation loop is also important to kinase inhibitors, including DFG motif and Tyr1230 in c-Met kinase. The DFG motif (Asp1222, Phe1223 and Gly1224) is conserved in kinase, interactions with which are beneficial for improving kinase inhibitory activity. Furthermore, Tyr1230 in c-Met kinase domain is unique so that compounds got interactions with it showed good selectivity. Type II c-Met kinase inhibitors are multi-targeted c-Met kinase inhibitors that pass the gatekeeper and occupy the deep hydrophobic back pocket. A significant movement of the A-loop is required for entering the c-Met back pocket which is compensated by additional interactions with the hydrophobic back pocket, often leading to compounds with high molecular weights and high lipophilicities. But hydrogen bond interactions with the hinge region are

the shared protein-ligand interactions of Type I and II c-Met kinase inhibitors. The representative c-Met inhibitors of Type I, II and other types in different clinical phase studies have been shown in Fig. 1. Especially, Crizotinib (c-Met/ALK) and Cabozantinib (VEGFR2/c-Met) have been approved for the treatment of non-small cell lung carcinoma [6]. Besides, some other c-Met kinase inhibitors have also been put into clinical studies, including PF-04217903 and JNJ-38877605 of Type I, BMS-777607, GSK1363089 and MGCD-265 of Type II.

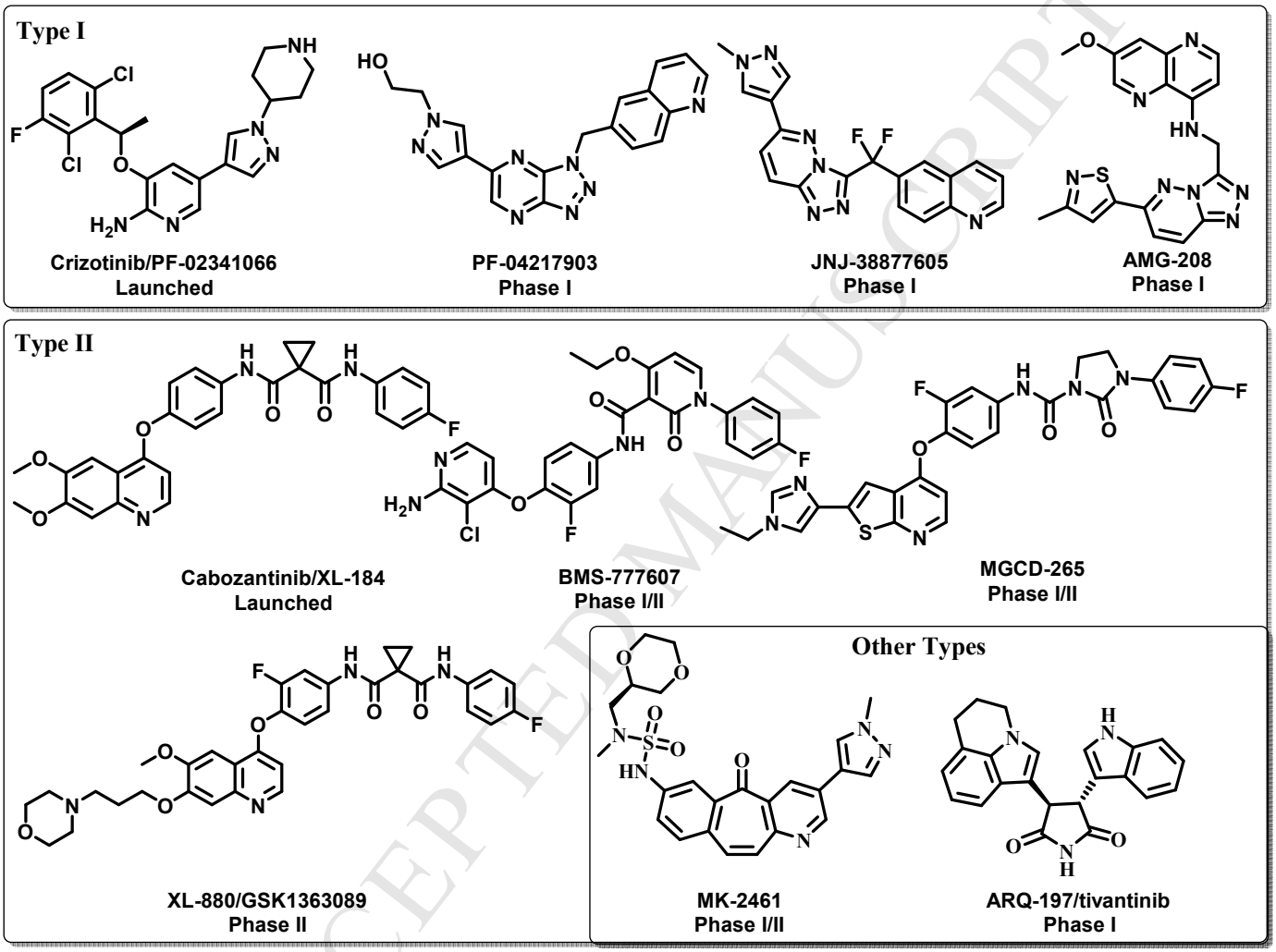


Fig. 1 The representative c-Met kinase inhibitors of different structural types.

Encouraged by these findings, a multiple workflow, including virtual screening, biological evaluation, X-ray crystallography, molecular dynamics simulation [8] and structural optimization, has been applied to identify novel c-Met inhibitors in this study. Complex-based pharmacophore model and molecular docking were used as the structure-based approaches, while the ligand-based approaches were ligand-based pharmacophore we have developed [9], 2D chemical similarity, and 3D-shape. These two types of approaches focused on different aspects and complemented each other. A series of novel and potent c-Met inhibitors have been identified after biological evaluation. The experimentally resolved crystal structure of

the best hit compound **HL-11** in complex with c-Met kinase domain validated the computational prediction. Structural optimization yielded compound **11e** with better c-Met kinase inhibitory activity and significantly improved anti-proliferative activity.

2. Materials and Methods

2.1 Structure-Based Virtual Screening

Molecular Docking. The available programs, Glide, Gold, Surflex and CDocker, were used for docking. As a test set, 11 ligands in the crystal structures of the kinase domain of c-Met were extracted and re-docked into the binding site. The PDB IDs for the 11 complex structures were 2RFS [10], 2WD1 [11], 2WGJ [12], 2WKM [13], 3A4P [14], 3CCN [15], 3CD8 [15], 3DKF [16], 3DKG [16], 3F66 [17], and 3I5N [18]. The root mean square deviation (RMSD) between the docked and crystal conformations of the ligand was a criteria to assess the docking accuracy.

Glide. The protein in the crystal structures was prepared using the *Protein Preparation Wizard* workflow. The generated receptor grid was centered on the ligand in the crystal structure, which was defined as the ligand-binding site search region. The compounds to be docked was confirmed by an enclosing box that was similar in size to the crystal ligand. The compounds were prepared with *LigPrep* and docked into the binding site using *Glide* standard/extra precision (SP/XP) mode. Default settings were used for all the other parameters. The best pose was determined by the protein-ligand interactions and Glide score. All these procedures were completed in Schrödinger 2009.

Gold. *Gold* uses a genetic algorithm to explore the binding site for bioactive ligands. Protein structures were prepared by adding hydrogen atoms and deleting water molecules as well as the ligand in Sybyl 6.9. A 10 Å radius surrounding the ligand was defined as the ligand binding site. Only the best conformation of the docked ligand was included in the output and other settings in the genetic algorithm were left default.

Surflex. An empirical scoring function was used to dock ligands into the binding site. The active site was defined as protomol, an idealized ligand that makes every potential interaction with the binding site. It was used to align the ligand fragments and also provide the direction to grow the fragments. All the preparations were completed in Sybyl 6.9.

CDocker. CDocker is a grid-based molecular docking method that employs CHARMm. During the docking procedure, the receptor keeps rigid while ligand is flexible. Protein and ligands were prepared in DS 2.5 (Accelrys Discovery Studio 2.5). The binding site was defined based on the volume occupied by the bound ligand. Only the best conformation of the docked ligand was output.

Generation of Complex-based Pharmacophore Models. We focused on developing a complex-based pharmacophore model for type I c-Met inhibitors. At the outset of this study, only 11 crystal structures of c-Met kinase domain in complex with type I inhibitors were available in PDB. All these complex structures were aligned by Cα atoms so that the active sites were thus aligned. Water molecules were retained in order to include the possible water-mediated hydrogen bonds.

Pharmacophore model was generated for each complex based on the aligned structures. The protein-ligand interactions were converted to pharmacophoric features, including hydrophobic centers, aromatic rings, and hydrogen bond donors and acceptors along with their direction vectors. Excluded volume spheres representing different spatial positions occupied by adjacent residues were automatically added to the final model. All features in the 11 pharmacophore models were clustered according to different types of interactions with the active site. Excluded volume spheres used as steric constraints were clustered based on the specific residues they represented. The cluster centers were identified using *Edit and Cluster Pharmacophore* tool in DS 2.5. Finally, the pharmacophore model was complemented with 3D-shape constraints to improve the selectivity of the c-Met inhibitors.

Validation of Complex-based Pharmacophore Models. Retrospective discovery of known c-Met inhibitors was used to validate the pharmacophores' capability of distinguishing active inhibitors from inactive compounds. A database included other 86 reported type I c-Met inhibitors and 538 randomly selected decoy compounds was constructed. Decoys were collected from FDA-approved small molecule drug without reported biological activities on c-Met in Drugbank (*https://www.drugbank.ca/*). Structures of all these molecules were prepared using SciTegic Pipeline Pilot 7.5. Multi-conformations of each compound were generated using the FAST conformational search protocol implemented in DS 2.5 with an energy threshold of 20 kcal/mol and a maximum of 255 conformers. A multi-conformation database included all the conformers of 624 compounds was resulted. All screening experiments were performed by using *Best Flexible Search* algorithm. Enrichment Factor (EF), Goodness of Hit (GH) and receiver operating characteristic (ROC) scores were used to evaluate the performance of pharmacophore screening.

