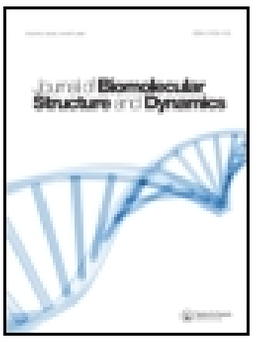


Journal of Biomolecular Structure and Dynamics

ISSN: 0739-1102 (Print) 1538-0254 (Online) Journal homepage: https://www.tandfonline.com/loi/tbsd20

Design, synthesis, biological evaluation, Common feature pharmacophore model and molecular dynamics simulation studies of ethyl 4-(phenoxymethyl)-2-phenylthiazole-5-carboxylate as Src Homology-2 Domain Containing Protein Tyrosine Phosphatase-2 (SHP2) inhibitors

Jingwei Wu, Weiya Li, Zhihui Zheng, Xinhua Lu, Huan Zhang, Ying Ma & Runling Wang

To cite this article: Jingwei Wu, Weiya Li, Zhihui Zheng, Xinhua Lu, Huan Zhang, Ying Ma & Runling Wang (2020): Design, synthesis, biological evaluation, Common feature pharmacophore model and molecular dynamics simulation studies of ethyl 4-(phenoxymethyl)-2-phenylthiazole-5-carboxylate as Src Homology-2 Domain Containing Protein Tyrosine Phosphatase-2 (SHP2) inhibitors, Journal of Biomolecular Structure and Dynamics, DOI: 10.1080/07391102.2020.1726817

To link to this article: https://doi.org/10.1080/07391102.2020.1726817

Accepted author version posted online: 10 Feb 2020.

Design, synthesis, biological evaluation，Common feature pharmacophore model and molecular dynamics simulation studies of ethyl 4-(phenoxymethyl)-2-phenylthiazole-5-carboxylate as Src Homology-2 Domain Containing Protein Tyrosine Phosphatase-2 (SHP2) inhibitors

Jingwei Wu [a, ‡], Weiya Li [a, ‡], Zhihui Zheng[b], Xinhua Lu[b], Huan Zhang [c], Ying Ma [a, *], Runling Wang [a, *]

[a]Tianjin Key Laboratory on Technologies Enabling Development of Clinical Therapeutics and Diagnostics (Theranostics), School of Pharmacy, Tianjin Medical University, Tianjin 300070, China

[b]New Drug Research and Development Center of North China Pharmaceutical Group Corporation, National Microbial Medicine Engineering and Research Center, Hebei Industry Microbial Metabolic Engineering &Technology Research Center, Key Laboratory for New Drug Screening Technology of Shijiazhuang City, Shijiazhuang, Hebei, China.

[c]Department of Pharmacy, Tianjin Medical University General Hospital, Tianjin 300052, China; Tel/Fax: +86-22-60362235;

‡ These authors contributed equally to this work

*Corresponding authors: Ying Ma, E-mail: maying@tmu.edu.cn, Runling Wang , E-mail:wangrunling@tmu.edu.cn

Keywords: SHP2; inhibitors; common feature pharmacophore; synthesis; molecular dynamics simulation

Abstract

SHP2 is a non-receptor protein tyrosine phosphatase (PTP) encoded by the PTPN11 gene involved in cell death pathway (PD-1/PD-L1) and cell growth and differentiation pathway (MAPK). Moreover, mutations in SHP2 have been implicated in Leopard syndrome(LS),Noonan syndrome (NS), juvenile myelomonocytic leukemia (JMML) and several types of cancer and solid tumors. Thus, SHP2 inhibitors are much needed reagents for evaluation of SHP2 as a therapeutic target. A series of novel ethyl 4-(phenoxymethyl)-2-phenylthiazole-5-carboxylate derivatives were designed and synthesized, and their SHP2 inhibitory activities (IC_{50}) were determined. Among the desired compounds, **1d** shares the highest inhibitory activity (IC_{50} = 0.99 μM) against SHP2. Additionally, a common feature pharmacophore model was established to explain the structure activity relationship of the desired compounds. Finally, molecular dynamics simulation was carried out to explore the most likely binding mode of compound **1d** with SHP2. In brief, the findings reported here may at least provide a new strategy or useful insights in discovering novel effective SHP2 inhibitors.

List of abbreviations

ANM (Anisotropic Network Mode); DMF (N,N-dimethylformamide); ESI (electro-spray ionization); JMML (juvenile myelomonocytic leukemia); LS (Leopard

syndrome); MD (molecular dynamics); MM-PBSA (molecular mechanics Poisson Boltzmann surface area); NMR (nuclear magnetic resonance); NS (Noonan syndrome); PDB (protein data bank); PTKs (protein tyrosine kinases); PTPs (protein tyrosine phosphatases); pNPP (para-nitrophenyl phosphate); RMSD (root means square deviation); RMSF (root mean square fluctuation); SH2 (Src homology-2); SHP2 (Src homology-2 domain containing protein tyrosine phosphatase-2); SAR (structure activity relationship); TLC (thin-layer chromatography);

1 Introduction

Lots of fundamental cellular events are regulated by protein dephosphorylation and phosphorylation controlled by protein-tyrosine phosphatases (PTPs) and protein-tyrosine kinases (PTKs), respectively(Taute et al., 2019). Not surprisingly, disturbance of the normal balance between PTK and PTP activity can result in abnormal tyrosine phosphorylation(Verma & Sharma, 2018), which has been linked to various human diseases, including cancer, cardiovascular and neurological disorders, metabolic and autoimmune diseases(Hendriks, Bourgonje, Leenders, & Pulido, 2018; Khan, Rhett, & O'Bryan, 2019). Because of opposing the action of PTKs, PTPs had been considered to be prime suspects for potential tumor suppressor genes(Ruckert, de Andrade, Santos, & Silveira, 2019). Hence, protein tyrosine phosphatases have been identified as high value therapeutic targets(Neel & Tonks, 2016; Van Huijsduijnen, Bombrun, & Swinnen, 2002).

The Src homology-2 domain (SH2) containing protein tyrosine phosphatase-2 (SHP2), is a non-receptor protein tyrosine phosphatase (PTP) encoded by the *PTPN11* gene(Chan & Feng, 2006). It was the first phosphatase to be recognized as a bonafide oncogene and plays a crucial role in regulation of signaling events such as cell death pathway (PD-1/PD-L1) and cell growth and differentiation pathway (MAPK)(Matozaki, Murata, Saito, Okazawa, & Ohnishi, 2009). SHP2 is expressed in every tissue and contains two *N*-terminal Src homology-2 (SH2) domains, a C-terminal domain and a PTP domain. Structural studies have revealed that SHP2 adopts an auto inhibited conformation under basal conditions, which associates the PTP and SH2 domains and leads to blockage of the catalytic site(Garcia Fortanet et al., 2016). Activation of SHP2 occurs via binding of phosphoproteins to the SH2 domain following cellular stimulation disrupts this interaction, relieving auto-inhibition and activating cancer dependent phosphatase(X Sun et al., 2018). Mutations in SHP2 that lead to alterations in the catalytic activity have been implicated in Leopard syndrome(LS), Noonan syndrome (NS), juvenile myelomonocytic leukemia (JMML) and several types of cancer and solid tumors(Y.-N. P. Chen et al., 2016; Lawrence et al., 2008). Moreover, SHP2 has been involved in the pathogenesis of gastric carcinoma caused by helicobacter pylori(Hatakeyama, 2004). Mutations in PTPN11 associated with JMML are often close to the N-SH2 domain and disrupt the auto-inhibitory closed conformation, leading to the increase of PTP activity(Xiaojun Sun et al., 2018). Therefore, down-regulating the PTP activity using SHP2 inhibitors represents a promising therapeutic method(Zeng et al., 2014).

