Journal of Medicinal Chemistry

Article

Discovery of an Orally Selective Inhibitor of Signal Transducer and Activator of Transcription 3 Using Advanced Multiple Ligand Simultaneous Docking

Wenying Yu, Chenglong Li, Wenda Zhang, Yuanzheng Xia, Shanshan Li, Jiayuh Lin, Keqin Yu, Mu Liu, Lei Yang, Jian-Guang Luo, Yijun Chen, Hongbin Sun, and Ling-Yi Kong

J. Med. Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.jmedchem.6b01489 • Publication Date (Web): 28 Feb 2017

Just Accepted

ACS Publications

Journal of Medicinal Chemistry is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Discovery of an Orally Selective Inhibitor of Signal Transducer and Activator of Transcription 3 Using Advanced Multiple Ligand Simultaneous Docking

Wenying Yu,[†‡] Chenglong Li,[§] Wenda Zhang,[‡] Yuanzheng Xia,[‡] Shanshan Li,[‡] Jia-yuh Lin[⊥], Keqin Yu[#], Mu Liu,[†‡] Lei Yang,[‡] Jianguang Luo,[‡] Yijun Chen,[∥] Hongbin Sun[@] and Lingyi Kong[*,‡]*

[†]State Key Laboratory of Natural Medicines, [‡]Department of Natural Medicinal Chemistry, [∥]Laboratory of Chemical Biology, [@]Center for Drug Discovery, China Pharmaceutical University, 24 Tong Jia Xiang, Nanjing 210009, China.

[§]Division of Medicinal Chemistry and Pharmacognosy, College of Pharmacy, The Ohio State University, Columbus, Ohio 43210, United States

[⊥]Department of Biochemistry and Molecular Biology, School of Medicine, University of Maryland, Baltimore, Maryland 21201, United States

[#]Water supply and Dramage, Nanchang Hangkong University, 696 Fenghe Ave. S., Nanchang 330046, China.

*Corresponding Author: Phone: +86-25-8327-1405; Fax: +86-25-8327-1405; E-mail: lykong@cpu.edu.cn

ABSTRACT. Targeting signal transducer and activator of transcription 3 (STAT3) is a potential anticancer strategy. However, STAT3 inhibitors with good selectivity and bioavailability are rare. The aim of this study was to discover selective direct STAT3 inhibitors with good drug-likeness. By advanced multiple ligand simultaneous docking (AMLSD) method, compound **9** was designed as an orally-bioavailable STAT3 inhibitor, which presented superior druggability and selectivity to other representative STAT3 inhibitors. **9** directly and selectively inhibited the pY705 site of STAT3 with an affinity (K_i) of 440 nM. The IC_{50} of **9** for breast cancer cell MDA-MB-231 was 184-fold lower than its IC_{50} for normal breast epithelial cell MCF-10A. **9** *in vivo* induced significant antitumor responses better than Gefitinib and its therapeutic index should be over 100, indicating good safety of **9**.

INTRODUCTION. Constitutive activated Signal Transducers and Activators of Transcription (STAT) 3 protein is frequently detected in many malignancies,[1-3] representing one of the most promising anticancer targets.[4, 5] The activation of JAK-STAT signaling pathway can be initiated by the binding of cytokines, hormones, and growth factors to the cell surface receptors. Activated receptors phosphorylate JAK kinases which then phosphorylate STAT3 protein. Phosphorylated STAT3 (p-STAT3) proteins form dimers which then translocate into nuclear and bind to DNA, causing cell replication and cancers.[6, 7] p-STAT3 can be dimerized by forming the salt bridges between the R609 residue in the SH2 domain and the pY705 residue of the native peptide. In the STAT3 SH2 domain, there are two "hot spots", pY705 (also named

pY+0) site and a nearby pY+X site.[8] The pY705 site is related to the disruption of STAT3 phosphorylation and dimerization of inhibitors, while the nearby pY+X site is associated to the selectivity of STAT3 inhibitors. Therefore, compounds targeting the pY705 and pY+X sites of STAT3 can selectively inhibit both STAT3 phosphorylation and dimerization. In addition, phosphorylation of Serine727 (pS727) residue of STAT3 enhances transcriptional activation and acetylation of Lys685 (Ac-K685) residue is crucial for methylation of tumor-suppressor gene promoters.[9, 10] Thus, proximal sites, pS727 and Ac-K685 sites, are also functional in the tumor cell proliferation.

STAT family includes seven members, among which, STAT3 and STAT5 are oncoproteins, but STAT1 plays opposite roles of STAT3 in carcinomas by triggering anti-proliferative effect in tumorigenisis.[3] STAT5 plays a critical role in lobuloalveolar proliferation, whereas STAT3 regulates lobuloalveolar apoptosis during pregnancy.[11] In addition, STAT3 is structurally most similar to STAT1 and STAT5. Hence, designing compounds with highly selective inhibitory activity of STAT3 over STAT1 and STAT5 is difficult and critical for precision anti-cancer therapy.

Strategies directly deactivating STAT3 mainly focus on the disruption of STAT3 dimerization and phosphorylation. Several series of direct STAT3 inhibitors have been discovered *via* both computational and experimental methods, represented by phosphor-peptide mimics,[12, 13] non-peptide small molecule inhibitors,[7, 8, 14-16] metal-based inhibitors,[6] and natural-product like inhibitors.[17, 18] Some of them have entered clinical trials.[19-23] However, there are none marketed STAT3 inhibitors, due to poor

selectivity, undesirable PK/PD properties or other reasons. Therefore, the improvement of selectivity and druggability will substantially advance the discovery of STAT3 inhibitors as precision anticancer agents.

Computer-aided drug design is widely utilized for the efficient identification and optimization of lead compounds.[24-27] Aiming to design druggable STAT3 inhibitors, multiple ligand simultaneous docking method was advanced in two aspects, *i.e.* the fragment library design and linker design (Figure 1). Using AMLSD, we designed compounds directly inhibiting both phosphorylation and dimerization of STAT3 protein offering an orally-available potent and selective STAT3 inhibitor, **9**, which presented superior druggability or selectivity to other representative non-peptide small STAT3 inhibitors (Figure 2 and Table 1).[7, 8, 14, 15, 19, 28, 29] For example, **9** has better druggability than **14**,[7] which violates the Lipinski's rules.

Figure 1, 2 & Table 1

RESULTS.

Design by AMLSD method. 1) Fragment libraries design. Drug scaffolds (MW<300) were categorized into six fragment libraries according to their physical properties, including polarity, aromaticity, bulkiness, charge, and acidity/alkalinity. **2)** Screening of fragment combinations (FCs). Naphthalene-5,8-dione-1-sulfonamide (fragment **1**) was paired with each fragment from the polar fragment libraries forming 963 FCs, which were simultaneously docked to the crystal structure of STAT3 SH2 domain (PDB: 1BG1). The resultant multi-fragment virtual screening was ranked by their

binding energies (Table S1, Supporting Information (SI)), and the binding modes of the 25 top ranked FCs were shown in Figure 3. **3)** Linker design. Possible linear linkers for the selected FCs were designed by the measured distances. Since the measured distances between two fragments were in the range of 3~5 Å in Table S1, the length of linkers should be within the range of 3~5 Å, thus all the proper linear linkers were listed in Figure 4. Suitable linkers for FCs were then designed based on two criteria, the feasibility of synthesis and the superposition of the docking modes between the linked molecules and the separate FCs. Finally, linker ***a1*** could connect two fragments through a C-N coupling reaction while the linked compounds could reposition to their FCs, thus we chose ***a1*** as an ideal linker to connect the FCs. **4)** Synergistic hit selection. FCs were linked and docked to the STAT3 protein. Their docking modes were repositioned to that of the pre-linked FCs. If linked compound could reposition to its FCs, what's more, the original contacts between FCs and protein could be maintained, and even more good contacts were formed after linkage, then we consider the FCs were synergistically linked. Finally, ten out of 25 compounds were chosen with synergistic effects (Figure 5 & Table S2, (SI)), for example, the docking modes of the linked compound **9** could be repositioned to the docking modes of its FCs (No.175) (Figure 6). Moreover, additional intermolecular hydrogen bonds were formed between the linker of **9** and Ser636 residue which enhanced the synergistic effects. **5)** ADME properties. Drug-like parameters including the ADME and toxicity properties of the hits were computed, predicating all compounds with decent drug-like properties.

Figure 3-6

In order to investigate the *in silico* selectivity of the 10 compounds over STAT1/3/5. Compounds were also docked to the crystal structures of STAT1 (PDB: 1BF5) and STAT5 proteins (PDB: 1Y1U).[30] Compound **1** was eliminated since it was also docked to STAT1 SH2 domain with free energy, -7.8 kcal/mol. Compounds **3**, **4**, **6** and **8** were excluded since they were weakly bound to STAT5 SH2 domain. Limited by the availability of the starting materials, compound **10** was not synthesized. Finally, compounds **2**, **5**, **7** and **9** were synthesized for further biological tests.

Chemistry. Synthesis of Designed Compounds. Fragment **1** and aromatic amines were catalyzed by $Cu(OAc)_2 \cdot H_2O$ in an acidic environment. Reaction system was heated to 65℃ and terminated when **1** was consumed completely. The atom numbering of **1** is shown in Scheme 1. The temperature of reaction was a key reason to cause the cyclized by-products since the sulfonamide group can attack the nearby 1-ketone and deplete a water molecule to form the cyclized products. A higher temperature can cause a lower yield of the target compounds. For the synthesis of compound **9**, aiming to selectively react to the aromatic amine group of the reactant, the piperazine amine was protected by the Cbz group initially and then deprotected by TFA.

Scheme 1

Cell Assays. In a preliminary *in vitro* evaluation of the designed compounds, **9** exhibited best potency than other compounds against a human basal-like breast cancer cell line, MDA-MB-231 (Figure 7A). However, its IC_{50} for normal breast epithelial cell

MCF-10A was 128.9μM, 184-fold higher than that for MDA-MB-231, 0.7μM. So **9** was highly selectively targeting the cancer cells. **14** was selected as the positive control since it is a representative orally available STAT3 inhibitor. Compound **9** was more potent than **14** to kill multiple types of cancer cells, *i.e.* UW426, UW288-1, BKPC3, MDA-MB-231 and U2OS cell lines (Figure 7B).

Figure 7

Compound 9 is a Direct STAT3 Inhibitor. 1) 9 was predicated as a direct STAT3 SH2 domain inhibitor by both docking and molecular dynamics (MD) simulations. The intermolecular interactions were revealed between compound **9** and STAT3. Based on the analysis of the simulations, the sulfamine group of fragment **1** formed hydrogen bonds with residues Arg609, Ser611, Ser612, and Ser613 in the pY705 site. In detail, the 8-ketone group of fragment **1** as a major pharmacophore interacted with residues Arg609 and Lys591. The linker amine group contacted with residues Ser636 and Glu638 in the loop. The piperazinyl fragment formed ionic bond with residue Glu594 and formed hydrogen bond with Gln635 in pY+X site (Figure 8). The RMSDs of the snapshots during the MD simulaltions also reconfirmed the stable binding between compound **9** and STAT3 (Figure S2, (SI)).