2.2 Ligand-Based Virtual Screening

Ligand-based Pharmacophore Modeling. A HypoGen ligand-based pharmacophore model has been developed and rigorously validated in our previous study [9]. The pharmacophore was generated using 23 known c-Met inhibitors of diverse structures. It consisted of two Hydrogen Bond Acceptors (HBA), one Hydrophobic center (HY) and one Ring Aromatic (RA). The capability of it in identifying active

compounds from inactive ones was illustrated with an enrichment factor of 21.07 and a ROC score of 0.896. It was suggested that such a phamacophore model could serve as a reliable tool for the discovery of new c-Met inhibitors.

2D Chemical Similarity and 3D-shape. Molecular similarity was calculated using SciTegic Pipeline Pilot 7.5. The c-Met inhibitors collected from literatures were used as the reference compounds. Tanimoto coefficient between the ChemDiv and reference compounds was calculated. 3D-shape was applied in combination with the complex-based pharmacophore model.

2.3 Crystal Structure Selection and Database Preparation

Crystal Structure Selection. Cross-docking was performed to select the most suitable crystal structure for the structure-based virtual screening. The complex structures were firstly superimposed onto the reference structure (PDB ID: 2WGJ). The co-crystal ligands were extracted and prepared before they were re-docked using the selected docking programs. RMSD values between the crystal and docked conformations were considered as the criteria for crystal structure selection. Because the docking accuracy may be affected by the type of the bound inhibitor, a representative structure for each structural type of inhibitor was used in the structure-based virtual screening.

Database Preparation. The ChemDiv compound library containing 0.7 million compounds was used for virtual screening in this study. For molecular docking, all the structures in the compound database were prepared using SciTegic Pipeline Pilot 7.5. They were protonated, added with partial charges, and then minimized with the MMFF94x force field to a gradient of 0.0001 kcal/mol Å. For the pharmacophore-based virtual screening, all the structures were subjected to conformation modeling using the Catalyst module of DS 2.5. Multi-conformations of each compound were generated using the *FAST* conformational search protocol implemented in DS 2.5 with an energy threshold of 20 kcal/mol and a maximum of 255 conformers. These compounds were converted into a searchable multi-conformation database.

2.4 Further Scoring

The compounds passed through the parallel virtual screening lines and cascade molecular docking were then further scored. First, they were energy minimized in MOE 2009 (Molecular Operating Environment, version 2009.10) to obtain their LigX_p*K*i binding affinity, which is mainly presented by the energy of hydrogen bonds and hydrophobic interactions. Second, they were scored using molecular mechanics, the generalized Born model, and the solvent accessibility (MM-GB/SA) method with OPLS_2005 and GB/SA

in MacroModel of Schrödinger 2009 to calculate the free energies (E_Macromodel) of the optimal chemical conformations. Third, another module implemented in Schrödinger 2009, Prime MM-GBSA which calculates ligand binding energies and ligand strain energies between a set of ligands and a single receptor, was also applied to calculate the binding free energies between the compounds and kinase domain of c-Met. Including Glide Gscore, four different scoring methods in total were applied to select the potential compounds from the docking results.

2.5 Kinase inhibitory and anti-proliferative assays

Chemical Source. The selected compounds were all purchased from Topscience (http://www.tsbiochem.com/). The structure and purity of all chemicals were confirmed by mass spectroscopy (MS) and liquid chromatography-tandem mass spectrometry (LC-MS/MS).

Biochemical kinase Assays. The c-Met kinase inhibitory activity was determined using Hot-SpotSM kinase assay by Reaction Biology Corp. (Malvern PA, USA). After the substrate was prepared in freshly prepared reaction buffer (20 mM Hepes pH 7.5, 10 mM $MgCl_2$, 1 mM EGTA, 0.02% Brij35, 0.02 mg/mL BSA, 0.1 mM Na_3VO_4, 2 mM DTT, 1% DMSO), 5 nM of human GST-tagged target kinase was delivered into the substrate solution and mixed gently. The testing compounds were dissolved in 100% DMSO to specific concentration and added into the kinase reaction mixture by Acoustic technology (Echo550; nanoliter range). The reaction mixture was incubated for 20 min at room temperature. ^{33}P-ATP (Specific activity 10 μCi/μL) was delivered into the reaction mixture to initiate the reaction and incubated at room temperature for 2 h. The kinase activities were detected by filter-binding method. IC_{50} values were obtained using GraphPad Prism (San Diego, CA).

AlamarBlue® Cell proliferation Assays. The human tumor cell lines used in preliminary anti-proliferation assays were A549 (lung), HCT116 (colon), MCF-7 (breast), PC3 (prostate) and SGC-7901 (gastric) initially, and EBC-1 (lung), MKN45 (gastric), NCI-H460 (lung), PC3, Colo-205 (colon) were used for anti-proliferation assays of new compounds obtained after structural optimization. All experiments were done under standard conditions (37 °C and 5% CO_2). Cells were seeded in 384-well plates in medium supplemented with 10% fetal bovine serum (FBS). After 24 hours, the wells were changed with serum-free medium [with 0.04% bovine serum albumin (BSA)]. Cells were treated with test compounds for 72 hours. Add 4 μl alamarBlue® reagent directly to cells in 40 μl culture medium. Incubate for 1 to 4 hours at 37 °C in a cell culture incubator, protected from direct light. Detect the fluorescence intensity at excitation wavelength of 540-570 nm (peak excitation is 570 nm) and fluorescence emission at

580-610 nm (peak emission is 585 nm) with the Safire microplate Reader (Tecan, Switzerland). The IC_{50} values were calculated by fitting the data to the equation for competitive inhibition using nonlinear regression method (GraphPad Prism, San Diego, CA).

2.6 Crystal Structure Determination

*Protein Purification and Crystallization***.** Production of the kinase domain (1038-1346 aa) of recombinant human c-Met followed the protocols of Wang [19]. with certain modifications. The cDNA fragment was cloned into the vector pET28a and the protein was co-expressed with catYopH subcloned in pET15b (164-468AA) [20]. The expressed c-Met kinase domain was passed through a Ni-NTA column (Qiagen) and further purified by QHP ion exchange column (GE) which eluted with 25 mM Tris pH 8.5, 100 mM NaCl, 10% glycerol, 1mM DTT. The protein was concentrated to ~10 mg/mL for further crystallization.

Cocrystallization of the c-Met kinase domain with compound **HL-11** was carried out by mixing a solution of the protein-ligand complex with an equal volume of precipitant solution (0.1M Tris pH7.5, 15% glycerol, 12% MPD,5% isopropanol, 15% PEG5KMME). The protein-ligand complex was prepared by adding the compound to the protein solution to a final concentration of 1 mM of **HL-11**. Cocrystallization utilized the vapour-diffusion method in hanging drops. Crystals were flash frozen in liquid nitrogen in the presence of well solution supplemented with 25% glycerol.

Structure Determination and Refinement. Data were collected at 100 K on beamline BL17U at the Shanghai Synchrotron Radiation Facility (SSRF), and were processed with the XDS [21] software packages. The structure was solved by molecular replacement, using the program PHASER [13] with the search model of PDB ID 4GG5 [22]. The structure was refined with PHENIX [23]. With the aid of the program Coot [24], compound **HL-11**, water molecules were fitted into to the initial F_o-F_c map. The complete statistics, as well as the quality of the solved structures, are shown in Table S1.

2.7 Molecular Dynamics Simulation

Molecular dynamics simulations were performed using AMBER 9.0 software package with the ff99SB force field to simulate c-Met in complex with its inhibitors and calculate their binding free energy. The co-complex structures were obtained by molecular docking. The ligands were firstly fully minimized by the AM1 method and electrostatic potentials computed at the HF/6-31G* level in the Gaussian 09 program. The RESP fitting technique in AMBER was used to determine the partial charges. The force-field parameters for the ligand were generated with the general AMBER force field (GAFF) by Antechamber.

Hydrogen atoms were assigned using the LEaP module, which sets ionisable residues to their default protonation states at a neutral pH value. Each complex was immersed in a cubic box of TIP3P water model with a 10 Å minimum solute-wall distance. A total of five Cl^- ions were added to neutralize each protein-ligand complex system. Energy was minimized in the solvated system employing 1000 steps of the steepest descent algorithm and 2000 steps of the conjugate gradient algorithm with a nonbonded cutoff of 10Å. The protocol for molecular dynamics simulation consisting of gradual heating, density equilibration, equilibration and production procedures in an isothermal isobaric ensemble (*NPT*, *P* = 1 atm and T = 298K) MD. The system was gradually heated from 0 to 298K in 50ps, followed by density equilibration at 298K for 500ps, and then constant equilibration at 298K for 500ps. Then each protein-ligand complex system underwent a process of equilibration procedure until the system achieved a continuous stable status, i.e. production stage. The time step was set to 2 fs while the snapshots were taken every 10 ps to record the conformation trajectory during production MD. The nonbonded interactions were treated with a 10Å cutoff. SHAKE algorithm was applied to constrain all bonds involving hydrogen atoms to their equilibrium length. The conformations of different systems were collected every 50ps after the system achieved their equilibrium status. The collected snapshots were used for structural and energetic analysis of each complex system.

Binding free energy calculations were also performed to investigate binding affinity between different ligands and the binding pocket. The methods like MMPBSA [25] and MM/GBSA [26] are usually used for investigating the energetic contribution of protein-ligand binding affinities. For different protein-ligand systems, the process of converged status was used for the binding free energy calculation. The solute and solvent dielectric constants were set as 1.0 and 80.0, respectively. The binding free energy of different ligands to the protein was calculated as follows:

$$\Delta G_{binding}=G_{complex}-[G_{protein}+G_{ligand}]$$

For different protein-ligand systems, the process for binding free energy calculations were also applied for hydrogen bond occupancy calculations. The hydrogen bond distance was set as 3.5Å and angle was 120.0°. Other parameters were kept default.