The organometallic compound sodium stibogluconate(Pathak & Yi, 2001) (**Figure 1**) is reported as an irreversible inhibitor of SHP2 in vitro, which induces tyrosine phosphorylation rapidly and enhances signaling by various cytokines. However, it is associated with a wide range of serious side effects, including hepato-, cardio-, and phlebo-toxicity. Screening of compounds in the National Cancer Institute (NCI) diversity set chemical library led to the discovery of compound NSC-87877(L. Chen et al., 2006) (**Figure 1**) with an IC_{50} of 0.3 μM against SHP2 in vitro. It shows selectivity over a series of other PTPs, but also inhibits DUSP26. PHPS1 (Hellmuth et al., 2008)(**Figure 1**) is identified by a high-throughput virtual screen against the SHP2 catalytic pocket and inhibits SHP2 with an IC_{50} of 0.7 μM. Screening of a library based on an indole salicylic acid scaffold led to the identification of Compound II-B08(X. Zhang et al., 2010) (**Figure 1**), which inhibits SHP2 with an IC_{50} of 5.5 μM and approximately threefold selectivity over PTP1B and SHP1. Another related irreversible inhibitor of SHP2 is TPI-1(Kundu et al., 2010), which is an analogue of cryptotanshinone and has an IC_{50} of 0.4 μM. But, it is likely to cause toxicity because of generation of reactive oxygen species and the oxidation of the catalytic cysteine residue of SHP2(Butterworth, Overduin, & Barr, 2014). Duan et al.(Duan et al., 2014) identified Comp#1 (**Figure 1**), a benzo[4,5]thiazolo[2,3-c][1,2,4]triazole derivative with an IC_{50} of 4.31 μM against SHP2. Recently, Chen et al.(Y.-N. P. Chen et al., 2016) reported a highly potent (IC_{50}=0.071 μM) SHP2 inhibitor, SHP099, an allosteric modulator that stabilizes the auto-inhibited conformation of SHP2 (**Figure 1**). Comp#2**,** a novel allosteric inhibitor of SHP2 reported by Xie et al.(Xie et al., 2017),

has an IC_{50} of 0.7 μM for SHP2 with more than 30-fold selectivity over other PTPs. RMC-4550(Koltun et al., 2018), a potent and selective SHP2 allosteric inhibitor (IC_{50}=0.538 nM), possesses an overwhelming advantage over SHP099 in terms of in vitro activity. Patrick Sarver et al.(Sarver et al., 2019) reported a screening method that led to the identification of SHP394, an orally efficacious inhibitor of SHP2, with high lipophilic efficiency, improved potency, and enhanced pharmacokinetic properties. Comp#3, a novel potent SHP2 inhibitor containing 3,4,6-trihydroxy-5-oxo-5*H*-benzo[7]annulene reported by Kim et al.(Kim et al., 2020), was obtained as an initial hit with an IC_{50} of 0.097 μM from highthroughput screening (HTS) study. A number of SHP2 inhibitors have been identified in recent years, however, the challenge is the lacking of a sufficiently well-defined SHP2 inhibitor that was approved to go to the market(Garcia Fortanet et al., 2016).

In this research, we designed the novel classes of SHP2 inhibitors by means of virtual screening, molecular docking and de novo design. With the aim to discover some potential SHP2 inhibitors, we evaluated the biological activity of the compounds in the ZINC chemical library that were obtained by virtual screening. Fortunately, compound ZINC169503038 was found to have an IC_{50} value of 4.30 μM against SHP2 in vitro enzyme activity. Encouraged by the promising finding, we further designed a series of ethyl 4-(phenoxymethyl)-2-phenylthiazole-5-carboxylate derivatives by means of scaffold hopping. Then, the target compounds were synthesized and their SHP2 inhibitory activities (IC_{50}) were determined. In order to obtain additional information of the relationship between structure and activity

develop more potent SHP2 inhibitors, common feature pharmacophore studies were performed. Finally, Molecular dynamic simulation study was carried out to explore the affinity of SHP2 and the inhibitor.

2 Materials and Methods

The crystal structure of SHP2 protein was downloaded from PDB Bank. Our calculations were carried out on Dell Precision TM T5500 computer with Discovery Studio 3.5 software package (http://accelrys.com/) and Gromacs 4.5.5.

2.1 Compounds design

2.1 .1 Virtual screening

Virtual screening by molecular docking, using a protein with an experimentally determined structure as a target, has become an established means for lead discovery and for enhancing efficiency in structure optimization(Monsen et al., 2017; J.-W. Wu et al., 2014). In the paper, docking-based virtual screening was carried out through LibDock module (Accelrys Discovery Studio package), which has been applied to the GSK validation data set for its efficiency(Rao, Head, Kulkarni, & LaLonde, 2007).

The receptor crystal structure of SHP2 (PDB ID: 3O5X) was downloaded from the protein data bank (PDB)(Prlić et al., 2016). Using the protocol of "Prepare Protein", the 3O5X was prepared by removing alternate conformations, deleting water and adding the hydrogen atoms. The ligands used for virtual screening were derived from ZINC database and were prepared by "Prepare Ligand" protocol. The binding sites sphere was defined using "Find Site From Current Selection" based on the key residues in the binding pocket, such as P-loop (residues 458-465), the Q loop

(residues 501-507) and the pY recognition loop (residues 277-284) (Zeng et al., 2014). All prepared ligands were docked to the binding pocket of diverse receptor conformations through LibDock module. According to the docking scores, the best candidates were purchased according to the ZINC code and evaluated the biological activity against SHP2 in vitro enzyme activity. The most potent SHP2 inhibitor was identified as potential scaffold used for the following research of scaf1.2 fold hopping.

2.1.2 Replace Fragment

"Replace Fragment" strategy has been used successfully in the design of inhibitors for many pharmaceutically relevant targets(Scott, Coyne, Hudson, & Abell, 2012). The aim of "Replace Fragment" is to discover structurally novel compounds starting from known active compounds by replacement of fragment structure and a lot of important drug discoveries were made by modifying structures of known drugs, such as zopiclone, quinpirole and etoricoxib(Böhm, Flohr, & Stahl, 2004; Wang et al., 2016). In our studies, repalcement was performed by "Replace Fragment" protocol of Discovery Studio 3.5. Firstly, build the fragment library that used for "Replace Fragment". Secondly, choose the possible cores that would be replaced by the prepared fragments. Lastly, replace the original fragment and identify the isosteric molecules with novel structure. When these steps were complete, the molecules were docked into the receptor pockets by "CDOCKER" module and the molecules perform higher binding affinities were selected for the following chemical synthesis(Mizwicki et al., 2012).

2.2 ADMET Prediction

The ADMET (absorption, distribution, metabolism, excretion, and toxicity) properties are a crucial aspect of clinical candidate quality. which were performed using DS v3.5 software to assess the pharmacokinetic properties of the compounds. Some important ADMET descriptors include log P, the aqueous solubility(Cheng & Merz, 2003), human intestinal absorption (HIA)(Egan & Lauri, 2002; Egan, Merz, & Baldwin, 2000),cytochrome P450 (CYP450) 2D6 inhibition(Susnow & Dixon, 2003; Wesson & Eisenberg, 1992), hepatotoxicity(Cheng & Dixon, 2003; Xia, Maliski, Gallant, & Rogers, 2004), blood brain barrier penetration (BBB)(Egan, Walters, & Murcko, 2002) and plasma protein binding(Dixon & Merz, 2001; Votano et al., 2006).

2.3 Chemistry

2.3.1 General

All the reagents were purchased from commercial suppliers and were used without further purification. All the reactions were monitored by thin-layer chromatography (TLC) on silica gel precoated F254 Merck plates, and spots were examined under UV light (254 nm or 365 nm). All column chromatography was performed using 200-300 mesh silica gel. Melting points were measured with an X-6 micro melting point apparatus and are uncorrected. ^{1}H NMR spectra and ^{13}C NMR spectra were recorded on a Bruker AV400 NMR spectrometer with DMSO-d_6 as solvent and known chemical shifts of residual proton signals of deuterated solvents as internal standard. High-resolution mass spectra (HR-MS) were determined with an Agilent Q-TOF 6510 mass spectrometer using direct injection method, and electro-spray ionization (ESI) was used as an ionization technique in either positive (ESI+) or negative (ESI-) mode. The synthetic route to the desired products **1a-1j** is

illustrated in Schemes 1.

Reagents and conditions: (ⅰ). ethyl 2-chloro-3-oxobutanoate, EtOH, reflux; (ⅱ). NBS, BPO, CCl_4, reflux; (ⅲ). 4-substitute phenol, K_2CO_3, DMF, 80℃.

2.3.2 General procedure for the synthesis of compounds **3a-3e**from compounds **2a-2e**

The synthetic procedure of compounds **3a-3e**: To a stirred solution of **2a-2e** (20 mmol) in EtOH (40 mL) and was added ethyl 2-chloro-3-oxobutanoate (20 mmol, 3.29g), the reaction mixture was heated at reflux in N_2 atmosphere until the completion of reaction as indicated by TLC analysis (typically within 5 h).