Figure 8

2) The direct binding between **9** and STAT3 protein were demonstrated by both Fluorescent Polarization (FP)-based competitive binding assay and Microscale Thermophoresis (MST) assay. **9** was proven to preferentially bind to the pY705 site of

STAT3, competitively disrupting the binding of the native high-affinity, fluorescence-labeled phosphorylated peptide, 5-FAM-SpYLPQTV, with a K_i value of 0.44 μM (Figure 9A). While, there was no direct binding between compound **9** and STAT1/5 proteins. Thus, **9** was determined as a selective STAT3 inhibitor by FP assays.

MST is a recently broadly used technology for the interaction analysis of biomolecules, which was used to confirm that **9** directly bound to the STAT3 protein with a K_i value of 2.42 μM±0.26 μM. The molar ratio of ligand/protein was determined as 1: 1 (Figure 9B).

Figure 9

Selectivity of Compound 9. 1) 9 *in silico* selectively docked to the pY site of STAT3 protein, distant from the pY sites of STAT1/5 proteins (Figure 10). **2) 9** selectively inhibited p-STAT3 at pY705 site without altering at S727 and K685 sites. The level of p-STAT3 (Y705) was reduced in a dose-dependent manner (Figure 11A), while the levels of acetylated STAT3 (K685) and p-STAT3 (S727) were not dose-dependently influenced by **9** and the expression of total STAT3 was not changed, indicating that **9** selectively decreased the level of pY705-STAT3 which was not due to a constitutional decrease of total STAT3 expression (Figure 11B). **3)** In western blot (WB) assays, **9** selectively inhibited p-STAT3 (Y705) (Figure 11A) and IL-6 induced phosphorylation of STAT3 (Figure 12A), without any inhibition on p-STAT1 (Y701) (Figure 11B) and IFN-γ induced phosphorylation of STAT1 (Figure 12B). **4)** By electrophoretic mobility shift assay (EMSA), **9** *in vitro* preferentially inhibited STAT3

homodimer DNA-binding activity in a time-dependent manner, over STAT5-STAT5 DNA-binding activity, indicating **9** was a selective STAT3 inhibitor. **9** weakly decreased the level of STAT1 dimer DNA-binding activity, while no observable inhibition by compound **9** of the IFN-γ induced STAT1-STAT1 DNA-binding (Figure 12C). Quantitative real-time RT-PCR showed that compound **9** decreased the expression of STAT3-specific oncogene, MMP9, while the expression of STAT1-specific gene, CDK2, treated with **9**, was not significantly influenced (Figure 12D). **5)** The stimulation of STAT3 upstream targets, such as EGFR, can not only induce the level of p-STAT3, but also induce the levels of p-AKT and p-ERK1/2 kinases. To investigate whether compound **9** is selective to the JAK-STAT signaling pathway, we also evaluated the levels of p-AKT and p-ERK1/2 kinases in bypass pathways. The phosphorylation levels of both kinases were not inhibited by **9** either (Figure 11C & Figure S8, (SI)). **9** showed better selectivity than **14** to p-AKT, since the level of p-AKT was decreased by **14** while it was not affected by **9** in MDA-MB-231 cancer cell line (Figure 11C).

In all, compound **9** selectively inhibited pY705-STAT3 protein.

Figures 10-12

Compound 9 Induced the Level of SHP-1. Interestingly, **9** was found to induce the expression of STAT3 phosphatase, SHP-1. Mechanism is unknown yet, however, the effect may partially contribute to the low IC_{50} values of **9** (Figure 11B&D).

Compound 9 Induced Apoptosis in Malignant Cancers. The inhibition of STAT3 phosphorylation by **9** seemed to be consistent with the induction of apoptosis as evidenced by the increase of cleaved caspase-3 in the MDA-MB-231 cancer cell line (Figure 11C).

Compound 9 Inhibited the Nuclear Translocation of p-STAT3 to Nucleus. The IL-6 induced level of p-STAT3 in the cytoplasm and nucleus treated with **9** was lower than that treated with **14** by immunofluorescence staining, suggesting **9** blocked STAT3 nuclear translocation with a better effect than **14** (Figure 13).

Figure 13

***In vivo* Study of Compound 9.** To investigate the safety profile of compound **9**, the acute toxicity of **9** was tested in ICR mice at doses of 200, 400, 600, 800 and 1000 mg/kg (*ip*, n = 10 per group). All the mice were vital and lively even treated at the highest dose (1000 mg/kg), thus the median lethal dose (LD_{50}) was unable to be obtained. Since 70% tumor weight was shrank at 10 mg/kg **9** (Figure S9), the therapeutic index of **9** should be over 100 (1000/10=100), indicating good safety of **9**.

1) 9 orally inhibited the growth of cancer xenografts with better therapeutic effects than first-line anticancer drug, Gefitinib, and with less toxicity than doxorubicin. The nude mice tumor xenografts were injected subcutaneously with human breast cancer MDA-MB-231 cells. **9** orally shrank 53% tumor weight at 2.5 mg/kg, 60% tumor weight at 5 mg/kg and 70% tumor weight at 10 mg/kg, meanwhile for Gefitinib, dosage at 40 mg/kg were needed to achieve similar therapeutic effects as **9** at 10 mg/kg,

indicating **9** was more active than the first-line anticancer Drug, Gefitinib. (Figure 14 and Figure S9). The body weight, post-treatment of either Gefitinib or **9**, increased normally (Figure 14D and Figure S9D) and no damage of the major organs (heart, liver, spleen, lung and kidney) was detected by H&E staining (Figure S9B).

Figure 14

The IC_{50} of **9** was firstly evaluated as 0.42 μM in S180 cancer cell line. In the mouse sarcoma tumor xenografts, 73% tumor weight was significantly ($P < 0.001$) shrank by **9** after a 21 days therapy (Figure 15A, C&E), showing much less side effect by **9** at 5 mg/kg than by Doxorubicin at 1 mg/kg. Dramatic loss of body weight was observed from the 13^{th} day treated by Doxorubicin; whereas for **9**, the increase of body weight was normal (Figure 15D&F). Additionally, the size of liver, spleen and kidney organs treated by Doxorubicin was notably smaller than that treated by **9** (Figure 15B).

Figure 15

2) 9 *in vivo* disabled STAT3 activity. The *in vivo* inhibitory activity of **9** against the overexpression level of p-STAT3 was studied by both immunohistochemical (IHC) analysis and WB. From the IHC results, **9** showed dose-dependent inhibitory effects on p-STAT3 (Figure 16). The results were confirmed by WB results, **9** dose-dependently inhibited the level of p-STAT3 while showing little influence on the level of total-STAT3 (Figure 17).

Figure 16 & 17

3) The oral bioavailability of compound **9** was measured as 44.7%±11.9%, comparable to Gefitinib, 47.4%±0.9% by PK study. The pharmacokinetic parameters of **9** were well comparable to Gefitinib too (Table 2). The change of plasma concentrations in rats through 2 min to 24 h was shown in Figure 18, treated by **9** (*i.v.* and *p.o.*). These results indicated that **9** as an orally bioavailable candidate. The method was fully validated according to the FDA guidelines for selectivity, linearity, lower limits of quantification (LLOQ), accuracy, precision, recovery, matrix effect and stability (Tables S8-S11).

Figure 18 & Table 2

PAINS analysis. All designed compounds were filtered by the online filter, http://zinc15.docking.org/patterns/home/, for Pan Assay Interference Compounds (PAINS). Quinone group was detected, which is a common structure in natural products. It is also a substructure for endogenous molecules, such as Vitamin K, and some approved drugs, such as Mitomycin, Menadione and Idebenone. To test the potential reactivity of the α,β-unsaturated carbonyl group as Michael acceptor site in the quinone group, compound **9** and cysteamine were mixed in DMSO or water solutions at 37℃ overnight, no reaction was found, monitored by TLC (see SI), indicating that **9** wasn't very reactive.

Experimental results proved good activity, selectivity and safety of compound **9**. The STAT3 siRNA transfection experiment demonstrated that the major antitumor target for compound **9** was STAT3 and compound **9** is selective to STAT3 protein (see SI).

For the cell based assays, both normal breast epithelial cell line, MCF-10A, and breast cancer cell line, MDA-MB-231, were treated with compound **9**. The IC_{50} for normal cells, 128.9 μM, was 184-fold higher than that for cancer cells, 0.7 μM, indicating compound **9** is selectively targeting the tumor cells with good safety. For the acute toxicity study, mice were vital and lively, orally treated by even 1 g/kg compound **9**, indicating its therapeutic index should be over 100. The growth of the body weight of mice was also not affected by compound **9**, however the body weight of mice was significantly decreased by Doxorubicin. No damage of the major organs (heart, liver, spleen, lung and kidney) was detected treated by **9** by H&E staining (Figure 14B). The size of liver, spleen and kidney organs treated by Doxorubicin was notably smaller than that treated by **9** (Figure 15B) too. All the *in vivo* results indicated that compound **9** had better safety than the first-line anticancer drug Doxorubicin. In addition, for the protein based assays, compound **9** has been proved as the direct inhibitor of STAT3 protein. Therefore, both the *in vitro* and *in vivo* results pointed to compound **9** as a nonpromiscuous selective STAT3 inhibitor.

DISCUSSION.

By AMLSD method, multiple fragments are simultaneously docked to one target, which can cover the chemical spaces of binding sites better than one ligand hopping. MLSD was advanced in fragment library design and evolving process. The fragment library design is crucial for a successful drug design. In our study, fragments were derived from the drug scaffolds aiming to design more druggable STAT3 inhibitors.

Old drugs might "repurpose" for new uses.[27] Building blocks of the clinical drugs share many interesting properties, thus there is a higher chance to find drug candidates by designing compounds derived from drug scaffolds.[27] In addition, AMLSD determines linkers reminiscent as the BRIGE program in LUDI.[31] For example, different xylenes were determined by distance measurement in our case. Since the intermolecular interactions between fragments have been considered by AMLSD, there is a higher chance that the linked compounds could be repositioned to their fragment combinations. For previous application of MLSD method, fragment library was designed according to the similarity of fragments to the scaffolds of known inhibitors. Then the selected fragments are grouped as FCs, resulting a very limited number of FCs before docking. The number of final linked compounds are small and determined by the variety of known inhibitors. If drug targets were lack of known inhibitors, then it would be hard to apply MLSD method to design compounds. AMLSD can overcome the above problems. The selection of fragments is no longer relied on the scaffolds of known inhibitors. Thus, AMLSD is more useful to design candidates for a target not so well-studied.

Compound **9** was found to selectively bind to the p-STAT3(705) protein rather than STAT1/5 proteins, p-AKT and p-ERK kinases in bypass pathways, and other proximal sites. **9** also selectively targeted the MDA-MB-231 cancer cells by majorly disabling p-STAT3 level. The IC_{50} of **9** for normal cells was 184-fold higher than that for breast cancer cells. For the animal study, the growth of the body weight of mice was not

affected by compound **9**, which had much less toxicity than the first-line anticancer drug Doxorubicin. The therapeutic index of **9** should be over 100. In protein, cell and animal levels, **9** was proved with good selectivity and safety.