2.8 Chemistry

^{1}H NMR and ^{13}C NMR spectra were obtained at room temperature using a Bruker Avance 300 spectrometer. DMSO-d_6 or $CDCl_3$ containing 0.03% tetramethylsilane (TMS) (99.8% D, Adamas, Inc.) was used as a solvent for NMR measurements. Chemical shifts (δ) for ^{1}H NMR are given in parts per million

(ppm) relative to residual DMSO (δ 2.50 ppm) or H_2O (δ 3.33 ppm). Chemical shifts (δ) for ^{13}C NMR are given in ppm relative to DMSO (δ 39.50 ppm). Splitting patterns are designated as follows: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad. High resolution mass spectrometry (HRMS) was obtained on a Q-TOF micro spectrometer. Melting points were determined with a Micro melting point apparatus. TLC plates were visualized by exposure to ultraviolet light.

2.8.1 4-((4-amino-5-mercapto-4H-1,2,4-triazol-3-yl)methyl)phenol (1)

2-(4-hydroxyphenyl)acetic acid (5.02 g, 33 mmol), hydrazinecarbothiohydrazide (3.80 g, 36 mmol) and CH_3SO_3H (6.91 g, 72mmol) were dissolved in 8 mL sulfolane and 6 mL H_2O. The solution was stirred at 90 ^oC for 24 h. After cooled to room temperature, the mixture was poured into 50 mL ice water. Saturated aqueous sodium carbonate solution was added to adjust the PH to 7-8. The mixture was filtered and dried in vacuo to give 3.60 g of **1** as blue solid: 50% yield. ^1H NMR (300 MHz, DMSO-d_6) δ 13.52 (s, 1H), 9.35 (s, 1H), 7.07 (d, J = 8.0 Hz, 2H), 6.69 (d, J = 7.3 Hz, 2H), 5.54 (s, 2H), 3.89 (s, 2H); m.p.: 206-208^oC; ESI-MS m/z: 223.2 $[M-H]^-$.

2.8.2 4-((6-(1H-indol-2-yl)-[1,2,4]triazolo[3,4-b][1,3,4]thiadiazol-3-yl) methyl)phenyl 1H-indole-2-carboxylate (2a)

1 (1.00 g, 4.50 mmol), 1*H*-indole-2-carboxylic acid (0.60 g, 5 mmol) were dissolved in 3 mL sulfolane, and 12 mL $POCl_3$ was added. The reaction mixture was stirred at 85 ^oC for 18 h, then cooled to room temperature, poured into 200 mL ice water. NaOH was added to adjust the PH to 7. The mixture was filtered and dried in vacuo to give 0.60 g of **2a** as white solid: 27% yield. ^1H NMR (300 MHz, DMSO-d_6) δ 9.35 (s, 1H), 8.60 (s, 1H), 8.09 (s, 1H), 7.14 (d, J = 7.9 Hz, 2H), 6.75 – 6.66 (m, 2H), 4.26 (s, 2H), 3.93 (s, 3H); m.p.: 280-281^oC; ESI-MS m/z: 491.3 $[M+H]^+$.

2.8.3 4-((6-(1-methyl-1H-pyrazol-4-yl)-[1,2,4]triazolo[3,4-b][1,3,4] thiadiazol-3-yl)methyl)phenyl 1-methyl-1H-pyrazole-4-carboxylate (2b)

Synthetic procedure was followed as described for compound **2a** to give **2b** as white solid: 27% yield. ^1H NMR (300 MHz, DMSO-d_6) δ 8.61 (s, 1H), 8.53 (s, 1H), 8.10 (s, 1H), 8.01 (s, 1H), 7.42 (d, J = 8.2 Hz, 2H), 7.18 (d, J = 8.0 Hz, 2H), 4.45 (s, 2H), 3.93 (s, 3H), 3.92 (s, 3H); m.p.: 167-168 ^oC; ESI-MS m/z: 421.2 $[M+H]^+$.

2.8.4 4-((6-(1H-indol-2-yl)-[1,2,4]triazolo[3,4-b][1,3,4] thiadiazol-3-yl)methyl)phenol (3a)

2a (0.49 g, 1.00mmol) was dissolved in 15 mL THF and NaOH (0.2 g, 5 mmol) was added. The mixture was stirred at room temperature for 24 h. 1N HCl was added to adjust the PH to 7 and the solution was

extracted with ethyl acetate (20 mL x 2). The organic layer was washed with brine (30 mL), dried over Na_2SO_4 and concentrated. The residue was purified by flash column chromatography (SiO_2, chloroform / MeOH) to give 0.32 g of **3a** as gray solid: 91% yield. ^{1}H NMR (300 MHz, DMSO-d_6) δ 12.28 (s, 1H), 7.66 (d, *J* = 8.1 Hz, 1H), 7.46 (dd, *J* = 12.9, 8.3 Hz, 3H), 7.28 (dd, *J* = 22.1, 13.2 Hz, 4H), 7.11 (t, *J* = 7.6 Hz, 1H), 4.44 (s, 2H); m.p.: 191-193 ^{o}C; ESI-MS m/z: 348.4 $[M+H]^+$.

2.8.5 4-((6-(1-methyl-1H-pyrazol-4-yl)-[1,2,4]triazolo[3,4-b][1,3,4] thiadiazol-3-yl)methyl)phenol (3b)

Synthetic procedure was followed as described for compound **3a** to give **3b** as white solid: 85% yield. ^{1}H NMR (300 MHz, DMSO-d_6) δ 9.35 (s, 1H), 8.60 (s, 1H), 8.09 (s, 1H), 7.14 (d, *J* = 7.9 Hz, 2H), 6.75 – 6.66 (m, 2H), 4.26 (s, 2H), 3.93 (s, 3H); m.p.: 243 ^{o}C (decomposed); ESI-MS m/z: 313.1 $[M+H]^+$.

2.8.6 2-(4-((6-(1H-indol-2-yl)-[1,2,4]triazolo[3,4-b][1,3,4]thiadiazol-3-yl)methyl)phenoxy)-N,N-dimethylacetamide (11a)

3a (94 mg, 0.30 mmol) and 2-chloro-N,N-dimethylacetamide (44 mg, 0.36 mmol) was dissolved in 10 mL DMF and Cs_2CO_3 (0.32 g, 0.90 mmol) was added. The reaction mixture was stirred at 80 ^{o}C for 3 h. After cooled to room temperature, 30 mL water was added to the reaction mixture, and the mixture was extracted with ethyl acetate (30 mL x 2). The organic layer was washed with brine (50 mL), dried over Na_2SO_4 and concentrated. The residue was purified by flash column chromatography (SiO_2, chloroform / MeOH) to give 30 mg **11-1** as white solid: 23% yield. ^{1}H NMR (300 MHz, DMSO-d_6) δ 8.59 (s, 1H), 8.08 (d, *J* = 0.8 Hz, 1H), 7.25 (d, *J* = 8.3 Hz, 2H), 6.88 (d, *J* = 8.6 Hz, 2H), 4.79 (s, 2H), 4.32 (s, 2H), 3.92 (s, 3H), 3.57 (s, 4H), 3.43 (s, 4H); ^{13}C NMR (75 MHz, DMSO-d_6) δ 167.06, 158.76, 157.10, 138.12, 129.56 (2C), 127.84, 126.32, 125.01, 121.54, 120.59, 114.74 (2C), 112.36, 107.37, 65.79, 38.71, 35.55, 34.89, 29.53; m.p.: 225-227^{o}C; HRMS (ESI^+) *m/z* 455.1244 (455.1261 calcd for $C_{22}H_{20}N_6NaO_2S^+$, $[M+Na]^+$).

2.8.7 2-(4-((6-(1-methyl-1H-pyrazol-4-yl)-[1,2,4]triazolo[3,4-b][1,3,4] thiadiazol-3-yl)methyl)phenoxy) -1-morpholinoethan-1-one (11b)

Synthetic procedure was followed as described for compound **11-1** to give **11-2** as white solid: 40% yield. ^{1}H NMR (300 MHz, DMSO-d_6) δ 8.59 (s, 1H), 8.08 (d, *J* = 0.8 Hz, 1H), 7.25 (d, *J* = 8.3 Hz, 2H), 6.88 (d, *J* = 8.6 Hz, 2H), 4.79 (s, 2H), 4.32 (s, 2H), 3.92 (s, 3H), 3.51-3.61 (brs, 4H), 3.48-3.40 (s, 4H); ^{13}C NMR (75 MHz, DMSO-d_6) δ 165.96, 146.54, 137.66, 131.74, 129.61(2C), 127.93, 114.72(2C), 65.99(2C), 65.83, 44.73, 41.55, 29.50; m.p.: 164-166^{o}C; HRMS (ESI^+) *m/z* 462.1307 (462.1319 calcd for $C_{20}H_{21}N_7NaO_3S^+$, $[M+Na]^+$).

2.8.8 2-((7-methoxyquinolin-4-yl)oxy)acetic acid (4)

2-hydroxyacetic acid (5.00 g, 65 mmol) and KOH (6.00 g, 90 mmol) were mixed and heated to 170 ^{o}C. 4-chloro-7-methoxyquinoline (5.00 g, 26 mmol) dissolved in DMSO (20 mL) was added dropwise. The reaction mixture was stirred at 170 ^{o}C for 2.5 h. After cooled to room temperature, the solution was poured to 50 mL ice water. Saturated aqueous sodium carbonate solution was added to adjust the PH to 7-8. The mixture was filtered and dried in vacuo to give 3.70 g of **4** as brown solid: 62% yield. m.p.: 223-225^{o}C; HRMS (ESI^{+}) *m/z* 234.0763 (234.0761 calcd for $C_{12}H_{12}NO_4^{+}$, $[M+H]^{+}$).