The reaction mixture was poured into ice-water (100 mL) and the mixture thus obtained was extracted with CH_2Cl_2 (100 mL × 3). The combined organic phases were washed with saturated brine (300 mL), dried over anhydrous Na_2SO_4 and evaporated on a rotary evaporator to afford a residue, which was purified by column chromatography through a short silica gel column to yield **3a-3e**.

Ethyl 2-(2-bromophenyl)-4-methylthiazole-5-carboxylate (**3a**), white solid, 4.37 g (67%), m.p. 93.1-94.6°C. ^{1}H NMR (DMSO-d_6, 400 MHz) δ: 1.32 (t, J = 7.2 Hz, 3H), 2.73 (s, 3H), 4.33 (q, J = 7.2 Hz, 2H), 7.46-8.16 (m, 4H).

Ethyl 2-(4-methoxyphenyl)-4-methylthiazole-5-carboxylate (**3b**), white solid, 3.94 g (71%), m.p. 78.2-78.9°C. ^{1}H NMR (DMSO-d_6, 400 MHz) δ: 1.31 (t, J = 7.2Hz, 3H), 2.68 (s, 3H), 3.84 (s，3H), 4.30 (q, J = 7.2 Hz, 2H), 7.08 (d, J = 8.4Hz, 2H), 7.95 (d, J = 8.4Hz, 2H).

Ethyl 4-methyl-2-(3-phenoxyphenyl)thiazole-5-carboxylate (**3c**), white solid, 4.68 g (69%), m.p. 72.1-73.4°C. ^{1}H NMR (DMSO-d_6, 400 MHz) δ: 1.31 (t, J=7.2Hz, 3H), 2.67 (s, 3H), 4.31 (q, J=7.2Hz, 2H), 7.10-7.52 (m, 9H).

Ethyl 2-(4-fluorophenyl)-4-methylthiazole-5-carboxylate (**3d**), white solid, 3.82 g

(72%), m.p. 88.8-89.9°C. ^{1}H NMR (DMSO-d_6, 400 MHz) δ: 1.30 (t, J=7.2Hz, 3H), 2.88 (s, 3H), 4.30 (q, J=7.2Hz, 2H), 7.60-8.03 (m, 4H). The ^{1}H NMR data were in good agreement with those reported(Sierra et al., 2007).

Ethyl 2-(4-bromophenyl)-4-methylthiazole-5-carboxylate (**3e**), white solid, 4.76 g (73%), m.p. 74.1-75.2°C. ^{1}H NMR (DMSO-d_6, 400 MHz) δ: 1.50 (t, J=7.2Hz, 3H), 2.89 (s, 3H), 4.50 (q, J=7.2Hz, 2H), 7.40-8.10 (m, 4H). The ^{1}H NMR data were in good agreement with those reported(Tiperciuc et al., 2012).

2.3.3 General procedure for the synthesis of compounds **4a-4e** from compounds **3a-3e**

To a stirred solution of **3a-3e** (10 mmol) and NBS (1.96 g, 11 mmol) in CCl_4 (50 mL) was added benzoyl peroxide (0.24 g, 1 mmol), the reaction mixture was heated at reflux in N_2 atmosphere until the completion of reaction as indicated by TLC analysis (typically within 8 h).

On cooling to room temperature, the reaction mixture was poured into ice-water (100 mL) and the mixture thus obtained was extracted with CH_2Cl_2 (50 mL × 3). The combined extracts were washed with 5% brine (100 mL), dried (Na_2SO_4) and evaporated on a rotary evaporator to give a residue, which was purified by column chromatography followed by trituration with n-hexane/EtOAc (1/3 by v/v) to afford **4a-4e**.

Ethyl 4-(bromomethyl)-2-(2-bromophenyl)thiazole-5-carboxylate (**4a**), white solid, 3.16 g (78%), m.p. 127.5-128.4°C. ^{1}H NMR (DMSO-d_6, 400 MHz) δ: 1.34 (t, 3H), 4.38 (q, 2H), 5.05 (s, 2H), 7.52 (m, 1H), 7.60 (m, 1H), 7.88 (m, 1H), 8.17 (m, 1H).

Ethyl 4-(bromomethyl)-2-(4-methoxyphenyl)thiazole-5-carboxylate (**4b**), white solid, 2.88 g (81%), m.p. 92.8-93.8°C. ^{1}H NMR (DMSO-d_6, 400 MHz) δ: 1.34 (t, J = 2.4Hz,3H), 3.86 (s, 3H), 4.35 (q, J=2.4 Hz, 2H), 5.00 (s, 2H), 7.10 (d, J=9.2Hz, 2H), 7.98 (d, J=9.2Hz, 2H).

Ethyl 4-(bromomethyl)-2-(3-phenoxyphenyl)thiazole-5-carboxylate (**4c**), white solid, 3.55 g (85%), m.p. 110.1-111.3°C. ^{1}H NMR (DMSO-d_6, 400 MHz) δ: 1.34(t, J=6.8Hz, 3H), 4.36 (q, J=6.8Hz, 2H), 5.00 (s, 2H), 7.12 (m, 2H), 7.21-7.24 (m, 2H), 7.44-7.48 (m, 2H), 7.55-7.57 (m, 2H), 7.72-7.90 (m, 1H).

Ethyl 4-(bromomethyl)-2-(4-fluorophenyl)thiazole-5-carboxylate (**4d**), white solid, 2.58 g (75%), m.p. 135.8-136.9°C. ^{1}H NMR (DMSO-d_6, 400 MHz) δ: 1.33 (t, 3H), 4.35 (q, J=7.2Hz, 2H), 5.10 (s, 2H), 8.21-8.42 (m, 4H).

Ethyl 4-(bromomethyl)-2-(4-bromophenyl)thiazole-5-carboxylate (**4e**), white solid, 3.32 g (82%), m.p. 141.7-143.0°C. ^{1}H NMR (DMSO-d_6, 400 MHz) δ: 1.36 (t, 3H), 4.32 (q, J=7.2Hz, 2H), 5.08 (s, 2H), 8.30-8.51 (m, 4H).

2.3.4 General procedure for the synthesis of compounds **1a-1i**.

A stirred mixture of **4a-4e** (5 mmol), 4-substitute phenol (6 mmol) and K_2CO_3 (1.38 g, 10 mmol) in DMF (20 mL) was heated at 80°C under N_2 atmosphere until the completion of reaction as indicated by TLC analysis (typically within 10 h).

On cooling to room temperature, the reaction mixture was poured into ice-water (30 mL). The mixture thus obtained was acidified to pH of 7 with concentrated hydrochloric acid and extracted with CH_2Cl_2 (30 mL × 3). The combined extracts were washed with 5% brine (50 mL), dried and evaporated to afford the crude product, which was purified by column chromatography through a short silica gel column followed by trituration with EtOAc/*n*-hexane to yield **1a-1i**.

Ethyl 2-(2-bromophenyl)-4-((4-chlorophenoxy) methyl) thiazole-5-carboxylate (**1a**), white solid, 1.95 g (85%), m.p. 67.5-68.7°C. ^{1}H NMR (DMSO-d_6, 400 MHz) δ: 1.28 (t, J=7.2Hz, 3H), 4.34 (q, J=7.2Hz, 2H), 5.49 (s, 2H), 7.07 (d, J=4.8Hz, 2H), 7.09 (d, J=4.8Hz, 2H), 7.33-7.36 (m, 2H), 7.46-7.50 (m, 2H). ^{13}C NMR (100 MHz, DMSO) δ: 166.23, 160.52, 157.17, 155.88, 134.26, 132.27, 132.01, 131.39, 129.19, 129.19,

128.18, 126.35, 124.67, 121.06, 116.38, 116.38, 63.72, 61.78, 13.94. MS, calcd for $C_{19}H_{15}BrClNO_3S$ ($[M+Na]^+$) 475.9522, found 475.9521.

Ethyl 4-((4-methoxyphenoxy)methyl)-2-(4-methoxyphenyl)thiazole-5-carboxylate (**1b**), white solid, 1.62 g (81%), m.p. 70.8-72.4°C. ^{1}H NMR (DMSO-d_6, 400 MHz) δ: 1.28 (t, J=7.2Hz, 3H), 3.71 (s, 3H), 3.85 (s, 3H), 4.32 (q, J=7.2Hz, 2H), 5.38 (s, 2H), 6.87 (d, J=9.2Hz, 2H), 6.98 (d, J=9.2Hz, 2H), 7.09 (d, J=8.4Hz, 2H), 7.97 (d, J=8.4Hz, 2H). ^{13}C NMR (100 MHz, DMSO) δ: 169.69, 161.93, 160.64, 157.79, 153.57, 152.39, 152.05, 151.09, 128.38, 124.61, 123.84, 115.64, 115.55, 114.64, 114.49, 114.49, 63.91, 61.52, 55.41, 55.23, 13.92. MS, calcd for $C_{21}H_{21}NO_5S$ ($[M+Na]^+$) 422.1038, found 422.1033.