Compound **9** was found as a direct small molecule inhibitor (MW<500) of STAT3 with a very low K_i value, 440 nM. **9** also showed very low IC_{50} values in different cancer cell lines. **9** had excellent aqueous solubility. The oral bioavailability of **9** was 44.7% by PK study, comparable to the first-line drug, Gefitinib, 47.4%. Thus, **9** was an orally-bioavailable STAT3 inhibitor. All of its physic-chemical properties follow the Lipinski's rules of five. Comprehensive assessing **9** and other representative STAT3 inhibitors, **9** showed outstanding features among all. Further experiments are needed to evaluate the potential of **9** to undertake clinical trials.

CONCLUSION. AMLSD was used to design new STAT3 inhibitors by screening FCs built from the library of drug scaffolds. Among these compounds, **9** was proved to directly and selectively bind to the pY705 site of STAT3 protein, which significantly decreased the level of STAT3 phosphorylation/dimerization as a dual functional inhibitor. **9** also significantly shrank the tumor weight of both breast cancer and sarcoma cancer xenographs *in vivo* with better activity than first-line drug, Gefitinib, and less toxicity than Doxorubicin. The therapeutic index of **9** should be over 100, indicating good safety of **9**.

EXPERIMENTAL SECTION

Design

Software. The design part was mainly performed in Dr. Chenglong Li's lab in the Ohio State University. MLSD software was written by Dr. Li's lab.[32] Schrödinger packages was purchased from its commercial supplier and installed in English version. Autodock4 was downloaded from its official website http://autodock.scripps.edu/ for free and installed in English.

Binding sites analysis. There are two "hot spots" of STAT3 SH2 domain, pY705 and pY+X sites. According to the crystal structure of the pY+X site of STAT3 (PDB: 1BG1), several key polar residues were embedded inside and formed a polar binding region. Based on the structural analysis of the crystallized STAT1, STAT3 and STAT5, the pY705 site is highly conserved, however the pY+X site varies within this family of proteins, which plays an important role in the selectivity of STAT inhibitors. We combined fragment 1 with each fragment of the polar fragment libraries to perform the hopping of FCs in the two adjacent pY705 and pY+X sites, purposing to design selective inhibitors of STAT3.

Preparation of Fragment Libraries. The fragment libraries in our study were generated from 1554 FDA approved drugs, 4959 experimental drugs, 84 nutraceutical drugs and 521 investigational drugs collected by Drugbank database (https://www.drugbank.ca/). Complete database of all drugs were downloaded and prepared by LigPrep of Schrödinger using the default parameters.[33] Fragments were extracted from the drugs using scaffold decomposition program by Canvas of Schrödinger using default parameters.[22] Fragments were calculated with MW and Polar

using physicochemical descriptors by Canvas of Schrödinger.[34] All the scaffolds were exported and the scaffolds with molecular weight higher than 300 were deleted. The total charge of the fragments was calculated by ligfilter descriptors. Other properties were divided manually. Fragments were reorganized using property filters by Canvas of Schrödinger as six fragment libraries, including polar_uncharged_nonaromatic fragment library, polar_uncharged_aromatic fragment library, polar_charged_basic fragment library, polar_charged_acidic fragment library, large and bulk fragment library, and nonpolar fragment library.

Screening FCs and the Evolving of FCs to Final Hits. Two fragments as one fragment combination were grouped to probe the two "hot spots" of STAT3 SH2 domain, pY705 and pY+X sites. Fragment **1**, which was the privileged structure of the two most potent STAT3 SH2 inhibitors, **16** and **17**,[8, 35] formed two dipole-ionic interactions with residue R609 in the pY705 site analyzed by the MD simulations (Figure S3, (SI)). Thus, fragment **1** was chosen for the inhibition of the pY705 site. While, fragments fitting to the pY+X site were screened from the newly designed fragment libraries. Since the pY+X site of STAT3 has a key polar binding region, therefore, fragments fitting to pY+X site were screened from the four polar fragment libraries, a total of 963 fragments, including 280 polar uncharged nonaromatic fragments, 532 polar uncharged aromatic fragments, 61 polar charged basic fragments, and 90 polar charged acidic fragments. The PDB file of each fragment from the four polar fragment libraries was grouped together with the PDB file of fragment **1** and

finally formatted as one PDB file, resulting 963 fragment combinations. Each FC includes two fragments.

The FCs best fit for pY705 and pY+X sites were docked to the crystal structure of STAT3 SH2 domain and scored by the PSO method,[32] which were then ranked by their docking energies.

The evolving from fragments to one compound was guided by the docking modes of the FCs, which was no longer determined by designers' chemical intuition or experience. The shortest distances between C6 or C7 of fragment **1** and the heavy atoms of the other fragment were around 3~5 Å (Table S1, (SI)), measured by Maestro of Schrödinger. For example, for compound **9**, the shortest distance between C6 of fragment **1** and the other 2-(piperazin-1-yl) aniline fragment were 3.671 Å (Figure 5). Thus, the length of possible linear linkers should be within the range, listed in Figure 4. Every linker was tried to link the FCs, and the linker ***a1*** was finally decided on the synthetic feasibility and its repositioning ability. Linked compounds were repositioned to the SH2 domain scored by Lamarckian Genetic Algorithm (LGA) of Autodock4 and ranked by binding energies. Final hits were selected based on both the docking energies and their repositioned docking modes, which were further synthesized and tested by all kinds of biological assays.

The docking procedure involved the preparation of the ligand and macromolecule using the Schrödinger software.[33, 36] AutoDockTools were used to assign Gasteiger charges to the ligands. AutoGrid maps were then calculated for all atom types in the

ligand set. Grid box was centered in the SH2 domain of STAT proteins. After 10 million energy evaluations, all the resulting conformations of the ligands in the binding pocket of the macromolecule were clustered into groups according to their conformations with a root-mean-square deviation threshold of 1.5 Å. The most significant low energy clusters were identified and binding energies were evaluated.[37]

ADME Prediction. Designed compounds were prepared by LigPrep wizard in the Schrödinger software. The ADME properties of the compounds were calculated using QikProp wizard in the Schrödinger software, including whether (I) less than 5 maximum possible metabolites would be formed *in vivo*; (II) the logP value for octanol/water was in the range of -2 to 2 indicating excellent aqueous solubility profile; (III) predicted logIC50 values for hERG K+ channel blocking was below -5; (IV) predicted Caco-2 and MDCK cell permeability values were acceptable; (V) predicted index of binding to human serum albumin was well within the recommended range of -1.5 to 1.5; (VI) predicted human oral absorption percentage was above 50%.

Chemistry.

Chemicals and Reagents. The solvents and reagents were purchased from commercial suppliers and were used as received. Silica gel was purchased from Sigma-Aldrich Chemical Co. (Milwaukee,WI). All the synthetic derivatives and compounds purchased were obtained with >96.0% purity by HPLC (UV detection at 254 and 280 nm). Positive control **14** was purchased from commercial suppliers.

Synthesis of Fragment 1. 1 was synthesized as previously described.[8]

Synthesis of Designed Compounds. Fragment **1** was reacted with aromatic amines as previously described,[8,38] which were catalyzed by $Cu(OAc)_2 \cdot H_2O$ in an acidic environment. Reaction system was heated to 65 ℃ and terminated when **1** was consumed completely. Compounds **2**, **5**, **7** and **11** were synthesized.

Compound 2, 5, 11. Reported[38]

*5-((1,4-dioxo-5-sulfamoyl-1,4-dihydronaphthalen-2-yl)amino)nicotinic acid **(7)***. Yield: 71%. ^{1}H NMR (500 MHz, DMSO) δ 8.31 (s, 1H), 8.04 (d, J = 7.6 Hz, 1H), 7.98 (d, J = 2.5 Hz, 1H), 7.94 (d, J = 7.5 Hz, 1H), 7.78 (t, J = 7.5 Hz, 1H), 7.44 (s, 1H), 7.40 (s, 2H), 5.62 (s, 1H), 5.34 (s, 2H). HRMS (ESI) of $C_{16}H_{11}N_3O_6SNa$ $[M + Na]^+$ calcd, 396.0266; found, 396.0270.

*5,8-dioxo-6-((2-(piperazin-1-yl)phenyl)amino)-5,8-dihydronaphthalene-1-sulfonamide **(9)***. 11 (0.15 mM, 80 mg) was dissolved in 5 mL TFA, heated to 50°C for 8 hours. As detected by TLC, after 11 had completely reacted, all the volatiles were removed under reduced pressure. The resulting crude product was applied to a column of silica gel. The column was eluted with CH_2Cl_2/MeOH (20:1 v/v), then further purified by recrystallization. The solvent was removed under reduced pressure to give product **9**. Yield: 92%. ^{1}H NMR (300 MHz, DMSO) δ 9.15 (s, 1H), 8.85 (m, 2H), 8.43 (d, J = 7.1 Hz, 1H), 8.09 (m, 2H), 7.63 – 7.05 (m, 4H), 5.80 (s, 1H), 3.60 (m, 4H), 3.14 (s, 5H). HRMS (ESI) of $C_{20}H_{21}N_4O_4S$ $[M + H]^+$ calcd, 413.1284; found, 413.1290.

Biological Assays and Biochemical Assays.

Cell Cultures. Human medulloblastoma cell lines (UW426, UW288-1 and DAOY) and human pancreatic cancer cell line (BkPC-3), were used as stored in the Center for Childhood Cancer, the Research Institute at Nationwide Children's Hospital, in 2014. Human breast cancer cell lines (basal-like subtype MDA-MB-231 and luminal A subtype MCF-7), human osteosarcoma cell line (U2OS), human normal breast epithelial cell line (MCF-10A) and mouse sarcoma cell line (S180) were purchased from ATCC in 2015. All cell lines were tested without mycoplasma contamination.

Human cancer cell lines, UW426, UW288-1, DAOY, BkPC-3, MDA-MB-231 and MCF-7 were maintained in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% penicillin/streptomycin FBS and stored in a humidified 37 °C incubator with 5% CO_2. Human bone osteosarcoma (U2OS), human breast cell line MCF-10A and mouse sarcoma cell line (S180) were maintained in RPMI-1640 medium supplemented with 10% FBS and antibiotics (100 U/mL of penicillin G and 100 μg/mL of streptomycin) and stored in a humidified 37 °C incubator with 5% CO_2.

Chemicals. Compounds were dissolved in sterile DMSO to make 20 mM stock solutions. Aliquots of the stock solutions were stored at −20 °C.