2.8.9 4-amino-5-(((7-methoxyquinolin-4-yl)oxy)methyl) -4H-1,2,4-triazole-3-thiol (5)

4 (2.00 g, 8.60 mmol), hydrazinecarbothiohydrazide (1.00 g, 9.50 mmol) and CH_3SO_3H (1.80 g, 19 mmol) were dissolved in 3 mL sulfolane and 3 mL H_2O. The solution was stirred at 90 ^{o}C for 24 h. After cooled to room temperature, the mixture was poured into 30 mL ice water. Saturated aqueous sodium carbonate solution was added to adjust the PH to 7-8. The mixture was filtered and dried in vacuo to give 1.80 g of **5** as brown solid: 69% yield. HRMS (ESI^{+}) *m/z* 304.0864 (304.0863 calcd for $C_{13}H_{14}N_5O_2S^{+}$, $[M+H]^{+}$).

2.8.10 6-(1H-imidazol-2-yl)-3-(((7-methoxyquinolin-4-yl)oxy)methyl) -[1,2,4]triazolo[3,4-b][1,3,4] thiadiazole (11c)

5 (1.00 g, 3.30 mmol) and 1*H*-imidazole-2-carboxylic acid (0.40 g, 3.60 mmol) were dissolved in 3 mL sulfolane, and 12 mL $POCl_3$ was added. The reaction mixture was stirred at 85 ^{o}C for 18 h, then cooled to room temperature, poured into 200 mL ice water. NaOH was added to adjust the PH to 7. The mixture was filtered and dried in vacuo to give 92 mg of **11-3** as gray solid: 7% yield. ^{1}H NMR (300 MHz, DMSO-d_6) δ 13.83 (s, 1H), 8.77 (d, *J* = 5.4 Hz, 1H), 8.01 (d, *J* = 9.0 Hz, 1H), 7.52 (s, 1H), 7.36 (d, *J* = 2.5 Hz, 1H), 7.28 (d, *J* = 5.2 Hz, 2H), 7.18 (dd, *J* = 9.2, 2.6 Hz, 1H), 5.82 (s, 2H), 3.90 (s, 3H); m.p.: 205-207^{o}C; HRMS (ESI^{+}) *m/z* 380.0914 (380.0924 calcd for $C_{17}H_{14}N_7O_2S^{+}$, $[M+H]^{+}$).

2.8.11 3-(((7-methoxyquinolin-4-yl)oxy)methyl)-6-(thiazol-4-yl) -[1,2,4]triazolo[3,4-b][1,3,4] thiadiazole (11d)

Synthetic procedure was followed as described for compound **11-3** to give 92 mg of **11-4** as yellow solid: 7% yield. ^{1}H NMR (300 MHz, Chloroform-*d*) δ 8.94 (d, *J* = 2.0 Hz, 1H), 8.74 (d, *J* = 5.4 Hz, 1H), 8.29 – 8.07 (m, 2H), 7.41 (d, *J* = 2.6 Hz, 1H), 7.18 – 7.08 (m, 1H), 7.04 (d, *J* = 5.4 Hz, 1H), 5.78 (s, 2H), 3.94 (s, 3H); m.p.: 211-213^{o}C; HRMS (ESI^{+}) *m/z* 397.0528 (380.0536 calcd for $C_{17}H_{13}N_6O_2S_2^{+}$, $[M+H]^{+}$).

2.8.12 6-(1H-indol-2-yl)-3-(((7-methoxyquinolin-4-yl)oxy)methyl) -[1,2,4]triazolo[3,4-b][1,3,4] thiadiazole (11e)

Synthetic procedure was followed as described for compound **11-3** to give 113 mg of **11-5** as light brown solid: 8% yield. ^{1}H NMR (300 MHz, DMSO-d_6) δ 12.33 (s, 1H), 8.75 (d, *J* = 5.2 Hz, 1H), 8.02 (d, *J* = 9.2

Hz, 1H), 7.67 (d, *J* = 8.0 Hz, 1H), 7.52 – 7.23 (m, 5H), 7.20 – 7.00 (m, 2H), 5.83 (s, 2H), 3.89 (s, 3H); m.p.: 258-259^{o}C; HRMS (ESI^{+}) *m/z* 429.1125 (429.1128 calcd for $C_{22}H_{16}N_6O_2S^{+}$, $[M+H]^{+}$).

2.8.13 6-(1H-benzo[d]imidazol-5-yl)-3-(((7-methoxyquinolin-4-yl)oxy) methyl)-[1,2,4]triazolo[3,4-b] [1,3,4]thiadiazole (11f)

Synthetic procedure was followed as described for compound **11-3** to give 156 mg of **11-6** as gray solid: 11% yield. ^{1}H NMR (300 MHz, DMSO-d_6) δ 13.12 (s, 1H), 8.79 (d, *J* = 5.5 Hz, 1H), 8.44 (s, 1H), 8.20 (s, 1H), 8.10 (d, *J* = 9.1 Hz, 1H), 7.78 (s, 2H), 7.40 – 7.28 (m, 2H), 7.21 (dd, *J* = 9.3, 2.6 Hz, 1H), 5.93 (s, 2H), 3.91 (s, 3H); m.p.: 249-251^{o}C; HRMS (ESI^{+}) *m/z* 430.1069 (380.0536 calcd for $C_{21}H_{16}N_7O_2S^{+}$, $[M+H]^{+}$).

3. RESULTS AND DISCUSSIONS

3.1 Virtual Screening

3.1.1 Selection of Docking Programs and PDBs

Table 1. RMSD values between the crystal and docked conformations of inhibitors [a]

PDB ID	CDocker	Glide(XP) [b]	Gold	Surflex
2RFS	0.49	0.34	1.52	1.33
2WD1	0.85	0.28	0.63	0.73
2WGJ	2.70	0.64	0.72	0.48
2WKM	2.12	1.59	1.03	2.54
3A4P	0.42	0.15	1.93	0.79
3CCN	0.36	0.15	0.24	0.43
3CD8	0.50	0.18	0.64	0.20
3DKF	1.11	0.28	0.28	0.21
3DKG	0.51	0.31	0.42	0.42
3F66	0.75	0.86	0.72	0.32
3I5N	0.48	0.25	0.40	0.31
Mean (Å)	0.93	0.46	0.78	0.71
Std. dev. (Å)	0.74	0.42	0.50	0.66

0Å<RMSD<0.5Å	0.5Å<RMSD<1Å	1Å<RMSD<1.5Å	1.5Å<RMSD<2Å	2Å<RMSD

[a]A color code bar is given according to the RMSD range. [b]Glide extra docking precision mode.

Four docking programs, including CDocker, Glide, Gold, and Surflex, were evaluated by re-generating the ligand binding conformation in the 11 crystal structures. The resulted RMSDs (Table 1) between the docked and crystal conformations of inhibitors suggested Glide and Gold were preferred for molecular docking. Glide outperformed the other programs as only one RMSD value was over 1.00 Å (1.59Å for 2WKM) and the average RMSD value was also the smallest. Gold was also considered as a suitable program since the docking results were all in the acceptable range. Therefore, Gold and Glide were selected for the molecular docking in this study.

Only limited number of crystal structures of c-Met kinase domain in complex with type I inhibitors were available at the outset of this study. All these available crystal structures of c-Met kinase domain in complex with type I inhibitors were evaluated. Considering both of the cross-docking accuracy and structure types of the ligands, crystal structures (PDB ID: 2WD1 [11], 2WGJ [12], 2WKM [12], 3A4P [14], and 3QTI [27]) were selected for cascade molecular docking (supplementary Table S2).

3.1.2 Complex-based Pharmacophore Generation and Validation

Complex-based pharmacophore models were generated based on the eleven crystal structures (Table 1) with different type I c-Met inhibitors (supplementary Fig. S1) to cover comprehensive pharmacophoric features. Structural alignment and feature cluster (Fig. 2) yielded different groups of pharmacophoric features, including hydrogen bond donors (HBD), hydrogen bond acceptors (HBA), hydrophobic center (HC) and aromatic ring (RA). HBD was not selected because of its small number and dispersed spatial distribution. Two representative HBA features reflected hydrogen bond interactions between the ligand and Met1160, Asp1222 were kept. Two main HC groups were found. One was located in the adenine binding pocket, and the other was between the hydrophobic residues Met1211 and Tyr1230. RA cluster center covered right the spatial position of the second HC. It was found that most of the chemical groups represented by the second HC were aromatic. Therefore, RA was more feasible to represent the chemical groups required in this region. Considering these, the first HC and the RA were kept in the pharmacophore model. Thus, the initial pharmacophore model (Hypo1) consisted of two HBA, one HC and one RA (supplementary Fig. S2.A).

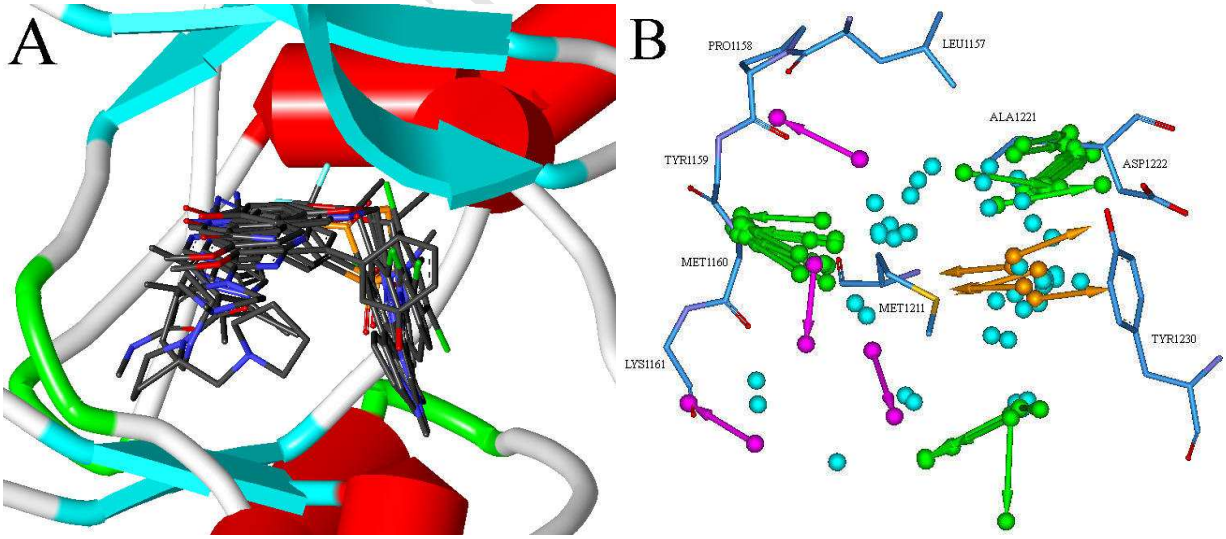


Fig. 2 (A) The 11 complex structures were well aligned in a common reference structure (PDB ID: 2WGJ). (B) The pharmacophoric features generated based on the aligned complex structures: HBA, green; HBD, magenta; HC, cyan; RA, orange.