Ethyl 2-(4-methoxyphenyl)-4-((p-tolyloxy)methyl)thiazole-5-carboxylate (**1c**), white solid, 1.69 g (88%), m.p. 97.8-99.2°C. ^{1}H NMR (DMSO-d_6, 400 MHz) δ: 1.26 (t, J=7.2Hz, 3H), 2.24 (s, 3H), 3.84 (s, 3H), 4.31 (q, J=7.2Hz, 2H), 5.39 (s, 2H), 6.94 (m, 2H), 7.09 (m, 4H), 7.97 (m, 2H). ^{13}C NMR (100 MHz, DMSO) δ 169.72, 161.95, 160.66, 157.72, 156.27, 129.79, 129.79, 129.52, 128.41, 128.41, 124.60, 123.89, 114.69, 114.69, 114.37, 114.37, 63.35, 61.55, 55.45, 20.04, 13.95. MS, calcd for $C_{21}H_{21}NO_4S$ ($[M+Na]^+$) 406.1089, found 406.1083.

Ethyl 4-((4-chlorophenoxy)methyl)-2-(4-methoxyphenyl)thiazole-5-carboxylate (**1d**), white solid, 1.60 g (79%), m.p. 95.3-97.0°C. ^{1}H NMR (DMSO-d_6, 400 MHz) δ: 1.26 (t, J=7.2Hz, 3H), 3.84 (s, 3H), 4.31 (q, J=7.2Hz, 2H), 5.44 (s, 2H), 7.08 (m, 4H), 7.35 (m, 2H), 7.95 (m, 2H). ^{13}C NMR (100 MHz, DMSO) δ: 169.81, 161.97, 160.61, 157.22, 157.15, 129.22, 129.22, 128.40, 128.40, 124.61, 124.54, 124.04, 116.34, 116.34, 114.68, 114.68, 63.72, 61.58, 55.44, 13.93. MS, calcd for $C_{20}H1_8ClNO_4S$ ($[M+ Na]^+$) 426.0543, found 426.0536.

Ethyl 4-((4-methoxyphenoxy)methyl)-2-(3-phenoxyphenyl)thiazole-5-carboxylate(**1e**),

white solid, 1.79 g (86%), m.p. 116.7-118.4°C. ^{1}H NMR (DMSO-d_6, 400 MHz) δ: 1.26 (t, J=7.2Hz, 3H), 3.70 (s, 3H), 4.30 (q, J=7.2Hz, 2H), 5.36 (s, 2H), 6.84-6.96 (m, 4H), 7.10-7.23 (m, 4H), 7.42-7.46 (m, 2H), 7.52-7.58 (m, 2H), 7.75-7.77 (m, 1H). ^{13}C NMR (100 MHz, DMSO) δ: 168.77, 160.50, 157.87, 157.63, 155.84, 153.60, 152.28, 133.56, 131.30, 130.25, 130.25, 125.38, 124.18, 121.75, 121.43, 119.23, 119.23, 115.66, 115.66, 115.43, 114.50, 114.50, 63.90, 61.76, 55.27, 13.94. MS, calcd for $C_{26}H_{23}NO_5S$ ([M+ Na]$^+$) 484.1195, found 484.1188.

Ethyl 2-(4-fluorophenyl)-4-((4-methoxyphenoxy)methyl)thiazole-5-carboxylate (**1f**), white solid, 1.61 g (83%), m.p. 73.5-74.5°C. ^{1}H NMR (DMSO-d_6, 400 MHz) δ: 1.26 (t, J=7.2Hz, 3H), 3.70 (s, 3H), 4.32 (q, J=7.2Hz, 2H), 5.39 (s, 2H), 6.86-6.99 (m, 4H), 7.36-7.41 (m, 2H), 8.07-8.10 (m, 2H). ^{13}C NMR (100 MHz, DMSO) δ: 168.46, 165.19, 162.70, 159.13 (d, J_{C-F}= 269.0 Hz), 153.60, 152.35, 129.03 (2C, d, J_{C-F} = 9.0 Hz), 128.48 (d, J_{C-F} = 3.0 Hz), 125.02, 116.34 (2C, d, J_{C-F} = 22.3 Hz), 115.56, 115.56, 114.46, 114.46, 63.88, 61.63, 55.19, 13.85. MS, calcd for $C_{20}H_{18}FNO_4S$ ([M+ Na]$^+$) 410.0838, found 410.0835.

Ethyl 4-((4-chlorophenoxy)methyl)-2-(4-fluorophenyl)thiazole-5-carboxylate (**1g**), white solid, 1.66 g (85%), m.p. 99.0-100.6°C. ^{1}H NMR (DMSO-d_6, 400 MHz) δ: 1.27 (t, J=7.2Hz, 3H), 4.32 (q, J=7.2Hz, 2H), 5.46 (s, 2H), 7.07-7.09 (m, 2H), 7.34-7.41(m, 4H), 8.06-8.07 (m, 2H). ^{13}C NMR (100 MHz, DMSO) δ: 168.62, 163.98 (d, J_{C-F} = 250.3 Hz), 160.46, 157.17, 157.17, 129.22, 129.22, 129.12 (2C, d, J_{C-F} = 9.0 Hz), 128.45 (d, J_{C-F} = 3.0 Hz), 125.28, 124.66, 116.45 (2C,d, J_{C-F} = 20.9 Hz), 116.34, 116.34, 63.69, 61.74, 13.90. MS, calcd for $C_{19}H_{15}ClFNO_3S$ ([M+ Na]$^+$) 414.0343,

found 414.0337.

Ethyl 2-(4-bromophenyl)-4-((p-tolyloxy)methyl)thiazole-5-carboxylate (**1h**), white solid, 1.90 g (88%), m.p. 91.9-92.8°C. ^{1}H NMR (DMSO-d_6, 400 MHz) δ: 1.27 (t, J=7.2Hz, 3H), 2.24 (s, 3H), 4.32 (q, J=7.2Hz, 2H), 5.40 (s, 2H), 6.92 (d, J=8.4Hz, 2H), 7.10 (d, J=8.4Hz, 2H), 7.75 (d, J=8.4Hz, 2H), 7.96 (d, J=8.4Hz, 2H). ^{13}C NMR (100 MHz, DMSO) δ: 168.45, 160.48, 157.81, 156.21, 132.36, 132.36, 130.96, 129.81, 129.81, 129.59, 128.49, 128.49, 125.44, 125.17, 114.39, 114.39, 63.33, 61.78, 20.05, 13.93. MS, calcd for $C_{20}H_{18}BrNO_3S$ ([M+ Na]$^+$) 454.0088, found 454.0083.

Ethyl 2-(4-bromophenyl)-4-((4-chlorophenoxy)methyl)thiazole-5-carboxylate (**1i**), white solid, 1.81 g (80%), m.p. 93.6-95.1°C. ^{1}H NMR (DMSO-d_6, 400 MHz) δ: 1.27 (t, J=7.2Hz, 3H), 4.32 (q, J=7.2Hz, 2H), 5.45 (s, 2H), 7.08 (d, J=8.8Hz, 2H), 7.35 (d, J=8.8Hz, 2H), 7.74 (d, J=8.4Hz, 2H), 7.95 (d, J=8.4Hz, 2H). ^{13}C NMR (100 MHz, DMSO) δ: 166.30, 160.54, 157.17, 155.94, 134.29, 132.36, 132.05, 131.44, 129.22, 129.22, 128.26, 126.40, 124.65, 121.07, 116.41, 116.41, 63.73, 61.81, 13.95. MS, calcd for $C_{19}H_{15}BrClNO_3S$ ([M+ Na]$^+$) 473.9542, found 473.9537.