Cell Viability Assay. Studies were performed as previously described.[8] UW426, UW288-1, BKPC3, MDA-MB-231, U2OS, S180 and MCF-10A cell lines were seeded in 96-well plates at a density of 3000 cells per well. The cells were incubated at 37 °C for a period of 24 h. For the preliminary assay, **2**, **5**, **7** and **9** (5 μM) were added in triplicate to the plates in the presence of 10% FBS, 24h. 3-(4,5-Dimethylthiazolyl)-2,5-

diphenyltetrazoliumbromide (MTT) was added to evaluate cell viability. For the IC_{50} measurements, different concentrations of **9** (0.1-10 μM for cancer cells and 3.125-400μM for MCF-10A), **14** (0.1-20 μM), and other compounds were added in triplicate to the plates in the presence of 10% FBS. For S180 cell line, cells were only treated with compound **9** (0.1-10 μM). MTT was added to evaluate cell viability. The absorbance was read at 595 nm. IC_{50} values were determined using SigmaPlot 9.0 Software (SySTAT Software, Inc., San Jose, CA).

Fluorescent Polarization. Studies were performed as previously described.[8,39] FP assay was used to analyze the ability of **9** to inhibit phosphopeptide binding to the STAT SH2 domain. STAT3 protein and peptides were purchased from Genscripts. STAT1/5 proteins were purchased from Abcam and their peptides were purchased from Taopushengwu, Inc., Shanghai, China.

Microscale Thermophoresis. A titration series of up to 16 dilutions of the inhibitor **9** was prepared. Here the concentration of the fluorescently labeled STAT3 protein was kept constant at 10-20 nM and the concentration of the titrant was varied. Therefore 2-20 μM of STAT3 protein were labeled with a fluorescent dye using Monolith NT™ Protein Labeling Kits. The labeling procedure and the subsequent removal of free dye were performed within 45 min. In the dilution series the highest concentration was chosen to be 20-fold higher than the expected K_i. 10 μL of the serial dilution of the non-labeled molecule was mixed with 10 μL of the diluted fluorescently

labeled molecule. Mixed samples were loaded into glass capillaries and the MST-analysis was performed using the Monolith.NT1[15].

Western Blot Analysis. Studies were performed as previously reported.[8] MDA-MB-231 cells were treated with **9** (2.5-10μM) or **14** (2.5-5μM) or DMSO at 60−80% confluence in the presence of 10% FBS for 12 h. UW426 cells were treated with **9** (1-2.5 μM) or **11** (1-2.5 μM) or DMSO at 60−80% confluence in the presence of 10% FBS for 24 h.

IL-6/IFN-γ Induction of STAT3/1 Phosphorylation. MDA-MB-231 cancer cells were seeded in 10 cm plates and allowed to adhere overnight. The following day, the cells were serum-starved. The cells were then left untreated or treated with **9** (2.5-5 μM) or DMSO. After 2 h, the untreated and treated cells were stimulated by IL-6/ IFN-γ (50 ng/mL). The cells were harvested after 30 min and analyzed by Western blot.

Quantitative real-time RT-PCR. Levels of mRNA expression were analyzed with the RT-PCR assay in MDA-MB-231 cells. Total RNA was isolated using an EASYspin Plus tissue/cell RNA extraction kit (Aidlab Biotechnologies Co. Ltd), which was quantified at 260 nm. 1μg RNA was reverse transcribed to cDNA using the Transcriptor First Strand cDNA Synthesis Kit (Roche Diagnostics, Basel, Switzerland). Thermal cycling conditions included 95 °C initial denaturation for 5 min, followed by 40 cycles of denaturation (10 s at 95 °C), annealing (15 s at 60 °C) and extension (15 s at 72 °C with a single fluorescence measurement), a melt curve program (60–95 °C with a 0.11 °C /s heat increase and continuous fluorescence measurement) and a cooling step to 40

°C. The Δ cycle threshold method was used for the calculation of relative differences in mRNA abundance with a LightCycler 480 (Roche Molecular Biochemicals, Mannheim, Germany). Data were normalized to the expression of GAPDH. The RT-PCR primers that were used in this study are listed in Table S4, (SI).

Immunofluorescence Staining. Studies were performed as previously described.[40]

Nuclear Extract Preparation and Gel Shift Assays. Nuclear extract preparations and EMSA were carried out as previously described.[41] MDA-MB-231 cells were pretreated with the **9** at 5μm at different time spots ranging from 0-24 h and then harvested for nuclear extract preparation for EMSA analysis at room temperature prior to incubation with the biotin-labeled probe for 30 min before subjecting to EMSA analysis. Serum-starved MDA-MB-231 cells were preincubated with **9** and then stimulated by IFN-γ, 50 ng/mL, and then harvested for nuclear extraction.

Wound-Healing Assay for Migration. Assays were performed as previously reported.[42]

***In Vivo* Tumor Studies.** The experiments were carried out in accordance with the guidelines issued by the Ethical Committee of China Pharmaceutical University (SYXK 2012-0035).

Mouse Xenograft Tumor Model. MDA-MB-231 human breast cancer cells (5 x 10^6) were injected subcutaneously into the flank area of 6-week-old female athymic SPF BALB/c nude mice which were purchased from Canvens (Changzhou, Jiangsu, China). After tumor volume reaching about 100 mm^3, mice were divided into five

treatment groups consisting of eight mice per group (each mouse generated one xenograft tumor): saline vehicle control, oral injection of **9** (2.5, 5 and 10 mg/kg or 6, 12, and 24 μmol/kg) and Gefitinib (40 mg/kg or 90μmol/kg). The inhibitors were directly dissolved in saline. Gefitinib was formulated with 0.05% tween80 to enhance solubility. Tumor growth was monitored by measuring length (L) and width (W) of the tumor every other day with a caliper. The tumor volume was calculated according to the formula: Tumor volume=$0.5236\times L\times W^2$. The treatment lasted for 32 days. The detailed dosage regimen is listed in Table S6, (SI). For dose 5 mg/kg, a repeated study was performed. Mice were divided into two groups, consisting of ten mice per group: saline vehicle control and oral injection of **9** (5 mg/kg).

S180 mouse sarcoma cancer cells (3×10^6) were injected subcutaneously into the flank area of 4-week-old male mice which were purchased from Canvens (Changzhou, Jiangsu, China). On the second day, mice were divided into two treatment groups consisting of ten mice per group (each mouse generated one xenograft tumor): saline vehicle control, oral injection of **9** (5 mg/kg) and *i.p.* injection of Doxorubicin (1 mg/kg or 1.8μmol/kg). Both compounds were directly dissolved in saline. Tumor growth was determined by measuring length (L) and width (W) of the tumor every other day with a caliper. The tumor volume was calculated according to the formula: Tumor volume=$0.5236\times L\times W^2$. The treatment lasted for 21 days. The detailed dosage regimen is listed in Table S7, (SI).

H&E Staining of Mice Organs. The five major organs (heart, liver, spleen, lung and kidney) of all groups were anatomized from the mice, which were then fresh fixed in 10% neutral-buffered formaldehyde for 48 h, embedded in paraffin, and sliced at 5 μm thickness. The sections were stained with haematoxylin and eosin (H&E), and examined by light microscopy.

IHC of Tumor Tissues. Immunohistochemical staining was carried out as previously described.[43] The fixed tumor sections were incubated with antibodies against p-STAT3. The DAB staining Kit (NanJing KeyGEN Biotechnology Co., Ltd., Nanjing, China) were used to visualize positive cells according to the manufacturers' recommendations.

The Acute Toxicity Test. Both genders of ICR mice (7 weeks, SLACCAS) were randomized and treated intraperitoneally (*i.p.*) with a single dose of **9** at 200, 400, 600, 800 and 1000 mg/kg, respectively. Each group contained 10 mice, and the animals were observed for abnormal behavior and mortality up to 2 weeks post treatment.

ASSOCIATED CONTENT

SUPPORTING INFORMATION

Supporting Information on design, biological assays, PAINS assay and methods, method validation of pharmacokinetic study and molecular formula strings. This material is available free of charge *via* the Internet at http://pubs.acs.org.

Accession Codes

STAT3β/DNA complex: Protein Data Bank ID 1BG1; STAT1/DNA complex: Protein Data Bank ID 1BF1; STAT5A: Protein Data Bank ID 1Y1U.

Authors will release the atomic coordinates of STAT3-**9**-(FC of **9**) complex, STAT1-**9** complex, STAT5-**9** complex and experimental data upon article publication.

AUTHOR INFORMATION

Corresponding Author

*E-mail: cpu_lykong@126.com; Phone: +86-25-8327-1405.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

We are grateful to Dr. Inhee Park, Dr. Sihui Long and Dr. Feng Zhang for paper editing. This work was supported by National Natural Science Foundation of China (Grant No. 81402791 and No. 81673298), Natural Science Foundation of the Jiangsu Higher Education Institutions of China (Grant No. BK20140670), Program for Changjiang Scholars and Innovative Research Team in University (IRT_15R63), Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD), College Students Innovation Project for the R&D of Novel Drugs (J1030830), Jiangsu Shuangchuang Talents for Oversea Ph.D.s.

ABBREVIATIONS USED

FBDD, fragment-based drug design; SH2 domain, Src Homology 2 domain; pY705, phosphor-tyrosine 705; PK/PD, Pharmacokinetic/Pharmacodynamic; FAM, carboxy-fluorescein; IL-6, Interleukin 6; AKT, Protein kinase B; H&E staining, Hematoxylin and eosin staining; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; Ac-STAT3, acetyl-STAT3; P-STAT3, phosphor-STAT3; T-STAT3, total-STAT3; FDA, Food and Drug Administration.

REFERENCES

1. Liu, C. Y.; Su, J. C.; Ni, M. H.; Tseng, L. M.; Chu, P. Y.; Wang, D. S.; Tai, W. T.; Kao, Y. P.; Hung, M. H.; Shiau, C. W.; Chen, K. F. Obatoclax analog SC-2001 inhibits STAT3 phosphorylation through enhancing SHP-1 expression and induces apoptosis in human breast cancer cells. *Breast Cancer Res. Treat.* **2014,** *146*, 71-84.

2. Wei, C.; Ball, S.; Lin, L.; Liu, A.; Fuchs, J. R.; Li, P.; Li, C.; Lin, J. Two small molecule compounds, LLL12 and FLLL32, exhibit potent inhibitory activity on STAT3 in human rhabdomyosarcoma cells. *Int. J. Oncol.* **2010,** 38.

3. Chen, H.; Yang, Z.; Ding, C.; Xiong, A.; Wild, C.; Wang, L.; Ye, N.; Cai, G.; Flores, R. M.; Ding, Y.; Shen, Q.; Zhou, J. Discovery of potent anticancer agent HJC0416, an orally bioavailable small molecule inhibitor of signal transducer and activator of transcription 3 (STAT3). *Eur. J. Med. Chem.* **2014,** *82*, 195-203.

4. Yue, P.; Turkson, J. Targeting STAT3 in cancer: how successful are we? *Expert Opin. Invest. Drugs* **2009,** *18*, 45-56.