To improve selectivity, the initial model Hypo1 was optimized by adding excluded volume spheres, which resembled sterically inaccessible regions within the binding site. Twelve excluded volume features were found and clustered in the ATP-binding site, whose spatial positions were respectively occupied by residues Lys1161, Met1160, Leu1157, Ala1108, Ile1084, Met1211, Leu1140, Gly1085, Tyr1230, Lys1110, Asn1209, and Asp1222. The resultant pharmacophore model was Hypo2 (supplementary Fig. S2.B). It was complemented with shape constraint based on the crystal binding conformation of Crizotinib to obtain a more restrict model, Hypo3 (supplementary Fig. S2.C).

It was suggested that *ligand pharmacophore mapping* module could generate low-energy conformations similar to bioactive conformations of active inhibitors in the binding pocket. With "*Best Fit*" selected, PF-02341066 was well mapped onto Hypo3. The RMSD value between the crystal and mapping conformation is 1.5689Å (supplementary Fig. S3) suggested acceptable predictive ability of this model for binding conformation.

Table 2. Statistics of pharmacophore model validation.

Parameter	Hypo1	Hypo2	Hypo3
Total molecules in database (D)	624	624	624
Total number of actives in database (A)	86	86	86
Total Hits (Ht)	394	287	12
Active Hits (Ha)	78	74	12
% Yield of actives [(Ha/Ht) × 100]	19.80	25.78	1
% Ratio of actives [(Ha/A) × 100]	90.70	86.05	13.95
Enrichment factor (EF) [(Ha × D)/(Ht × A)]	1.44	1.87	7.26
False Negatives [A - Ha]	8	12	0
False Positives [Ht - Ha]	316	213	0
Goodness of Hit Score (GH)[a]	0.15	0.24	0.78

[a] [(Ha/4HtA) (3A + Ht)) × (1 - ((Ht - Ha)/(D -A))]; GH score of >0.6 indicates a functioning model.

Enrichment factor (EF), Goodness of Hit Score (GH) and ROC scores were applied as the criteria to assess the performance of these three models by retrospective discovery of active c-Met inhibitors from inactive ones. Most of the active compounds could be identified by Hypo1 and Hypo2 (Table 2), but also many false positives. For Hypo3, only 12 compounds were retrieved, but all active ones, indicating strong reliability of this model in rejecting inactive compounds. Hypo3 was more statistically significant than Hypo1 and Hypo2 (Table 2). The GH score was 0.78 while the accepted value of GH must ≧ 0.5 for a screening model to be considered as well defined. ROC score was 0.807 (supplementary Fig. S4), indicating that in eight out of ten cases, one randomly selected active c-Met inhibitors is better ranked than an inactive one. Considering all of these, Hypo3 was the most suitable model for virtual screening.

3.1.3 Virtual Screening Workflow

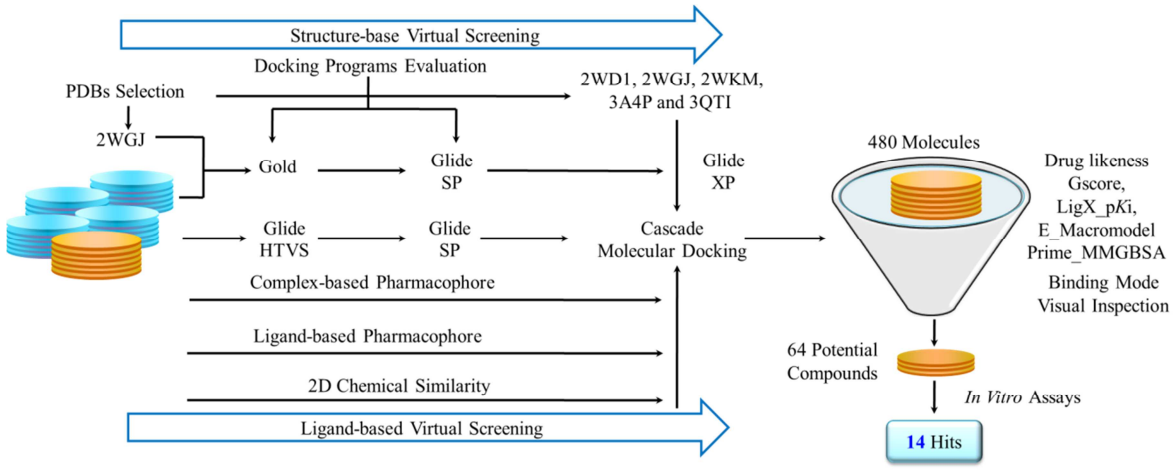


Fig. 3 The discovery process for novel c-Met inhibitors.

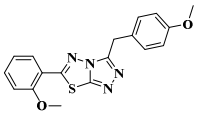
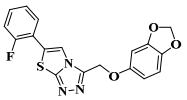
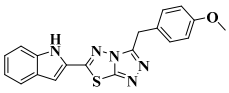
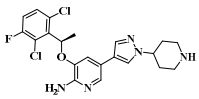
ChemDiv compound database was used for virtual screening (Fig. 3). The first parallel line was the cascade molecular docking using Gold and Glide. Four known c-Met inhibitors (IC_{50}: 2, 20, 100, and 640 nM) were used as positive control. Their lowest Gold Fitness 55.00 and highest Glide Gscore -7.3 were set as the threshold values to discard the compounds poorly scored in the Gold and Glide HTVS docking procedures. The first 10% of the Glide SP docking results were then put into the cascade docking using the selected PDBs. The second parallel line was performed by Glide independently. The first 10% ranked compounds of former mode were put into the next docking mode. The third parallel line was the complex-based pharmacophore model. The first 10% compounds mapped well on Hypo3 were put into cascade docking. Ligand-based pharmacophore model was used as another query for screening the compound database. The first 10% compounds mapped well on it were put into cascade docking. The other parallel line is 2D chemical similarity. Compounds with Tanimoto coefficient less than 0.5 were kept. After all the parallel lines, cascade molecular docking using the selected crystal structures was applied to all the retrieved compounds. The compounds with preferred binding pose in all the selected crystal structures were considered to be potential hits. Through visual inspection, 480 novel compounds were obtained. After filtered by further scoring (Fig. 3) and drug-likeness, the first 64 compounds of each structural type by different scoring methods were selected and purchased from vendors in sufficient purity and quantity.

3.2 *In Vitro* Biological Evaluation of Virtual Hits

After biological evaluation, fourteen hit compounds were found, with Crizotinib as the positive control. Among them, compounds **HL-9**, **HL-10**, and **HL-11** sharing similar scaffold got the IC_{50} values of 11.19, 3.31 and 0.19 μM (Table 3), respectively. The other hit compounds (data not shown) were in micromolar

range, which were in the process of further structural optimization [28]. However, the three hit compounds did not show significant inhibitory activity against the selected cancer cell lines (Table 3), indicating further structural optimizations were needed. Compound **HL-11** showed high selectivity against KDR and B-raf (supplementary Table S3). It was supposed that this specificity might be closely related with the 'U' shaped binding mode of type I inhibitors. But this hypothesis should be validated by experimental approaches like X-ray crystallography to disclose binding mode of **HL-11**.

Table 3. Structures and activities of the hit compounds.

Comp.	ChemDiv IDNUMBER	Structure	Kinase Inhibition	Tumor Cell lines Inhibition Rate (%) under 10μM				
			c-Met	A549	HCT116	MCF-7	PC3	SGC-7901
HL-9	D727-0159		11.19	10.42	8.2	35.7	57.74	15.53
HL-10	D727-0656		3.31	20.11	4.73	45.41	19.43	12.37
HL-11	D727-0806		0.19	6.08	0.64	14.94	55.1	12.26
Crizotinib	—		0.001053	0.1343 [a]	0.2536 [a]	0.045 [a]	ND [b]	0.3213 [a]

[a] IC_{50} values in μM; [b]not determine

3.3 X-ray Crystallographic study

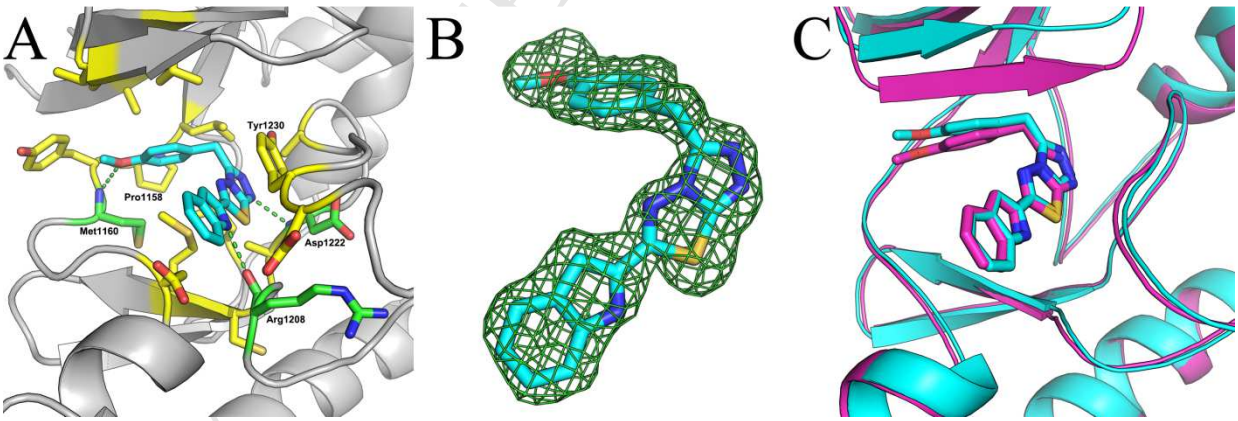


Fig. 4 A) Interactions of **HL-11** with the kinase domain of c-Met revealed by the crystal structure (PDB ID: 5YA5). The compound was shown by cyan sticks. Hydrogen bonds were shown as green dash line. The residues hydrogen bonding to **HL-11** were coloured in green while other surrounding residues were colored in yellow. B) A (F_o-F_c) difference electron-density map contoured at 3.0 σ for **HL-11** in its complex structure with c-Met kinase determined by the X-ray diffraction. C) The crystal-structure determined (cyan) and docked (magenta) conformations of **HL-11** were aligned in the binding site.