2.4 Enzymatic assays

Human recombinant SHP2 (PTP-domain), LAR, SHP1, MEG2, CDC25B or TCPTP was expressed in E coli and purifed by Ni-NTA affinity chromotagraphy in our laboratory. The phosphatase activity was assayed using p-nitrophenyl phosphate (pNPP) as a substrate at 37 °C in 60 μL buffer containing 2 mM EDTA, 2 mM dithiothreitol and 60 mM calcium citrate. The sample was screened in a 96-well format at 20 μM compound concentration. The reaction was initiated by the addition

of 0.1μg enzyme. After preincubation for 2 min at 37°C, 50 μL of buffer with 40 μL pNPP was added and incubated. The reaction was quenched by adding 10μL of 0.2 N NaOH. The amount of product p-nitrophenol was determined from the absorbance at 405 nm detected by a microplate spectrophotometer. The experiments were performed in triplicate and inhibitory activity was expressed as half maximal inhibitory concentration IC_{50}(J. Wu, Sun, Zhou, Ma, & Wang, 2019).

2.5 Pharmacophore generation

The 3D-common feature pharmacophore hypothesis of these SHP2 inhibitors was identified by applying the HipHop tool embedded in the Accelrys Discovery Studio 3.5 software (http://accelrys.com/) to analyze the structure-activity relationships (SAR) of SHP2 protein-ligand interactions.

At first, the 3D structures of chemical compounds were optimized with CHARMm force field(Brooks et al., 2009) by applying the Full Minimization module embedded in the Accelrys Discovery Studio 3.5 software. Based on the bioactivity of the compounds, the molecular property Principal to the molecules was set as a value. “2” meant that the molecular was active; “1” meant that the molecular was moderately active; “0” meant that the molecular was inactive. MaxOmitFeat was used to indicate how many features were allowed to miss for each molecule. Then, Molecular Overlay was used to automatically align the conformational for each compound by 50% steric filed and 50% electrostatic filed. In addition, the Input ligands conformation generation option was set to “BEST”, the Maximum Conformations was set to 200, the Energy Threshold was set to 10, and the other parameters were kept default values.

With the aid of Edit and Cluster Features tool, the list of pharmacophore feature types for the compounds was represented. Four features, namely hydrogen bond acceptors (HBA), hydrogen bond doners (HBD), Ring_Aromatic (RA) and hydrophobe area (HYP), were used to build the SHP2 inhibitors hypotheses. The minimum numbers of HBA, HBD, RA and HYP were set to 1 and the maximum numbers of them were set to 2. Moreover, pharmacophore model computation was initialized using the HipHop algorithm. Finally, pharmacophoric hypothesis was generated from these aligned structures.

2.6 Molecular Dynamics Simulations

The mechanism of how the compound inhibited the SHP2 protein at molecular level were studied by applying the "GROMACS 4.5.5 package(Loeffler & Winn, 2012; Pandey, Roy, & Doerksen, 2020)"with AMBER99B force field(Smith, Rao, Segelken, & Cruz, 2015) for 20 ns. The topology files were firstly generated. The protein topology file was produced automatically and the ligand topology file were generated by utilizing ACPYPE Server(W Sousa da Silva & Vranken, 2012) with the atom type setting as default GAFF(Generalised Amber Force Field). Secondly, the box and solvate was defined(Sadeghi-Kaji, Shareghi, Saboury, & Farhadian, 2020; H. Zhang et al., 2020). The dodecahedron box was defined as the unit cell and was filled with explicit single-point charge (SPC) water molecules, which were more than 3000 and less than 4000. Thirdly, the Na^+ /Cl^- ions were added to neutralize the system by replacing some solvent molecules randomly. Fourthly, the system was minimized by using the steepest decent method. The minimization was stopped when the maximum

force < 10.0 kJ/mol.The system was then heated to 310 K during a 200 ps NVT simulation with 1 fs time step. The pressure was then equilibrated to 1 atm during a 500 ps NPT simulation with 2 fs time step. In both simulations, all heavy atoms were position restrained with the force constant of 1000 kJ/(mol nm^2)(Mohammadi, Shareghi, Akbar Saboury, & Farhadian, 2020). The position restraint was gradually removed during a 1 ns simulation with a time step of 2 fs. After that, molecular dynamics study was initiated by using the Particle Mesh Ewald algorithm (PME)(Nocito & Beran, 2018) and (Uusitalo, Ingólfsson, Marrink, & Faustino, 2017) with a time step of 2 fs. The short-range electrostatic cutoff and the short-range van der Waals cutoff were 1.0 nm, respectively(Maroli & Kolandaivel, 2020).

Free energy were calculated on the MD trajectory using g_mmpbsa(Kumari, Kumar, Open Source Drug Discovery, & Lynn, 2014). The calculation formula is as follows:(Saiz-Urra, Cabrera, & Froeyen, 2011):

$$\Delta G_{bind} = G_{complex} - G_{receptor} - G_{ligand} \quad (1)$$

$$\Delta G_{bind} = E_{gas} + G_{sol} - T\Delta S \quad (2)$$

$$E_{gas} = E_{int} + E_{vdw} + E_{ele} \quad (3)$$

$$G_{sol} = G_{GB} + G_{SA} \quad (4)$$

$$G_{SA} = \gamma_{SASA} \quad (5)$$

where T is for temperature and S is for the total solute entropy; E_{gas} signifies gas-phase energy, which is the sum of internal energy (E_{int}), electrostatic (E_{ele}), and van der Waals contributions (E_{vdw}). The solvation free energy (G_{sol}) can be further decomposed into polar and nonpolar solvation states. The polar solvation contribution

(G_{GB} and G_{PB}) is determined by solving Poisson Boltzmann (PB) and Generalized Borne (GB) equations. The nonpolar solvation contribution (G_{SA}) is estimated using 0.0072 kcal $mol^{-1}A^{-2}$ as the value for constant γ and the solvent-accessible surface area (SASA) determined using a water probe radius of 1.4 Å. Dielectric constant values for solute and solvent were set to 1 and 80, respectively(Mehla & Ramana, 2016).

3. Results and discussion

3.1 Compounds design

As shown in **Figure 2**. The high-throughput virtual screen of approximately 3 million compounds in ZINC database against the SHP2 catalytic domain led to the identification of several candidates. These compounds were then purchased from reagent companies and evaluated the biological activity against SHP2. Fortunately, we discovered a potent SHP2 inhibitor, ZINC169503038, which inhibits SHP2 with an IC_{50} of 4.30 μM. In the paper, scaffold hopping was performed by "Replace Fragment" protocol of Discovery Studio 3.5. The first step was to replace scaffold A (colored in red, **Figure 2**), which generated 5 scaffolds. The second step was to replace scaffold B (colored in blue, **Figure 2**), which generated 4 scaffolds. Therefore, we have a total of 20 different compounds for the ZINC169503038 derivatives thus generated. These compounds were then re-docked into the binding pocket of diverse receptor (3O5X) conformations through "CDOCKER" module. The "CDOCKER interaction energy" was an important index for evaluating the binding affinities. The docking score of top

10 compounds (**1a-1i**) and ZINC169503038 were listed in **Table 1**, the compound **1d** (- CDOCKER interaction energy = 50.82 Kcal/moL) had the higher docking score than ZINC169503038 (CDOCKER interaction energy = 30.95 Kcal/moL).

The two-dimensional (2D) diagram of compound **1d-**3O5X interaction was shown in **Figure** 3. Pink boxes such as Tyr279, Lys366, Trp423, Gly427, Arg465 and Gln510 represented hydrogen bonding, charge or polar interactions, while green boxes like Val428, Cys459, Gly464, Ala509, His426, Thr507, Arg362, Lys364, Ser460, Gln506, Ala461and Thr357 were involved in van der waals interactions. The 4-chlorophenyl of compound **1d** formed cation-Pi interaction with Lys366. Moreover, the N and O atoms of compound **1d** formed seven potential H-bonding with the side chains of Glu510, Trp423 Arg465 and Lys366 (see blue dashed arrow directed towards the electron donor in **Figure** 3). In short, compound **1d** not only formed the cation-Pi interaction but also more hydrogen bonds with the SHP2 protein, resulting in the strong binding affinities.

Furthermore, the binding modes of 3O5X and compound **1d** obtained from the docking simulation was shown in **Figure 4C**, it was obvious that the key residues (Arg362, Lys364, Lys366, His426, Arg465 and Gln510) for the binding interactions between the ligand and the receptor were fully consistent with the previous results. As shown in **Figure 4C**, the three O atoms of compound **1d** forms six hydrogen bonds with Gln510, Arg465 and Lys364 while the N atom of compound **1d** interacts with residue Gln510 via a strong H-bond. These H-bonds significantly enhance the

inhibitor-receptor interaction and stabilize the ethyl 4-(phenoxymethyl)-2-phenylthiazole-5-carboxylate derivatives complex within the active site to a great extent.