5. Bharadwaj, U.; Eckols, T. K.; Kolosov, M.; Kasembeli, M. M.; Adam, A.; Torres, D.; Zhang, X.; Dobrolecki, L. E.; Wei, W.; Lewis, M. T.; Dave, B.; Chang, J. C.; Landis, M. D.; Creighton, C. J.; Mancini, M. A.; Tweardy, D. J. Drug-repositioning screening identified piperlongumine as a direct STAT3 inhibitor with potent activity against breast cancer. *Oncogene* **2014**, 1-14.

6. Ma, D. L.; Liu, L. J.; Leung, K. H.; Chen, Y. T.; Zhong, H. J.; Chan, D. S.; Wang, H. M.; Leung, C. H. Antagonizing STAT3 dimerization with a rhodium(III) complex. *Angew. Chem. Int. Ed. Engl.* **2014,** *53*, 9178-9182.

7. Zhang, X.; Yue, P.; Page, B. D.; Li, T.; Zhao, W.; Namanja, A. T.; Paladino, D.; Zhao, J.; Chen, Y.; Gunning, P. T.; Turkson, J. Orally bioavailable small-molecule inhibitor of transcription factor STAT3 regresses human breast and lung cancer xenografts. *Proc. Natl. Acad. Sci. U. S. A.* **2012,** *109*, 9623-9628.

8. Yu, W.; Xiao, H.; Lin, J.; Li, C. Discovery of novel STAT3 small molecule inhibitors via in silico site-directed fragment-based drug design. *J. Med. Chem.* **2013,** *56*, 4402-4412.

9. Schuringa, J. J.; Schepers, H.; Vellenga, E.; Kruijer, W. Ser727-dependent transcriptional activation by association of p300 with STAT3 upon IL-6 stimulation. *FEBS Lett.* **2001,** *495*, 71-76.

10. Lee, H.; Zhang, P.; Herrmann, A.; Yang, C.; Xin, H.; Wang, Z.; Hoon, D. S.; Forman, S. J.; Jove, R.; Riggs, A. D.; Yu, H. Acetylated STAT3 is crucial for methylation of tumor-suppressor gene promoters and inhibition by resveratrol results in demethylation. *Proc. Natl. Acad. Sci. U. S. A.* **2012,** *109*, 7765-7769.

11. Clevenger, C. V. Roles and regulation of STAT family transcription factors in human breast cancer. *Am. J. Pathol.* **2004,** *165*, 1449-1460.

12. Mandal, P. K.; Gao, F.; Lu, Z.; Ren, Z.; Ramesh, R.; Birtwistle, J. S.; Kaluarachchi, K. K.; Chen, X.; Bast, R. C., Jr.; Liao, W. S.; McMurray, J. S. Potent and selective phosphopeptide mimetic prodrugs targeted to the src homology 2 (SH2) domain of signal transducer and activator of transcription 3. *J. Med. Chem.* **2011,** *54*, 3549-3563.

13. Mandal, P. K.; Ren, Z.; Chen, X.; Xiong, C.; McMurray, J. S. Structure-affinity relationships of glutamine mimics incorporated into phosphopeptides targeted to the SH2 domain of signal transducer and activator of transcription 3. *J. Med. Chem.* **2009,** *52*, 6126-6141.

14. Song, H.; Wang, R.; Wang, S.; Lin, J. A low-molecular-weight compound discovered through virtual database screening inhibits STAT3 function in breast cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** *102*, 4700-4705.

15. Siddiquee, K.; Zhang, S.; Guida, W. C.; Blaskovich, M. A.; Greedy, B.; Lawrence, H. R.; Yip, M. L.; Jove, R.; McLaughlin, M. M.; Lawrence, N. J.; Sebti, S.

M.; Turkson, J. Selective chemical probe inhibitor of STAT3, identified through structure-based virtual screening, induces antitumor activity. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104*, 7391-7396.

16. Fletcher, S.; Page, B. D.; Zhang, X.; Yue, P.; Li, Z. H.; Sharmeen, S.; Singh, J.; Zhao, W.; Schimmer, A. D.; Trudel, S.; Turkson, J.; Gunning, P. T. Antagonism of the STAT3-STAT3 protein dimer with salicylic acid based small molecules. *ChemMedChem* **2011,** *6*, 1459-1470.

17. Liu, L. J.; Leung, K. H.; Chan, D. S.; Wang, Y. T.; Ma, D. L.; Leung, C. H. Identification of a natural product-like STAT3 dimerization inhibitor by structure-based virtual screening. *Cell Death Dis.* **2014,** *5*, e1293.

18. Furqan, M.; Akinleye, A.; Mukhi, N.; Mittal, V.; Chen, Y.; Liu, D. STAT inhibitors for cancer therapy. *J. Hematol. Oncol.* **2013,** *6*, 90.

19. Oh, D. Y.; Lee, S. H.; Han, S. W.; Kim, M. J.; Kim, T. M.; Kim, T. Y.; Heo, D. S.; Yuasa, M.; Yanagihara, Y.; Bang, Y. J. Phase I study of OPB-31121, an oral STAT3 inhibitor, in patients with advanced solid tumors. *Cancer Res. Treat.* **2015,** *47*, 607-615.

20. Kim, M. J.; Nam, H. J.; Kim, H. P.; Han, S. W.; Im, S. A.; Kim, T. Y.; Oh, D. Y.; Bang, Y. J. OPB-31121, a novel small molecular inhibitor, disrupts the JAK2/STAT3 pathway and exhibits an antitumor activity in gastric cancer cells. *Cancer Lett.* **2013,** *335*, 145-152.

21. Okusaka, T.; Ueno, H.; Ikeda, M.; Mitsunaga, S.; Ozaka, M.; Ishii, H.; Yokosuka, O.; Ooka, Y.; Yoshimoto, R.; Yanagihara, Y.; Okita, K. Phase 1 and pharmacological trial of OPB-31121, a signal transducer and activator of transcription-3 inhibitor, in patients with advanced hepatocellular carcinoma. *Hepatol. Res.* **2015,** *45*, 1283-1291.

22. Bendell, J. C.; Hong, D. S.; Burris, H. A., 3rd; Naing, A.; Jones, S. F.; Falchook, G.; Bricmont, P.; Elekes, A.; Rock, E. P.; Kurzrock, R. Phase 1, open-label, dose-escalation, and pharmacokinetic study of STAT3 inhibitor OPB-31121 in subjects with advanced solid tumors. *Cancer Chemother. Pharmacol.* **2014,** *74*, 125-130.

23. Hayakawa, F.; Sugimoto, K.; Harada, Y.; Hashimoto, N.; Ohi, N.; Kurahashi, S.; Naoe, T. A novel STAT inhibitor, OPB-31121, has a significant antitumor effect on leukemia with STAT-addictive oncokinases. *Blood Cancer J.* **2013,** *3*, e166.

24. Zhong, H. J.; Lu, L. H.; Leung, K. H.; Wong, C. C. L.; Peng, C.; Yan, S. C.; Ma, D. L.; Cai, Z. W.; Wang, H. M. D.; Leung, C. H. An iridium(III)-based irreversible protein-protein interaction inhibitor of BRD4 as a potent anticancer agent. *Chem. Sci.* **2015,** *6*, 5400-5408.

25. Leung, C. H.; Chan, D. S.; Yang, H.; Abagyan, R.; Lee, S. M.; Zhu, G. Y.; Fong, W. F.; Ma, D. L. A natural product-like inhibitor of NEDD8-activating enzyme. *Chem. Commun.* **2011,** *47*, 2511-2513.

26. Bajorath, J. Computer-aided drug discovery. *F1000Res.* **2015,** 4.

27. Hu, Y.; Stumpfe, D.; Bajorath, J. Computational exploration of molecular scaffolds in medicinal chemistry. *J. Med. Chem.* **2016,** *59*, 4062-4076.

28. Zhang, X.; Yue, P.; Fletcher, S.; Zhao, W.; Gunning, P. T.; Turkson, J. A novel small-molecule disrupts STAT3 SH2 domain-phosphotyrosine interactions and STAT3-dependent tumor processes. *Biochem. Pharmacol.* **2010,** *79*, 1398-1409.

29. Lin, L.; Hutzen, B.; Li, P. K.; Ball, S.; Zuo, M.; DeAngelis, S.; Foust, E.; Sobo, M.; Friedman, L.; Bhasin, D.; Cen, L.; Li, C.; Lin, J. A novel small molecule, LLL12, inhibits STAT3 phosphorylation and activities and exhibits potent growth-suppressive activity in human cancer cells. *Neoplasia* **2010,** *12*, 39-50.

30. Szelag, M.; Czerwoniec, A.; Wesoly, J.; Bluyssen, H. A. Identification of STAT1 and STAT3 specific inhibitors using comparative virtual screening and docking validation. *PLoS One* **2015,** *10*, e0116688.

31. Bohm, H. J. The computer program LUDI: a new method for the de novo design of enzyme inhibitors. *J. Comput. Aided Mol. Des.* **1992,** *6*, 61-78.

32. Li, H.; Li, C. Multiple ligand simultaneous docking: orchestrated dancing of ligands in binding sites of protein. *J. Comput. Chem.* **2010,** *31*, 2014-2022.

33. Schrödinger Release 2015-2: *LigPrep,* version 3.4, Schrödinger, LLC, New York, NY. 2015.

34. Schrödinger Release 2015-2: *Canvas,* version 2.4, Schrödinger, LLC, New York, NY. 2015.

35. Bhasin, D.; Etter, JP.; Chettiar, SN.; Mok, M.; Li, PK. Antiproliferative activities and SAR studies of substituted anthraquinones and 1,4-naphthoquinones. *Bioorg. Med. Chem. Lett*. **2013,** *23*, 6864-6867.

36. Schrödinger Release 2015-2: *Protein Preparation Wizard; Epik version 3.2*; *Impact version 6.7*; *Prime version 4.0,* Schrödinger, LLC, New York, NY. 2015.

37. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. Autodock4 and AutoDockTools4: automated docking with selective receptor flexiblity. *J. Comput. Chem.* **2009,** *16*, 2785-2791.

38. Yu, W.; Li, C. Regioselective one-pot C-N coupling of substituted naphthoquinones: selective intramolecular ring fusion of sulfonamides. *Tetrahedron* **2014,** *70*, 459-464.

39. Muller, J.; Schust, J.; Berg, T. A high-throughput assay for signal transducer and activator of transcription 5b based on fluorescence polarization. *Anal. Biochem.* **2008,** *375*, 249-254.

40. Zhao, W.; Jaganathan, S.; Turkson, J. A cell-permeable STAT3 SH2 domain mimetic inhibits STAT3 activation and induces antitumor cell effects in vitro. *J. Biol. Chem.* **2010,** *285*, 35855-35865.

41. Shahani, V. M.; Yue, P.; Fletcher, S.; Sharmeen, S.; Sukhai, M. A.; Luu, D. P.; Zhang, X.; Sun, H.; Zhao, W.; Schimmer, A. D.; Turkson, J.; Gunning, P. T. Design, synthesis, and in vitro characterization of novel hybrid peptidomimetic inhibitors of STAT3 protein. *Bioorg. Med. Chem.* **2011,** *19*, 1823-1838.