The crystal structure of c-Met kinase domain in complex with **HL-11** has been determined at 1.89 Å resolution (PDB ID: 5YA5; supplementary Table S1). It was found that **HL-11** utilized a 'U' shaped crystal

binding pose (Fig. 4A). Unlike Crizotinib, although it took such a conformation, the protein-ligand interactions mainly focused on the side of Arg1208, Asp1222 and Tyr1230. Hydrogen bonds were generated between **HL-11** and the two residues, Arg1208 and Asp1222. Hydrophobic and π-π stacking interactions between the scaffold of **HL-11** and the side-chain of Tyr1230 contributed greatly to the binding. The hydrogen bond between **HL-11** and the main-chain of Met1160 was considered to be critical for c-Met kinase inhibitors. The electrostatic density of **HL-11** (Fig. 4B) showed that it was possible for the methyl to rotate in different directions, but not disturb this hydrogen bond. Therefore, **HL-11** perfectly fitted into the binding cavity of c-Met kinase.

Importantly, the crystal binding conformation of **HL-11** turned out to be very similar to the computational prediction (RMSD = 0.4147Å; Fig. 4C), proving the accuracy of molecular docking. Based on this experimental validation, **HL-9** and **HL-10** were supposed to take similar binding mode of **HL-11** due to their similar structures and docking conformations. Molecular dynamics simulations were applied to find their differences for further structural optimization.

3.4 Molecular Dynamics Simulation

Molecular dynamics simulations were performed to elucidate the differences of protein-ligand interactions between compound **HL-10** and **HL-11**. Both of the simulations were performed for 10 ns (Fig. 5). For **HL-11**, the protein-ligand system kept equilibrium all through the process while it converged after 7.3 ns for **HL-10**. The predicted root mean square fluctuation (RMSF) values (supplementary Fig. S5) illustrated that the protein kept stable in the process of modeling, especially the conservative residues, like the hinge region, Asp1222 and Tyr1230. Both of the binding free energy values for **HL-11** and **HL-10** calculated with MM/GBSA and MM/PBSA methods illustrated that it was more stable for protein-ligand binding of **HL-11** than **HL-10**, consistent with their c-Met kinase inhibitory activity. The individual energy terms (Table 4) illustrated that van der waals energy contributions and polar desolvation energy were the major reasons for the distinct binding affinity of **HL-11** and **HL-10**. Hydrogen bond analysis found that both of these two compounds formed conserved hydrogen bonds with the amino acid residues Met1160 and Asp1222 (Table 5). The occupancies of the hydrogen bond with Met1160 were 98.51% and 92.66%, respectively. But the occupancies of the other hydrogen bond with Asp1222 were 32.18% and 74%, indicating that it was more stable for the hydrogen bond with **HL-11** than **HL-10**. This was supposed to be one of the reasons that **HL-10** was less potent than **HL-11**. Besides, this also illustrated that hydrogen bond between the ligand and Asp1222 was very important for their activity.

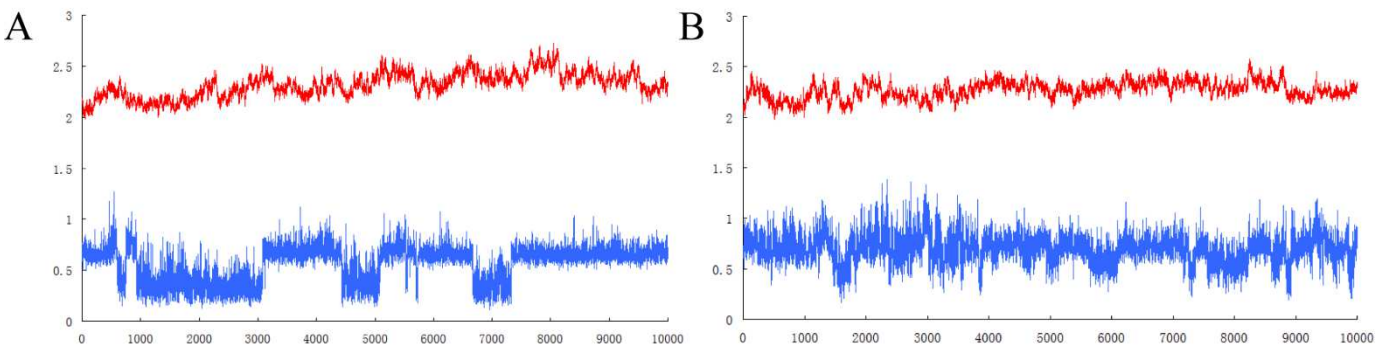


Fig. 5 RMSD of protein-ligand complex in 10 ns, red: backbone atoms of c-Met kinase domain; blue: heavy atoms of the ligand. A) **HL-10**. B) **HL-11**.

Table 4. Predicted binding free energies and individual energy terms of **HL-10** and **HL-11** in complex with c-Met kinase domain (kcal/mol^{-1}), calculated by MM/PBSA and MM/GBSA.

Contribution	**HL-10**	**HL-11**		
	Mean/(kcal/mol)	Mean/(kcal/mol)	Mean/(kcal/mol)	Std.
ΔE_{ele}	-6.50	2.67	-16.91	2.71
ΔE_{vdw}	-47.42	2.95	-49.19	2.26
ΔE_{intra}	0.00	0.00	0.00	0.00
ΔE_{gas}	-53.92	4.05	-66.10	2.93
ΔG_{np}	-6.00	0.14	-5.70	0.12
ΔG_{pb}	34.98	3.16	36.47	2.06
ΔG_{pbele}	28.48	3.45	19.57	2.38
$\Delta G_{pb,tot}$	-24.94	3.93	-35.33	2.62
ΔG_{gb}	22.46	1.97	27.75	1.96
ΔG_{gbele}	15.96	1.66	10.84	1.53
$\Delta G_{gb,tot}$	-37.46	3.00	-44.05	2.22

E_{ele}: electrostatic energy term; E_{vd}: van der waals energy term; E_{intra}: internal energy contributions generated because of conformation changes; $E_{gas} = E_{ele} + E_{vdw} + E_{intra}$; G_{np}: non-polar desolvation free energy term; G_{pb}: polar desolvation free energy calculated using MM/PBSA; $G_{pbele} = E_{ele} + G_{pb}$; $G_{pb,tot} = E_{gas} + G_{pb} + G_{np}$; G_{gb}: polar desolvation free energy calculated using MM/GBSA; $G_{gbele} = E_{ele} + G_{gb}$; $G_{gb,tot} = E_{gas} + G_{gb} + G_{np}$.

Table 5. Statistics of hydrogen bonds formation between compound **HL-11**, **HL-10** and the corresponding residues of c-Met.

Compd.	Hbond	Distance (Å)	Angle (°)	Occupied (%)
HL-10	O9@MOL	3.095 (±0.17)	2.966(±0.14)	23.77(±11.69)	98.51
	N14@MOL	3.263 (±0.15)	3.206 (± 0.17)	47.59 (±8.98)	32.18
HL-11	O21@MOL	2.966(±0.14)	3.095 (±0.17)	22.98 (±12.15)	92.66
	N7@MOL	3.206 (± 0.17)	3.263 (±0.15)	40.88 (±11.14)	74.00

Occupied (%): Percentage of snapshots with H-bond formation during MD simulations. The larger the occupied value, the more stable the hydrogen bonds are. Percentages lower than 50% mean the hydrogen bonds are not stable.

3.5 Structural Optimization

3.5.1 Molecular Design and Chemical Synthesis

Further structural optimization was performed to develop more potent c-Met kinase inhibitors based on the experimental and computational results (Fig. 6) [8, 9, 28, 29]. The hydrogen bonds with Met1160 and

Asp1222 were retained. For R^1, we proposed that it should be extended to the solvent accessible region to get hydrophilic interactions (compound **11a** and **11b**). Besides, we also supposed that R^1 should be bulky and hydrophobic groups to fully occupy the adenine binding pocket (compound **11c~11f**). Linker group was used to keep the 'U' shaped binding conformation of the entire ligand. We selected the commonly used linker groups to replace the original one (compound **11c~11f**). The hydrogen bond between R^2 and Arg1208 was not stable in the process of molecular dynamics simulation. Compounds with different R^2 groups were designed to investigate whether it was important or not. The hydrophobicity of R^2 was also investigated. These newly designed compounds were listed in Table 6. Their general synthetic routes for compound **11a**~**11f** were outlined in Scheme 1 and 2. The procedures and characterization were detailed in section 2.8, and the NMR spectra was in the supplementary material.