3.2 ADMET Pridiction

There are two important factors that affect oral bioavailability: one is Human Intestinal Absorption (HIA), the other is solubility. Solubility plays a key role in the pharmacological activity of a compound. HIA has a pronounced effect on the therapeutic effect of drugs. All compounds (**1a-1i**) showed good human intestinal absorption due to appropriate aqueous solubility and LogP. Plasma protein plays an important role in drug distribution. All compounds (**1a-1i**) were found to be highly bound with plasma protein. CYP2D6 is one of the important enzymes involved in drug metabolism, which is associated with the metabolism and elimination of approximately 25% of clinically used drugs. Nine compounds were predicted to be non-inhibitors of cytochrome P450 2D6 (CYP2D6). For hepatotoxicity, all compounds were predicted non-toxic. For brain/blood barrier (BBB), compound **1g**, **1h** and **1i** had a very good penetrant level, compound **1a**, **1b**, **1c** and **1e** had a good penetrant level, while compound **1d** and **1f** showed a moderate penetrant level. Therefore, as mentioned above, the values for the ADMET properties of all compounds listed in **Table 1** are within the acceptable range for human beings, indicating these compounds found in this study can be utilized as candidates for the purpose of developing new drugs.

3.3 Chemistry

Firstly, the synthetic route to desired products **1a-1i** was shown in **Scheme 1**. Cyclization of benzothioamide derivatives (**2a-2e**) with ethyl 2-chloro-3-oxobutanoate in EtOH at reflux smoothly afforded **3a-3e** according to known procedures. Secondly, bromination of ethyl 4-methyl-2-phenylthiazole-5- carboxylate derivatives (**3a-3e**) with N-Bromosuccinimide in the presence of benzoyl peroxid as catalyst in carbon tetrachloride at reflux afforded **4a-4e**. Thirdly, O-alkylation of ethyl 4-(bromomethyl) -2-phenylthiazole-5-carboxylate derivatives with substituted phenol in the presence of K_2CO_3 in DMF at 80°C produced the final desired products **1a-1i**(J.-w. Wu et al., 2019).

3.4 *In vitro* SHP2 inhibitory activity

The results of *in vitro* inhibitory assay of 10 synthesized compounds (**1a-1i**) as well as **comp#1** as positive control against SHP2 were summarized in **Table 2**. Comp#1 was identified by Duan et al. against the SHP2 catalytic pocket and inhibits SHP2 with an IC_{50} of 4.31 μM (**Figure 1**). Most of the synthesized compounds (**1a-1i**) were more active than the positive control except compound **1a** and **1g**. It was obvious that half of the desired compounds (**1b-1e**) showed improved SHP2 inhibitory activity (IC_{50} values were range from 0.99 μM to 3.90 μM) as compared with ZINC169503038 (IC_{50}=4.30 μM). The rest of the desired compounds (**1f-1i**) were only slightly less active (IC_{50} values were range from 4.60 μM to 11.48 μM) than ZINC169503038 except **1a**, displayed an IC_{50} of 35.56 μ M, which is 8-fold less potent than ZINC169503038. As shown in **Table 2**, compounds **1b-1d** showed higher and similar inhibitory activity (IC_{50} values were range from 0.99 μM to 1.25 μM), leading to the

identification of 4-methoxyphenyl as the optimal scaffold A. It should be noted that the compounds **1a** (scaffold A= 2-bromophenyl, IC_{50}=35.56 μM) and **1f-1i** (scaffold A= 4-bromophenyl or 4-fluorophenyl, IC_{50}=4.6-11.48 μM) with halogen substituted benzene exhibited lower IC_{50} values than ZINC169503038. These observations strongly indicated that the scaffold A is very important t for the significant impacts on inhibitory activity. Among the four scaffold B groups, 4-methoxyphenyl seemed the worse one (**1g** < **1f**, **1i** < **1h**).

As shown in **Table 3**, we also explored the selectivity of compound **1a-1i** for other phosphatases, such as src homologous phosphatase-1 (SHP1), homogeneous T cell protein tyrosine phosphatase (TCPTP), maternal-effect germ-cell defective 2 (MEG2), cell division cycle 25 homolog B (CDC25B) and leukocyte antigen-related phosphatase (LAR). It could be seen from the **Table 3** that all compounds (**1a-1i**) seem to be inactive (IC_{50} > 100 μM) against SHP1, MEG2, CDC25B and LAR. Compound **1c** and **1d** showed strong selectivity (more than 33-fold) for SHP2 over TCPTP, compound **1f**, **1g** and **1h** showed moderate selectivity (more than 3-fold) while compound **1i** showed poor selectivity (less than 2-fold). In short, most of compounds (**1a-1h**) have some selective inhibition for SHP2 over SHP1, TCPTP, MEG2, CDC25B and LAR. However, compound **1i** shows no selectivity for SHP2 over TCPTP.

In summary, systematic SAR exploration of a ZINC169503038-based hit led to the discovery of ethyl 4-((4-chlorophenoxy)methyl)-2-(4-methoxyphenyl)thiazole -5-carboxylate (**1d**) as highly active SHP2 inhibitors, which was 4-fold more active

than parent ZINC169503038 in SHP2 inhibitory assay (IC_{50} values for **1d** was 0.96 0.99 μM, against SHP2 vs 4.30μM for ZINC169503038). The hypothesis was tested in the following 3D-QSAR study.

3.5 Pharmacophore hypotheses generated by the HipHop method

The 3D-common feature pharmacophore hypothesis of these SHP2 inhibitors was identified by applying the HipHop tool to analyze the SAR of SHP2 protein-ligand interactions. The reliability of the generated pharmacophore model was assessed by aligning the docked pose of the reference compound **1d** to the conformation generated in Hiphop. Comparing the two conformations with compound **1d** in the top ten hypotheses and SHP2, the RMSD was found to be 0.38 Å, 0.67 Å, 0.93Å, 1.24 Å, 1.57 Å, 1.89 Å, 1.91 Å, 2.10 Å, 2.10 Å, 2.46 Å, respectively. As shown in **Figure 4**, comparing the conformations of compound **1d** in the Hypo-1 and in the active sites of SHP2, the RMSD was only 0.38 Å, indicating that the Hypo-1 was a highly reliable pharmacophore model of SHP2 inhibitors. Thus, the top one hypothesis-Hypo-1 had the least RMSD value and possessed the highest ranking score was chosen for the further study. The generated 3D common feature pharmacophore hypothesis had two HBAs, one RA, two HYPs. Fit value could provide information in whether the molecule could fit into the pharmacophore hypothesis. As shown in **Table 4**, the fit value of the most active compound **1d** was 4.96, whereas, the least active compound 1a had the fit value of 1.60. The most active compound **1d** with IC_{50} 0.99 μM and least active compound **1a** with IC_{50} 35.56 μM were aligned with Hypo-1 to verify whether the hypothesis could distinguish the active compounds from the inactive

compounds. It could be seen from **Figure 5** the most active compound **1d** mapped all the features in Hypo-1, whereas the least active compound **1a** only mapped the two HBA and RA features and failed to map the two HYP features. In order to verify the reliability of the Hypo-1, we compared the pharmcophore model with the active site of SHP2. The two HBAs were pointed to the Arg465 and Gln 510, respectively, and the RA was pointed to the Lys366, while one HYP was oriented towards the His426, and the other HYP was pointed to the Arg362 and Lys364. Thus, the chemical features of Hypo-1 could interact with key residues in the active site of SHP2. The generated pharmacophore model revealed that the two HBAs, one RA, and two HYPs features might play an important role in binding to the active site of SHP2. According to these, two HBA atoms or groups at the specific positions were the essentials in the molecule to bind to the SHP2 protein, while bulky aromatic moieties with appropriate active shape in the molecule were the necessary to fit into the one RA and two HYPs features.