42. Liu, A.; Liu, Y.; Xu, Z.; Yu, W.; Wang, H.; Li, C.; Lin, J. Novel small molecule, XZH-5, inhibits constitutive and interleukin-6-induced STAT3 phosphorylation in human rhabdomyosarcoma cells. *Cancer Sci.* **2011,** *102*, 1381-1387.

43. Wang, H.; Jiang, M.; Cui, H.; Chen, M.; Buttyan, R.; Hayward, S. W.; Hai, T.; Wang, Z.; Yan, C. The stress response mediator ATF3 represses androgen signaling by binding the androgen receptor. *Mol. Cell. Biol.* **2012,** *32*, 3190-3202.

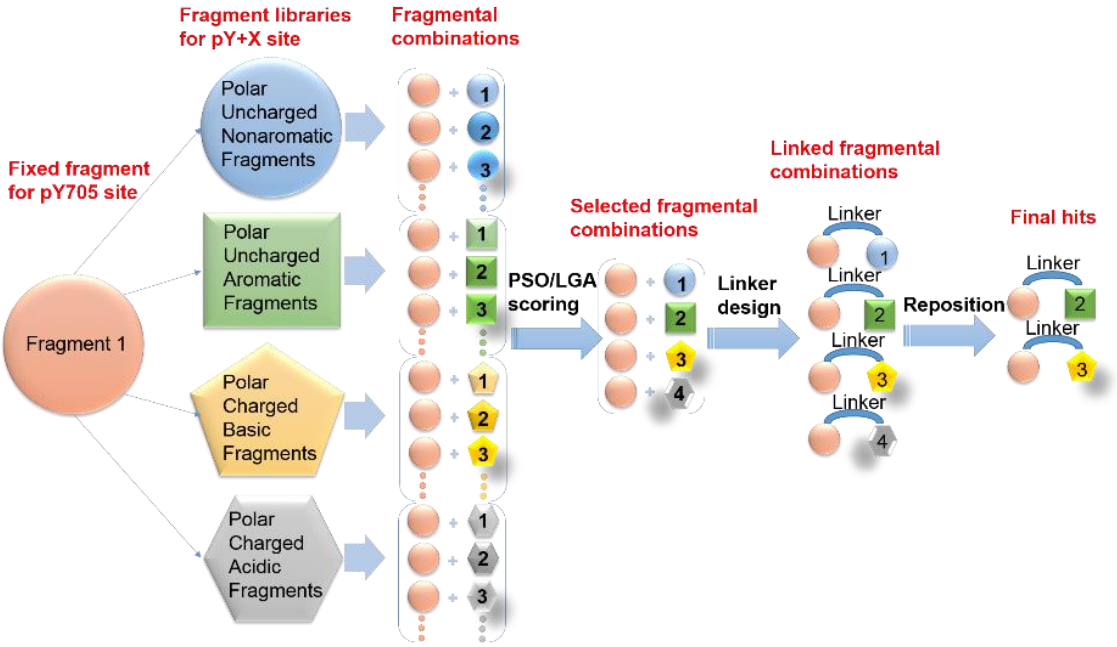


Figure 1. Strategy of AMLSD for STAT3 target. It proceeds as follows: (1) Fragment **1** paired with each fragment from the polar fragment libraries form a new library of the fragment combinations, which are simultaneously docked to the pY705 and pY+X sites of the STAT3 protein (PDB:1BG1) and scored by PSO/LGA functions. (2) The best scored FCs are linked according to the analysis of their docking modes and re-docked to the STAT3 protein. (3) The top scored linked compounds are repositioned to their docking modes of the unlinked FCs. Final hits are selected based on the criteria of both their repositioned docking modes and scored energies.

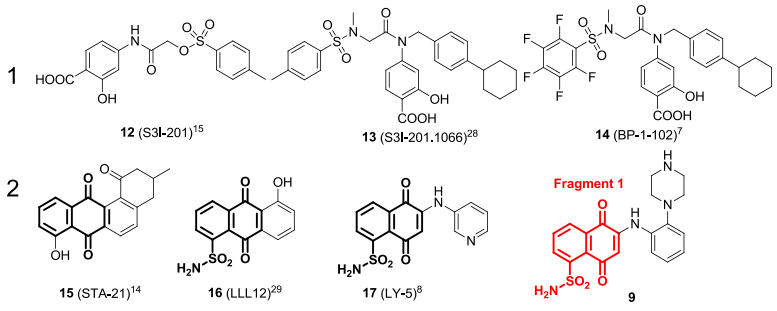


Figure 2. Representative small molecule STAT3 inhibitors. **9** was designed by this study and fragment **1** was colored as red.

Table 1. The comparison among **9** and other known STAT3 inhibitors.

STAT3 inhibitors	Linpiski's rules of five		Orally available for *in vivo* study	Selectivity between STAT	
	MW<500	LogP[a]<5		STAT3/1	STAT3/5
9	O	O	O	O	O
12	O	O	*i.v.*	O	Weak
13	550	5.4	*i.v.*	O	O
14	626	5.7	O	O	O
15	O	O	*i.v.*	O	O
16	O	O	*i.p.*	O	N/A
17	O	O	*i.p.*	O	Weak

O: Marked if following the titled requirement, otherwise, the violations were listed in detail. Not: Means "not selective". Weak: Means "weak selectivity". N/A: Marked if the data was not reported. [a]: LogP was calculated by Chemdraw.

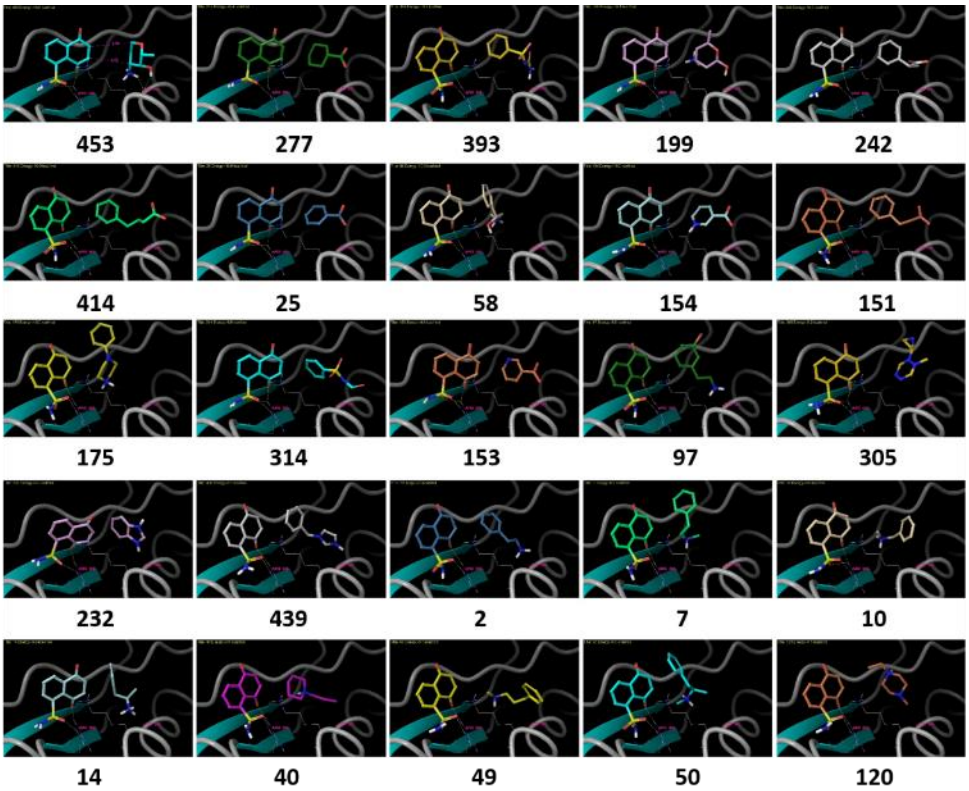


Figure 3 The docking modes of the 25 FCs selected by AMLSD, titled by their screening numbers in the library of FCs.

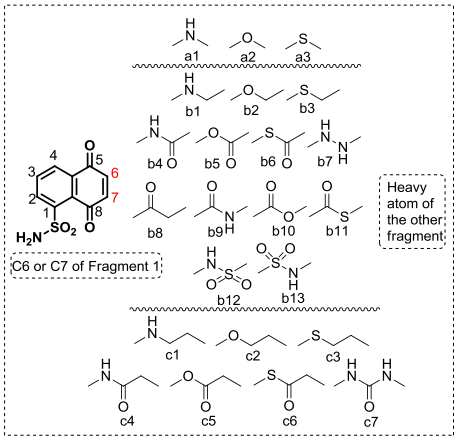


Figure 4. Possible linear linkers connecting the C6 or C7 of fragment **1** to the heavy atoms of the other fragment. The length of linkers is in the range of 3~5 Å.

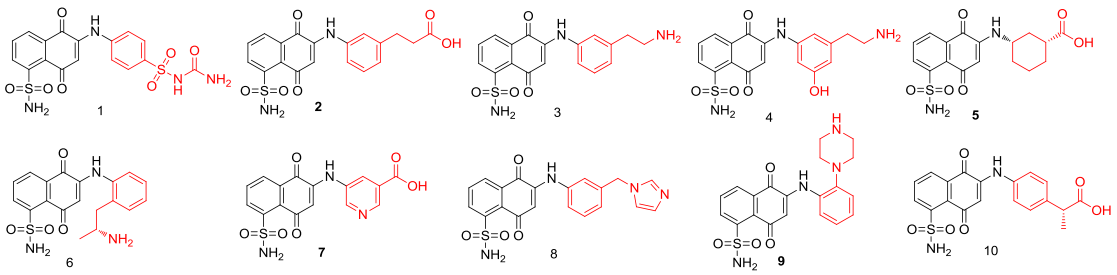
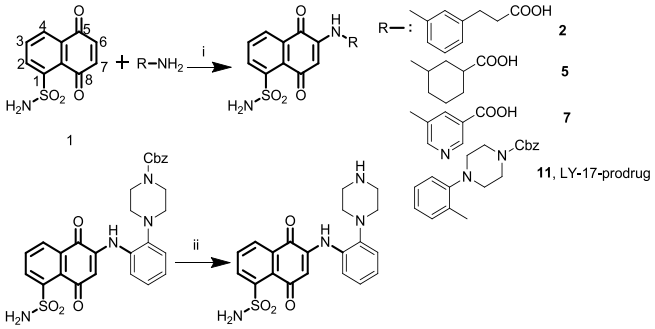


Figure 5. Chemical structures of the linked compounds designed by AMLSD.

Scheme 1. Synthesis of designed inhibitors **2**, **5**, **7** and **9** [a]



[a]Reagents and conditions: (i) $Cu(OAc)_2 \cdot H_2O$, AcOH and H_2O (v:v 1:10), 65℃. (ii) TFA, 50°C.