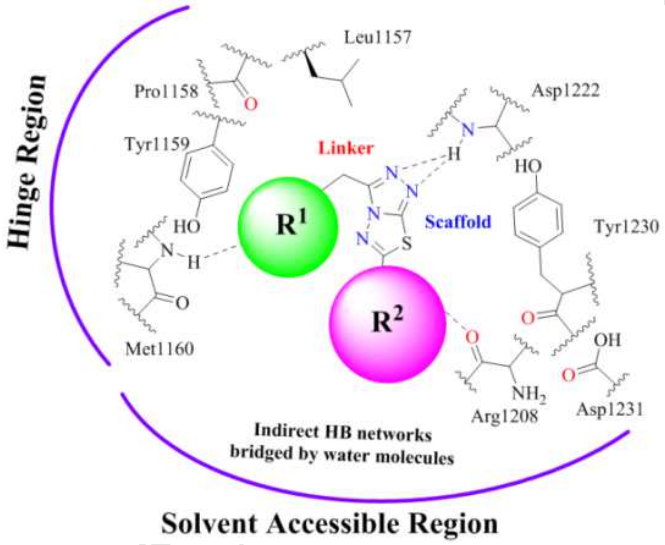
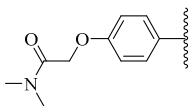
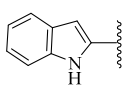
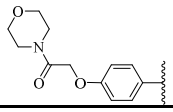
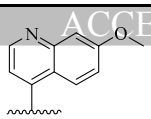
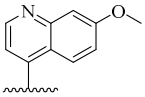
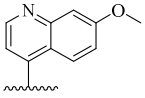
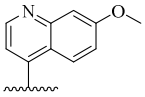


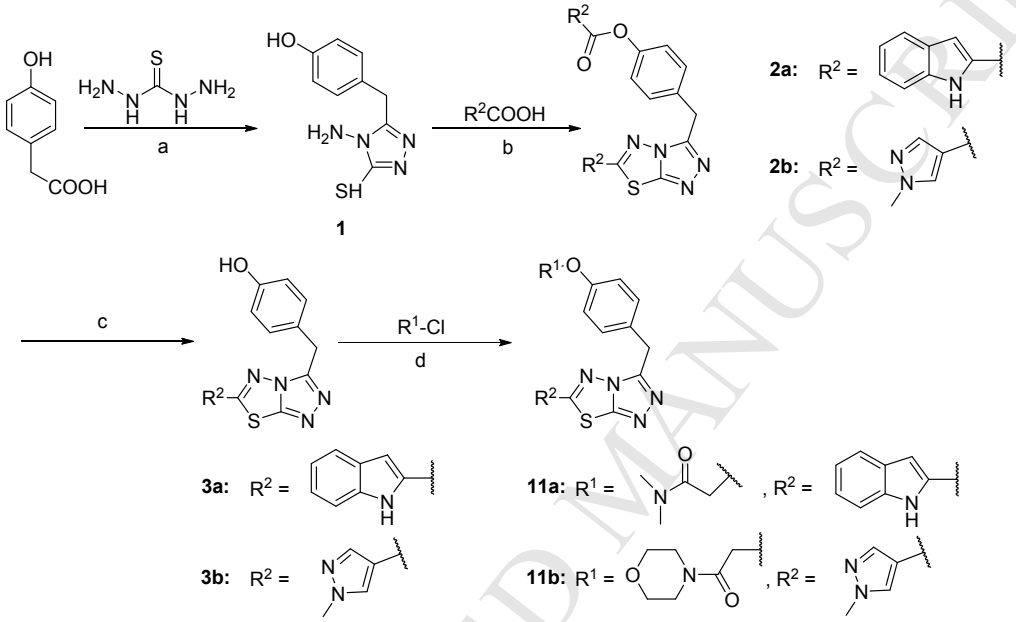
Fig. 6 Binding mode for structural optimization.

Docking conformations of all these new compounds in the binding site of c-Met kinase domain were shown in Fig. S6 in the supplementary material. They kept similar binding mode with the hit compound **HL-11** and the extending of R^1 and R^2 in the binding site was consistent with our design. Hydrogen bond interactions were found between all these new compound and Asp122, Met1160. Furthermore, compound **11a, 11c** and **11e** kept hydrogen bond interactions with the residue Arg1208 as expected.

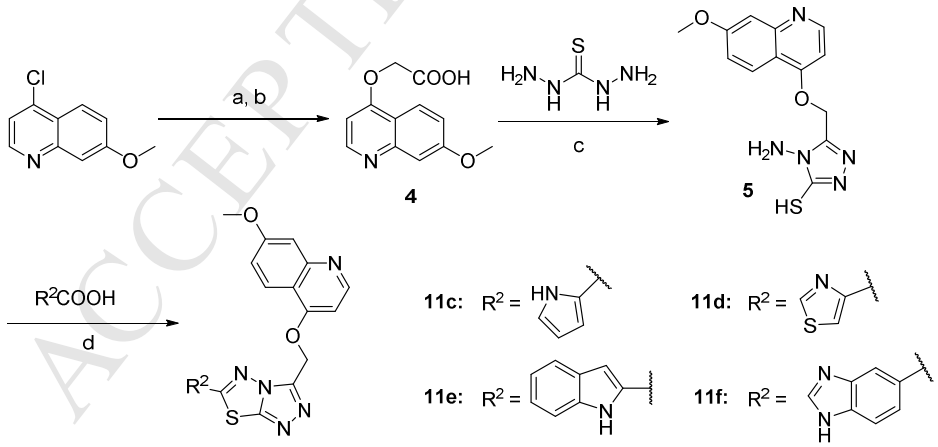
Table 6. The new designed compounds and their *in vitro* c-Met kinase inhibitory activity.

Comp.	R^1	Linker	R^2	c-Met (IC_{50}: μM)
11a		$-CH_2-$		6.51
11b		$-CH_2-$		5.49

11c		-OCH_2-	1.15
11d		-OCH_2-	0.154
11e		-OCH_2-	0.0793
11f		-OCH_2-	0.0139



Scheme 1. Reagents and conditions: a) CH_3SO_3H, sulfolane, H_2O, 90°C; b)$POCl_3$, sulfolane, 85 °C; c) NaOH, THF, rt; d) CS_2CO_3, DMF, 80°C.



Scheme 2. Reagents and conditions: a) 2-hydroxyacetic acid, KOH; b) 6M HCl, rt; c) CH_3SO_3H, sulfolane, H_2O, 90°C; d) $POCl_3$, sulfolane, 85°C.

3.5.2 *In Vitro* Biological Evaluation of 11a~11f

From Table 6, the optimization to extend R^1 into the solvent accessible region gave negative results. Compound **11a** and **11b** only got micromolar inhibitory activity. On the opposite, the bulky and hydrophobic R^1 group (compound **11c-11f**) gave the expected results. Comparison between the c-Met

kinase inhibitory activity of compound **11a** and **11e** validated the two hypotheses for R^1 optimization. For the linker, it didn't make significant differences because both of them could keep the 'U' shaped binding conformations. But they could be used for expanding the structural diversity. The different R^2 groups showed certain hydrophobicity and aromatic property was preferred to strengthen the protein-ligand interactions, including π-π interaction with Tyr1230. From the results of compound **11d** and **11e**, hydrogen bond between the ligand and Arg1208 should be favored for improving c-Met inhibitory activity. Furthermore, compound **11e** and **11f** with the best c-Met kinase inhibitory activity also got improved anti-proliferative activity (Table 7) compared with the starting point. Compound **11e** showed comparable inhibitory activity against NCI-H460, PC3 and Colo-205 to Crizotinib. But their anti-proliferative activity against EBC-1 and MKN45 were still needed to be improved.

Table 7. Anti-proliferative activity (IC_{50}: μM) of **11e** and **11f**.

Comp.	EBC-1	MKN45	NCI-H460	PC3	Colo-205
11e	0.506	1.171	2.055	2.970	2.242
11f	3.17	1.02	ND[a]	ND[a]	ND[a]
Crizotinib	0.013	0.022	2.244	9.787	2.449

[a] not determined

4. CONCLUSION

A new series of novel Type I c-Met inhibitors have been identified in this study with the multi workflows, including virtual screening, biological evaluations, X-ray crystallography, molecular dynamics simulation and structural optimization. The crystal binding conformation of the best hit compound **HL-11** was highly consistent with the computational result. Molecular dynamics simulations elucidated the main detailed differences between compound **HL-10** and **HL-11**. Based on this starting point, hypothesis for structural optimization was proposed and novel analogues of compound **HL-11** were designed. Compound **11d**~**11f** showed better *in vitro* c-Met inhibitory activity and anti-proliferative activity. Compound **11e** showed significantly improved anti-proliferative activity against several cancer cell lines. The experimental results illustrated that bulky hydrophobic R^1 groups that got hydrogen bonds with Met1160 were beneficial for better c-Met inhibitory activity. Aromatic and hydrophobic R^2 groups were preferred to strengthen the protein-ligand interactions. It was also favoured if additional stable hydrogen bond between the ligand and Arg1208. All these experimental findings proved the reliability and efficiency of our *in silico* methods.

Acknowledgments

This work was financially supported by the National Science Foundation of China (81473078 and 81703367), the scientific research funds for new teacher of China Pharmaceutical University

(3014070086) and the Fundamental Research Funds for the Central Universities (2632017PY12). This work was also supported by the "111 Project" from the Ministry of Education of China, the State Administration of Foreign Expert Affairs of China (No. 111-2-07).

Appendix A. Supplementary data

Supplementary data related to this article can be found at

References

[1] M.F. Di Renzo, R.P. Narsimhan, M. Olivero, S. Bretti, S. Giordano, E. Medico, P. Gaglia, P. Zara, P.M. Comoglio, Expression of the Met/HGF receptor in normal and neoplastic human tissues, Oncogene 6 (1991) 1997-2003.

[2] E. Gherardi, W. Birchmeier, C. Birchmeier, G. Vande Woude, Targeting MET in cancer: rationale and progress, Nat. Rev. Cancer 12 (2012) 89-103.