The results of the compounds mapping onto features were given in **Table 4**. the **1b**, **1c**, and **1d** compounds were discovered as the most 4-(phenoxymethyl)-2-phenylthiazole-5-carboxylate derivatives with the best fit value of 4.80, 4.95, and 4.96, respectively, for the generated pharmacophore hypothesis, fitting to all the mapped common features in the anticipated model with a specified "Pharmprint" value of "11111". The fit values and "Pharmprint" values of these SHP2 inhibitors based on the pharmacophoric features in the Hypo-1 model were given in **Table 4.** The compounds having 4-methoxy group at the benzene ring in scaffold A,

for example compound **1b**, compound **1c** and compound **1d** rather than holding a 4-nitro, 3-phenoxy, and 4-bromol group at the benzene ring in scaffold A, such as compond **1e** and compound **1h** possessed a higher fit values, meaning that they possessed a better match with all the features in the pharmacophoric features. This explains why compounds with 4-methoxy group at the benzene ring in scaffold A, such as compound **1b**, compound **1c** and compound **1d** had more significant bioactivity than compounds holding a 4-nitro, 3-phenoxy, and 4-bromol group at the benzene ring in scaffold A. Owing to matching with all the features except the HYP features, the compounds with 2-bromol and 4-fluoro group at the benzene ring in scaffold A showed a lower fit value than compounds 4-methoxy, 4-nitro, 3-phenoxy, and 4-bromol group at the benzene ring in scaffold A, such as compound **1a**,compound **1f** and compound **1g**, which was accordance with the SHP2 inhibitory activity results. This observation indicated why structure with the 4-methoxy group at the benzene ring in scaffold A was more favourable than the other group, including 2-bromol,4-nitro, 3-phenoxy, 4-fluoro and 4-bromol groups, for increasing potency in this set of compounds.

3.6 Molecular Dynamics Simulations

To explore the mechanism of how compound **1d** inhibit SHP2, compound **1d** in complex with SHP2 and SHP2 without ligand were subjected to 20 ns MD simulations, respectively. As shown in **Figure 6**, the RMSD values of SHP2-compound **1d** and SHP2 without ligand systems were stable from 2 ns to 20 ns. The RMSD value was a key criterion applied to evaluate the stability of the

protein-complex and the protein systems. During the simulation, the RMSD values of SHP2 with ligand compound **1d** and without ligand systems were found to be relatively stable about 0.21 Å and 0.26 Å, respectively. As we could see from the **Figure 6**, the RMSD value of SHP2 with ligand compound **1d** was smaller than that of SHP2, indicating that the complexes of SHP2 with ligand compound **1d** was more stable than SHP2. Thus, the compound **1d** could make the SHP2 protein stable.

As shown in panel B of **Figure 6**, the root mean square fluctuations for all the side-chain atoms of the receptors were also computed within 20 ns molecular dynamics simulations, to perform an in-depth study of the interactions of the SHP2 binding domain with the inhibitor. The larger the RMSF fluctuated, the more unstable the residues were, whereas the smaller the RMSF fluctuated, the more stable the residues were. In the active site-pTyr recognition loop (residues 271-284), the WPD-loop (residues 421-431) and the Q-loop ((residues 501-507), the order of the RMSF values was 3O5X-compound **1d** system < 3O5X without ligand system. The results suggested that the compound **1d** had effect on the three regions (pTyr recognition loop, WPD-loop and Q-loop) with decreasing the interactions of the residues.

To provide further understanding of compound **1d** in inhibiting the fluctuation of loops of SHP2 (**Figure 7**), the Anisotropic Network Mode (ANM) was performed by the VMD software(Kotamarthi, Yadav, & Ainavarapu, 2015; Sarkar, 2018) The arrow represented the direction and amplitude of the motion. The motions of the pTyr

recognition loop (residues residues 271-284 colored in red), the WPD-loop (residues 421-431 colored in blue) and the Q-loop (residues 501-507 colored in purple) in the SHP2-compound **1d** system were all smaller than that in the SHP2 without ligand system, showing that compound7d could inhibit the fluctuations of the pTyr recognition loop, the WPD-loop and the Q-loop. So, the ligand (compound **1d**) binding to the receptor (3O5X) might increase the stability of the residues in the active site, causing the inhibition activities. The analysis of ANM was in accordance with RMSF analysis.

For better understanding of the specific interactions between protein systems, the binding free energy was calculated using MM/PBSA method, which is consisting of four terms, including the polar solvation free energy, the non-polar solvation free energy, the van der waals interaction energy and the electrostatic energy. As we could see from the **Table 5**, The binding energy of compound **1d** with SHP2 is -76.737 kcal/mol, while the binding energy of compound **1a** with SHP2 is -66.949 kcal/mol, indicating that the system of SHP2-compound **1d** seems more stable than SHP2-compound **1a**, which was in accord with the bioactivity results. Polar solvation free energy made a positive contribution to the total free binding energy while non-polar solvation free energy, electrostatic interactions and Van der Waals made a negative contribution to the total interaction energy, which meant that Van der Waals, electrostatic interactions and non-polar solvation energy together are in favor of the stability of the complex system. We could clearly see that the Van der Waal

interactions and the non-polar solvation free energy were the major favorable contributions to the binding energies between SHP2 and compound **1d/1a**.

For the purpose of discerning the contributions of binding free energy of each residue, the interaction energies are decomposed into individual residue contributions using energy decomposition scheme(Baker, Sept, Joseph, Holst, & McCammon, 2001). The residues with the top 10 interaction energies for compound **1d** with SHP2, respectively, are shown in **Figure 8**. The binding energy is defined as the sum of van der Waals interactions and short range electrostatic interactions between the ligand and the receptor. The key residues were found in three significantly different regions: the active site-pTyr recognition loop (residues 271-284), WPD-loop (residues 421-431), P-loop (residues 501-507), which was in accordance with the RMSF results in **Figure 5**. As shown in **Figure 8**, Ile282, Pro284, Tyr279 and Phe285 in pTyr recognition loop made large contributions to the binding affinity. Their total energy contributions are -6.20 kcal/mol, -3.07 kcal/mol, -1.99 kcal/mol and -1.69 kcal/mol, respectively. Gly503, Ser502 and Met504 in Q-loop makes remarkable contributions to the binding affinity, their total energy contributions are -4.29 kcal/mol, -2.52 kcal/mol and -1.65 kcal/mol, respectively. In a word, all of these are very important reasons to explain why compound **1d** showed good activities.

4 Conclusion

In conclusion, a series of novel ethyl 4-(phenoxymethyl)-2-phenylthiazole-5-carboxylate derivatives targeting SHP2 were designed with the aid of scaffold hopping approach. The designed compounds were synthesized, and their activities

against SHP2 were evaluated. The results showed that the ethyl 4-(phenoxymethyl)-2-phenylthiazole-5-carboxylate derivatives inhibited SHP2 within the micromolar range (0.99 μM ~ 35.56 μM). Compound **1d**, displaying the best inhibitory activity with an IC_{50} value of 0.99 μM. The 3D-common feature pharmacophore hypothesis of these SHP2 inhibitors was identified by applying the HipHop tool to analyze the SAR of SHP2 protein-ligand interactions. Furthermore, compound **1d** was validated by the method of molecular dynamics simulations. In this paper, the various post-dynamics analyses were used to explore the effect of the compound **1d** on the SHP2, such as the RMSD, RMSF, ANM and binding energy analyses. The results suggested that the compound **1d** could make the SHP2 protein stable, that's why compound **1d** showed good activities.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (Grant No. 81773569), the Natural Science Foundation of Tianjin (Grant No. 18JCQNJC13700), the Cooperation and Exchange Project of the National Natural Science Foundation of China (Grant No.81611130090), the Science & Technology Development Fund of Tianjin Education Commission for Higher Education (Grant No. 2017KJ229).

Conflict of interest

There is no conflict of interest.