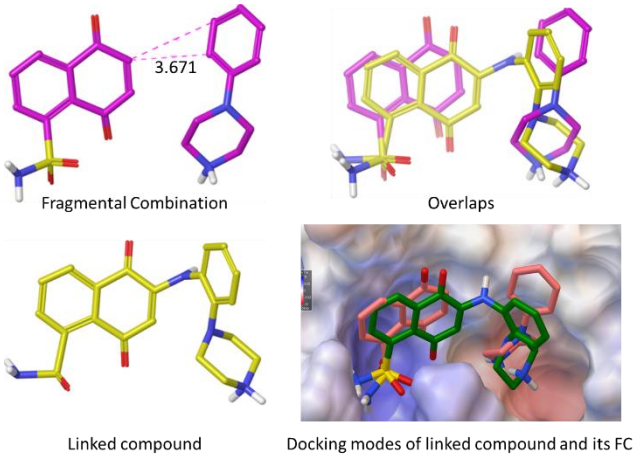


Figure 6. The linked compound **9** was repositioned to the docking mode of its FC to compare their overlaps. The shortest distance between the two fragments was 3.671 Å.

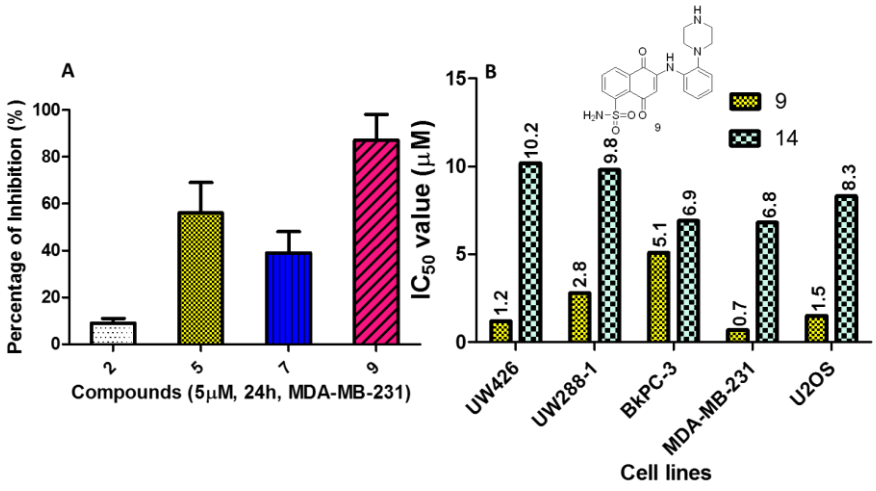


Figure 7. IC_{50} of **9** and positive control **14** against multiple cancer cell lines.

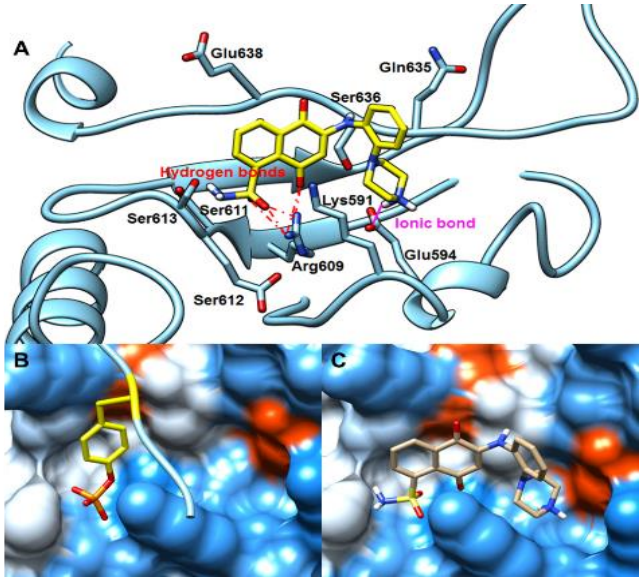


Figure 8. (A) Interactions between **9** and key residues of STAT3 protein. (B) Interactions between the native peptide and the pY705 site. (C) The docking modes of **9** to the same binding sites as the native peptide.

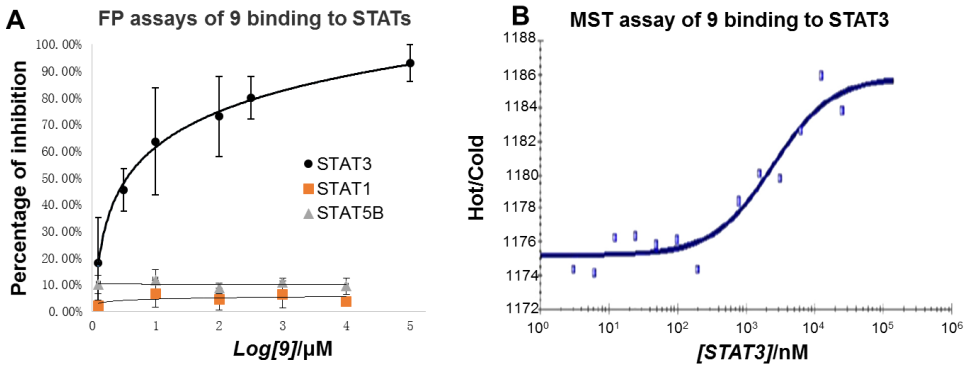


Figure 9. (A) **9** competitively and selectively inhibited the binding of fluorescent-labeled peptide to the pY705 site of STAT3 by the FP assays. (B) **9** directly bound to the STAT3 protein by the MST assays.

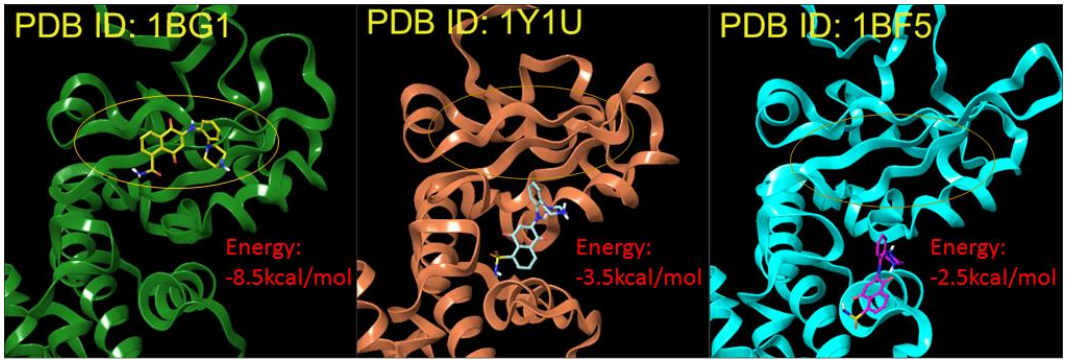


Figure 10. 9 was docked to the crystal structures of STAT1 (PDB: 1BF5)/3(PDB: 1BG1)/5(PDB: 1Y1U), their docking energies were marked in red. The yellow line circled region was the pY site of proteins.

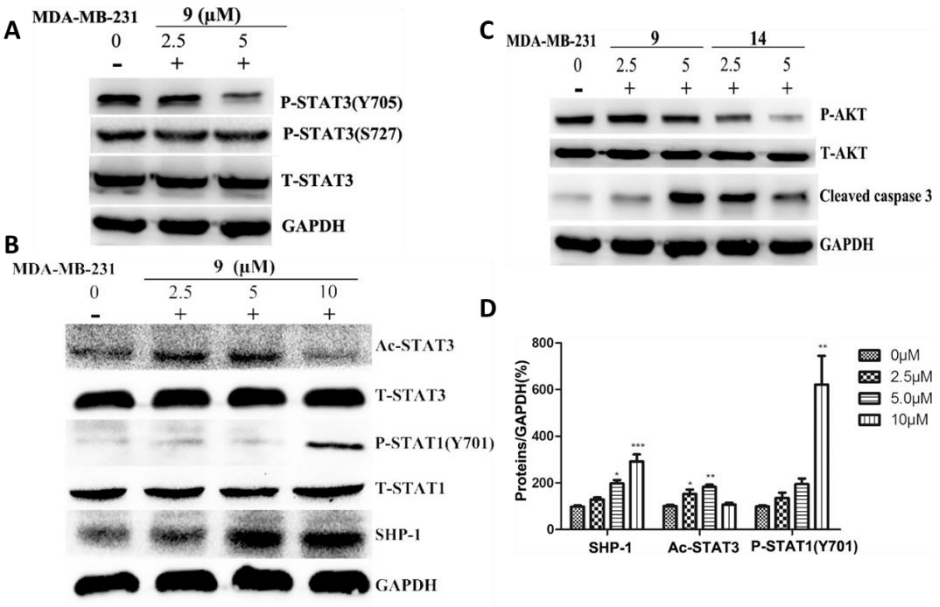


Figure 11. (A&B) Immunoblots of P-STAT3(Y705), P-STAT3(S727), T-STAT3, Ac-STAT3, P-STAT1(Y701), T-STAT1, SHP-1 or GAPDH from whole-cell lysates from MDA-MB-231 cells were treated with **9**. (C) MDA-MB-231 cells were treated with **9** and **14**, probing for P-AKT, T-AKT, Cleaved caspase 3, or GAPDH. (D) Western blot results from (B) was calculated and represented as the percent of control. (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

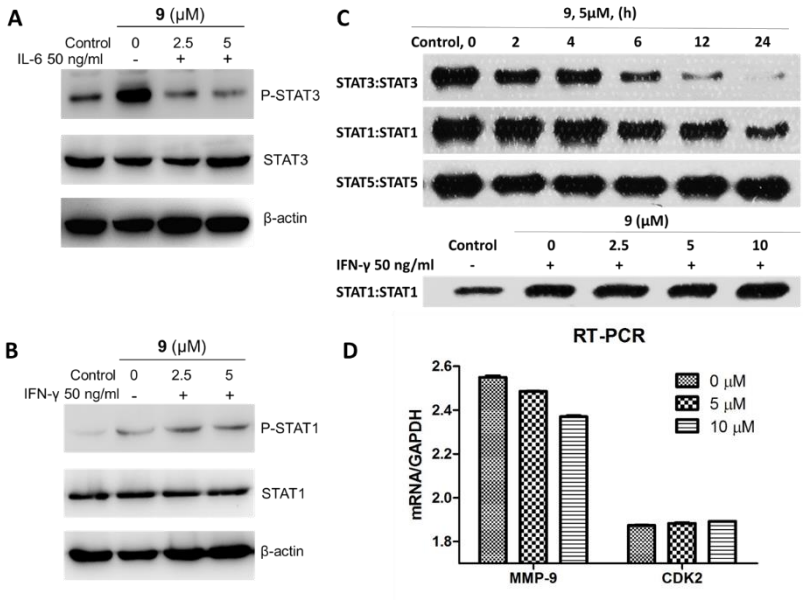


Figure 12. (A&B) MDA-MB-231 cancer cells were treated with IL-6, IFN-γ, and **9**, probing for P-STAT3(Y705), T-STAT3, P- STAT1(Y701), T-STAT1, or β-actin. (C) MDA-MB-231 cells were treated with **9** at 5 μM for 0~24h or IFN-γ and different concentrations of **9** by EMSA assays. (D) The expression of MMP9 and CDK2 genes treated with **9** was measured by RT-PCR.