[3] C. Birchmeier, W. Birchmeier, E. Gherardi, G.F. Vande Woude, Met, metastasis, motility and more, Nat. Rev. Mol. Cell Biol. 4 (2003) 915-925.

[4] C. Boccaccio, P.M. Comoglio, Invasive growth: a MET-driven genetic programme for cancer and stem cells, Nat. Rev. Cancer 6 (2006) 637-645.

[5] P.C. Ma, M.S. Tretiakova, A.C. MacKinnon, N. Ramnath, C. Johnson, S. Dietrich, T. Seiwert, J.G. Christensen, R. Jagadeeswaran, T. Krausz, E.E. Vokes, A.N. Husain, R. Salgia, Expression and mutational analysis of MET in human solid cancers, Genes Chromosomes Cancer 47 (2008) 1025-1037.

[6] J.J. Cui, Targeting receptor tyrosine kinase MET in cancer: small molecule inhibitors and clinical progress, J. Med. Chem. 57 (2014) 4427-4453.

[7] S. Peters, A.A. Adjei, MET: a promising anticancer therapeutic target, Nat. Rev. Clin. Oncol. 9 (2012) 314-326.

[8] H.L. Yuan, J. Zhuang, S.H. Hu, H.C. Li, J.X. Xu, Y.N. Hu, X. Xiong, Y.D. Chen, T. Lu, Molecular modeling of exquisitely selective c-Met inhibitors through 3D-QSAR and molecular dynamics simulations, J. Chem. Inf. Model. 54 (2014) 2544-2554.

[9] W. Tai, T. Lu, H. Yuan, F. Wang, H. Liu, S. Lu, Y. Leng, W. Zhang, Y. Jiang, Y. Chen, Pharmacophore modeling and virtual screening studies to identify new c-Met inhibitors, J. Mol. Model. 18 (2012) 3087-3100.

[10] S.F. Bellon, P. Kaplan-Lefko, Y. Yang, Y. Zhang, J. Moriguchi, K. Rex, C.W. Johnson, P.E. Rose, A.M. Long, A.B. O'Connor, Y. Gu, A. Coxon, T.S. Kim, A. Tasker, T.L. Burgess, I. Dussault, c-Met inhibitors with novel binding mode show activity against several hereditary papillary renal cell carcinoma-related mutations, J. Biol. Chem. 283 (2008) 2675-2683.

[11] J. Porter, S. Lumb, R.J. Franklin, J.M. Gascon-Simorte, M. Calmiano, K.L. Riche, B. Lallemand, J. Keyaerts, H. Edwards, A. Maloney, J. Delgado, L. King, A. Foley, F. Lecomte, J. Reuberson, C. Meier, M. Batchelor, Discovery of 4-azaindoles as novel inhibitors of c-Met kinase, Bioorg. Med. Chem. Lett. 19 (2009) 2780-2784.

[12] J.J. Cui, M. Tran-Dube, H. Shen, M. Nambu, P.P. Kung, M. Pairish, L. Jia, J. Meng, L. Funk, I. Botrous, M. McTigue, N. Grodsky, K. Ryan, E. Padrique, G. Alton, S. Timofeevski, S. Yamazaki, Q. Li, H. Zou, J. Christensen, B. Mroczkowski, S. Bender, R.S. Kania, M.P. Edwards, Structure based drug design of crizotinib (PF-02341066), a potent and selective dual inhibitor of mesenchymal-epithelial transition factor (c-MET) kinase and anaplastic lymphoma kinase (ALK), J. Med. Chem. 54 (2011) 6342-6363.

[13] Afinogenov, II, [History of the pharmacy of the Main Hospital], Voen. Med. Zh. 318 (1997) 57-60.

[14] H. Nishii, T. Chiba, K. Morikami, T.A. Fukami, H. Sakamoto, K. Ko, H. Koyano, Discovery of 6-benzyloxyquinolines as c-Met selective kinase inhibitors, Bioorg. Med. Chem. Lett. 20 (2010) 1405-1409.

[15] B.K. Albrecht, J.C. Harmange, D. Bauer, L. Berry, C. Bode, A.A. Boezio, A. Chen, D. Choquette, I. Dussault, C. Fridrich, S. Hirai, D. Hoffman, J.F. Larrow, P. Kaplan-Lefko, J. Lin, J. Lohman, A.M. Long, J. Moriguchi, A.

O'Connor, M.H. Potashman, M. Reese, K. Rex, A. Siegmund, K. Shah, R. Shimanovich, S.K. Springer, Y. Teffera, Y. Yang, Y. Zhang, S.F. Bellon, Discovery and optimization of triazolopyridazines as potent and selective inhibitors of the c-Met kinase, J. Med. Chem. 51 (2008) 2879-2882.

[16] S.G. Buchanan, J. Hendle, P.S. Lee, C.R. Smith, P.Y. Bounaud, K.A. Jessen, C.M. Tang, N.H. Huser, J.D. Felce, K.J. Froning, M.C. Peterman, B.E. Aubol, S.F. Gessert, J.M. Sauder, K.D. Schwinn, M. Russell, I.A. Rooney, J. Adams, B.C. Leon, T.H. Do, J.M. Blaney, P.A. Sprengeler, D.A. Thompson, L. Smyth, L.A. Pelletier, S. Atwell, K. Holme, S.R. Wasserman, S. Emtage, S.K. Burley, S.H. Reich, SGX523 is an exquisitely selective, ATP-competitive inhibitor of the MET receptor tyrosine kinase with antitumor activity in vivo, Mol. Cancer Ther. 8 (2009) 3181-3190.

[17] J. Porter, S. Lumb, F. Lecomte, J. Reuberson, A. Foley, M. Calmiano, K. le Riche, H. Edwards, J. Delgado, R.J. Franklin, J.M. Gascon-Simorte, A. Maloney, C. Meier, M. Batchelor, Discovery of a novel series of quinoxalines as inhibitors of c-Met kinase, Bioorg. Med. Chem. Lett. 19 (2009) 397-400.

[18] A.A. Boezio, L. Berry, B.K. Albrecht, D. Bauer, S.F. Bellon, C. Bode, A. Chen, D. Choquette, I. Dussault, M. Fang, S. Hirai, P. Kaplan-Lefko, J.F. Larrow, M.H. Lin, J. Lohman, M.H. Potashman, Y. Qu, K. Rex, M. Santostefano, K. Shah, R. Shimanovich, S.K. Springer, Y. Teffera, Y. Yang, Y. Zhang, J.C. Harmange, Discovery and optimization of potent and selective triazolopyridazine series of c-Met inhibitors, Bioorg. Med. Chem. Lett. 19 (2009) 6307-6312.

[19] W. Wang, A. Marimuthu, J. Tsai, A. Kumar, H.I. Krupka, C. Zhang, B. Powell, Y. Suzuki, H. Nguyen, M. Tabrizizad, C. Luu, B.L. West, Structural characterization of autoinhibited c-Met kinase produced by coexpression in bacteria with phosphatase, Proc. Natl. Acad. Sci. U S A 103 (2006) 3563-3568.

[20] M.A. Seeliger, M. Young, M.N. Henderson, P. Pellicena, D.S. King, A.M. Falick, J. Kuriyan, High yield bacterial expression of active c-Abl and c-Src tyrosine kinases, Protein Sci. 14 (2005) 3135-3139.

[21] W. Kabsch, Xds, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 125-132.

[22] K. Wu, J. Ai, Q. Liu, T. Chen, A. Zhao, X. Peng, Y. Wang, Y. Ji, Q. Yao, Y. Xu, M. Geng, A. Zhang, Multisubstituted quinoxalines and pyrido[2,3-d]pyrimidines: synthesis and SAR study as tyrosine kinase c-Met inhibitors, Bioorg. Med. Chem. Lett. 22 (2012) 6368-6372.

[23] P.D. Adams, P.V. Afonine, G. Bunkoczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, PHENIX: a comprehensive Python-based system for macromolecular structure solution, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 213-221.

[24] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of Coot, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 486-501.

[25] A. Onufriev, D. Bashford, D.A. Case, Exploring protein native states and large-scale conformational changes with a modified generalized born model, Proteins 55 (2004) 383-394.

[26] P.A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D.A. Case, T.E. Cheatham, 3rd, Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models, Acc. Chem. Res. 33 (2000) 889-897.

[27] R. Tiedt, E. Degenkolbe, P. Furet, B.A. Appleton, S. Wagner, J. Schoepfer, E. Buck, D.A. Ruddy, J.E. Monahan, M.D. Jones, J. Blank, D. Haasen, P. Drueckes, M. Wartmann, C. McCarthy, W.R. Sellers, F. Hofmann, A drug resistance screen using a selective MET inhibitor reveals a spectrum of mutations that partially overlap with activating mutations found in cancer patients, Cancer Res. 71 (2011) 5255-5264.

[28] L. Zhang, B. Zhang, J. Zhao, Y. Zhi, L. Wang, T. Lu, Y. Chen, Structure-based design, synthesis, and evaluation of 4,5,6,7-tetrahydro-1H-pyrazolo[4,3-c]pyridine derivatives as novel c-Met inhibitors, Eur. J. Med. Chem. 138 (2017) 942-951.

[29] H.L. Yuan, T. Lu, T. Ran, H.C. Liu, S. Lu, W.T. Tai, Y. Leng, W.W. Zhang, J. Wang, Y.D. Chen, Novel strategy for three-dimensional fragment-based lead discovery, J. Chem. Inf. Model. 51 (2011) 959-974.

HIGHLIGHTS

- The best hit compound **HL-11** identified with the reliable virtual screening workflow exhibited potent c-Met kinase inhibitory activity.
- Comparison of the hit compounds with different c-Met inhibitory activity by molecular dynamics simulations suggested the key protein-ligand interactions for structural optimization.
- The experimentally resolved crystal binding mode of compound **HL-11** was highly consistent with the computational prediction.
- Structural optimization of compound **HL-11** resulted in **11e** with better c-Met kinase inhibitory activity and improved anti-proliferative activity.