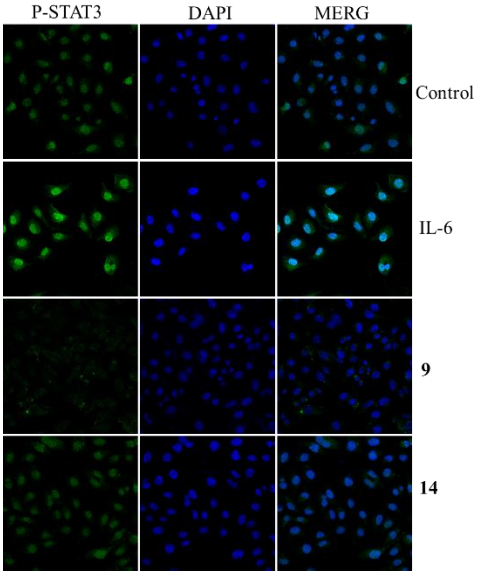


Figure 13. MDA-MB-231 cells was stimulated with IL-6. The migration of p-STAT3 to nucleus was treated with **9** and **14** (5 μM).

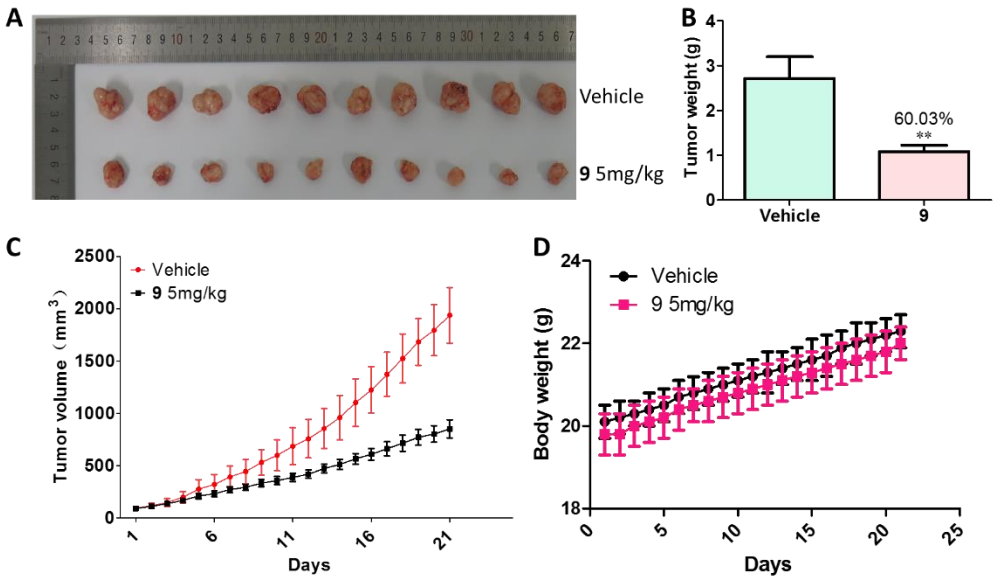


Figure 14. (A) The anatomical nude mice's tumor tissues untreated and treated by **9**. (B) The inhibiting percentage of mice's tumor weight treated by **9**. (C) The growing curves of tumor volume treated by **9**. (D) The growing curves of mice's body weight treated by **9**.

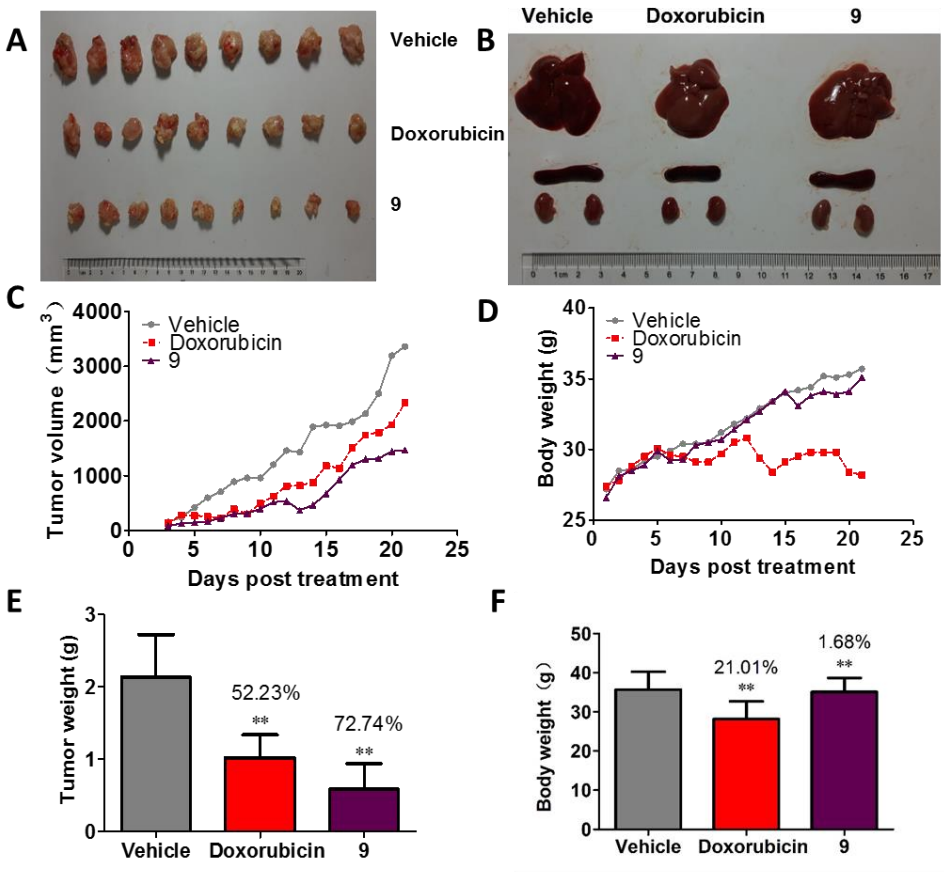


Figure 15. (A) The anatomical mice's tumor tissues treated by **9** and Doxorubicin. (B) The anatomical major organs, liver, spleen and kidney, treated by **9** and Doxorubicin. (C) The growing curves of mice's tumor volume. (E) The inhibiting percentage of mice's tumor weight. (D&F) The growing curves and the inhibiting percentage of mice's body weight.

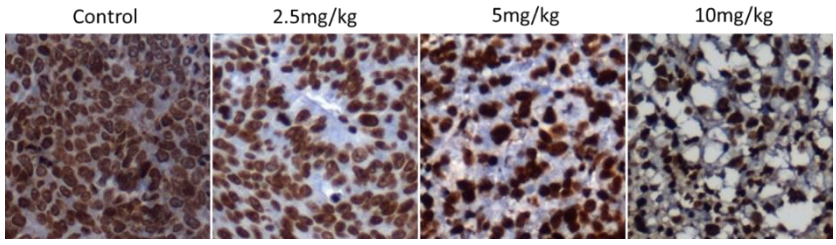


Figure 16. The IHC analysis of p-STAT3 in tumor tissues treated with **9**.

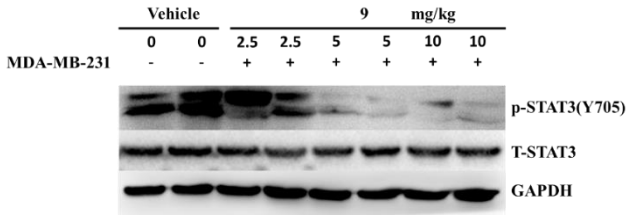


Figure 17. Immunoblots of P-STAT3(Y705), T-STAT3 or GAPDH from whole-cell lysates from tumor tissues treated with 2.5, 5, 10 mg/kg compound **9**. Positions of proteins in gel are shown with control lane.

Table 2. Main pharmacokinetic parameters of **9**.

		Gefitinib		**9**	
Parameter	Units	*i.v.* (1 mg/kg)	*p.o.* (10 mg/kg)	*i.v.* (1 mg/kg)	*p.o.* (10 mg/kg)
$T_{1/2}$	h	2.39±1.18	2.17±0.93	0.71±0.34	1.98±0.23
T_{max}	h	0.033	1	0.033	0.5
C_{max}	ng/ml	621.3±83.5	1202.7±140.0	359.0±77.5	540.3±147.0
C_0	ng/ml	714.8±137.4		382.3±73.6	
AUC_{0-t}	h*ng/ml	490.6±73.7	2356.79±43.5	189.0±54.4	850.0±252.9
$AUC_{0-\infty}$	h*ng/ml	500.8±72.4	2371.1±44.7	199.2±56.3	889.0±236.9
V_z	ml/kg	6678.0±2397.4	13238.7±5903.4	4944.2±925.1	32092.5±7550.5
Cl	ml/h/kg	2028.3±300.1	4218.5±79.2	5261.7±1289.7	11766.8±2612.3
$AUMC_{0-t}$	h*h*ng/ml	1098.0±734.0	5649.7±677.1	109.3±59.4	1622.6±718.2
$AUMC_{0-\infty}$	h*h*ng/ml	1261.5±845.2	5903.9±687.9	149.0±85.1	2002.1±575.9
MRT_{0-t}	h	2.12±1.17	2.40±0.30	0.55±0.14	1.87±0.37
$MRT_{0-\infty}$	h	2.40±1.32	2.49±0.31	0.71±0.20	2.25±0.25
F		47.4%±0.9%		44.7%±11.9%	

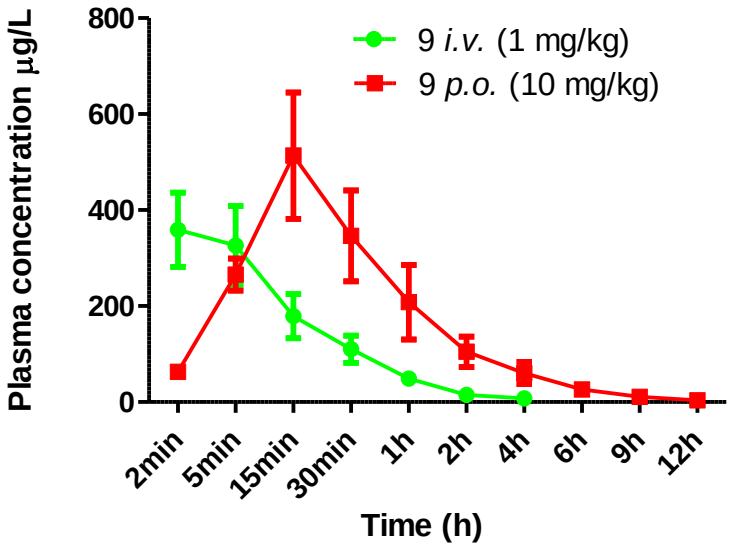


Figure 18. Plasma concentration−time profiles of **9** in rats following *i.v.* and *p.o.* administration.

Insert Table of Contents Graphic and Synopsis Here