Rererence

Böhm, H.-J., Flohr, A., & Stahl, M. (2004). Scaffold hopping. *Drug discovery today: Technologies, 1*(3), 217-224

Baker, N. A., Sept, D., Joseph, S., Holst, M. J., & McCammon, J. A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A, 98*(18), 10037-10041. https://doi.org/10.1073/pnas.181342398

Brooks, B. R., Brooks III, C. L., Mackerell Jr, A. D., Nilsson, L., Petrella, R. J., Roux, B., . . . Boresch, S. (2009). CHARMM: the biomolecular simulation program. *Journal of computational chemistry, 30*(10), 1545-1614

Butterworth, S., Overduin, M., & Barr, A. J. (2014). Targeting protein tyrosine phosphatase SHP2 for therapeutic intervention. *Future medicinal chemistry, 6*(12), 1423-1437

Chan, R. J., & Feng, G.-S. (2006). PTPN11 is the first identified proto-oncogene that encodes a tyrosine phosphatase. *Blood, 109*(3), 862-867

Chen, L., Sung, S.-S., Yip, M. R., Lawrence, H. R., Ren, Y., Guida, W. C., . . . Wu, J. (2006). Discovery of a novel shp2 protein tyrosine phosphatase inhibitor. *Molecular pharmacology, 70*(2), 562-570

Chen, Y.-N. P., LaMarche, M. J., Chan, H. M., Fekkes, P., Garcia-Fortanet, J., Acker, M. G., . . . Cooke, V. G. (2016). Allosteric inhibition of SHP2 phosphatase inhibits cancers driven by receptor tyrosine kinases. *Nature, 535*(7610), 148

Cheng, A., & Dixon, S. L. (2003). In silico models for the prediction of dose-dependent human hepatotoxicity. *J Comput Aided Mol Des, 17*(12), 811-823

Cheng, A., & Merz, K. M., Jr. (2003). Prediction of aqueous solubility of a diverse set of compounds using quantitative structure-property relationships. *J Med Chem, 46*(17), 3572-3580. https://doi.org/10.1021/jm020266b

Dixon, S. L., & Merz, K. M., Jr. (2001). One-dimensional molecular representations and similarity calculations: methodology and validation. *J Med Chem, 44*(23), 3795-3809. https://doi.org/10.1021/jm010137f

Duan, Y.-Q., Ma, Y., Wang, X.-J., Jin, Y.-Y., Wang, R.-L., Dong, W.-L., . . . Wang, S.-Q. (2014). Design potential selective inhibitors for treating cancer by targeting the Src homology 2

(SH2) domain-containing phosphatase 2 (Shp2) with core hopping approach. *Protein and peptide letters, 21*(6), 556-563

Egan, W. J., & Lauri, G. (2002). Prediction of intestinal permeability. *Adv Drug Deliv Rev, 54*(3), 273-289. https://doi.org/10.1016/S0169-409X(02)00004-2

Egan, W. J., Merz, K. M., Jr., & Baldwin, J. J. (2000). Prediction of drug absorption using multivariate statistics. *J Med Chem, 43*(21), 3867-3877. doi: 10.1021/jm000292e

Egan, W. J., Walters, W. P., & Murcko, M. A. (2002). Guiding molecules towards drug-likeness. *Curr Opin Drug Discov Devel, 5*(4), 540-549

Garcia Fortanet, J., Chen, C. H.-T., Chen, Y.-N. P., Chen, Z., Deng, Z., Firestone, B., . . . Fridrich, C. (2016). Allosteric inhibition of SHP2: identification of a potent, selective, and orally efficacious phosphatase inhibitor. *Journal of medicinal chemistry, 59*(17), 7773-7782

Hatakeyama, M. (2004). Oncogenic mechanisms of the Helicobacter pylori CagA protein. *Nature Reviews Cancer, 4*(9), 688

Hellmuth, K., Grosskopf, S., Lum, C. T., Würtele, M., Röder, N., von Kries, J. P., . . . Birchmeier, W. (2008). Specific inhibitors of the protein tyrosine phosphatase Shp2 identified by high-throughput docking. *Proceedings of the National Academy of Sciences, 105*(20), 7275-7280

Hendriks, W., Bourgonje, A., Leenders, W., & Pulido, R. (2018). Proteinaceous regulators and inhibitors of protein tyrosine phosphatases. *Molecules, 23*(2), 395

Khan, I., Rhett, J. M., & O'Bryan, J. P. (2019). Therapeutic targeting of RAS: New hope for drugging the "undruggable". *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 118570

Kim, B., Jo, S., Park, S. B., Chae, C. H., Lee, K., Koh, B., & Shin, I. (2020). Development and structure-activity relationship study of SHP2 inhibitor containing 3, 4, 6-trihydroxy-5-oxo-5H-benzo [7] annulene. *Bioorganic & Medicinal Chemistry Letters, 30*(1), 126756

Koltun, E. S., Aay, N., Buckl, A., Jogalekar, A. S., Kiss, G., Marquez, A., . . . Semko, C. M. (2018). RMC-4550, an allosteric inhibitor of SHP2: Synthesis, structure, and anti-tumor activity: AACR.

Kotamarthi, H. C., Yadav, A., & Ainavarapu, S. R. K. (2015). Small peptide binding stiffens the

ubiquitin-like protein SUMO1. *Biophysical journal, 108*(2), 360-367

Kumari, R., Kumar, R., Open Source Drug Discovery, C., & Lynn, A. (2014). g_mmpbsa--a GROMACS tool for high-throughput MM-PBSA calculations. *J Chem Inf Model, 54*(7), 1951-1962. https://doi.org/10.1021/ci500020m

Kundu, S., Fan, K., Cao, M., Lindner, D. J., Zhao, Z. J., Borden, E., & Yi, T. (2010). Novel SHP-1 inhibitors tyrosine phosphatase inhibitor-1 and analogs with preclinical anti-tumor activities as tolerated oral agents. *The Journal of Immunology, 184*(11), 6529-6536

Lawrence, H. R., Pireddu, R., Chen, L., Luo, Y., Sung, S.-S., Szymanski, A. M., . . . Wu, J. (2008). Inhibitors of Src homology-2 domain containing protein tyrosine phosphatase-2 (Shp2) based on oxindole scaffolds. *Journal of medicinal chemistry, 51*(16), 4948-4956

Loeffler, H. H., & Winn, M. (2012). *Large biomolecular simulation on hpc platforms III. AMBER, CHARMM, GROMACS, LAMMPS and NAMD*.

Maroli, N., & Kolandaivel, P. (2020). Comparative study of stability and transport of molecules through cyclic peptide nanotube and aquaporin: A molecular dynamics simulation approach. *Journal of Biomolecular Structure and Dynamics, 38*(1), 186-199. https://doi.org/10.1080/07391102.2019.1570341

Matozaki, T., Murata, Y., Saito, Y., Okazawa, H., & Ohnishi, H. (2009). Protein tyrosine phosphatase SHP-2: a proto-oncogene product that promotes Ras activation. *Cancer science, 100*(10), 1786-1793

Mehla, K., & Ramana, J. (2016). Travelers' Diarrhea-Associated Enterotoxigenic Escherichia coli gyrA Mutants and Quinolone Antibiotic Affinity: A Molecular Dynamics Simulation and Residue Interaction Network Analysis. *OMICS, 20*(11), 635-644. https://doi.org/10.1089/omi.2016.0104

Mizwicki, M. T., Menegaz, D., Zhang, J., Barrientos-Durán, A., Tse, S., Cashman, J. R., . . . Fiala, M. (2012). Genomic and nongenomic signaling induced by 1α, 25 (OH) 2-vitamin D 3 promotes the recovery of amyloid-β phagocytosis by Alzheimer's disease macrophages. *Journal of Alzheimer's Disease, 29*(1), 51-62

Mohammadi, M., Shareghi, B., Akbar Saboury, A., & Farhadian, S. (2020). Spermine as a possible endogenous allosteric activator of carboxypeptidase A: multispectroscopic and molecular simulation studies. *Journal of Biomolecular Structure and Dynamics, 38*(1), 101-113.

https://doi.org/10.1080/07391102.2019.1567387

Monsen, R. C., Deleeuw, L., Maguire, J., Dean, W. L., Gray, R. D., Chaires, J. B., & Trent, J. O. (2017). Structure-based Drug Discovery: Computational Virtual Screening

Neel, B. G., & Tonks, N. K. (2016). *Protein tyrosine phosphatases in cancer*: Springer.

Nocito, D., & Beran, G. J. (2018). Massively Parallel Implementation of Divide-and-Conquer Jacobi Iterations Using Particle-Mesh Ewald for Force Field Polarization. *Journal of chemical theory and computation, 14*(7), 3633-3642

Pandey, P., Roy, K. K., & Doerksen, R. J. (2020). Negative allosteric modulators of cannabinoid receptor 2: protein modeling, binding site identification and molecular dynamics simulations in the presence of an orthosteric agonist. *Journal of Biomolecular Structure and Dynamics, 38*(1), 32-47. https://doi.org/10.1080/07391102.2019.1567384

Pathak, M. K., & Yi, T. (2001). Sodium stibogluconate is a potent inhibitor of protein tyrosine phosphatases and augments cytokine responses in hemopoietic cell lines. *The Journal of Immunology, 167*(6), 3391-3397

Prlić, A., Kalro, T., Bhattacharya, R., Christie, C., Burley, S. K., & Rose, P. W. (2016). Integrating genomic information with protein sequence and 3D atomic level structure at the RCSB protein data bank. *Bioinformatics, 32*(24), 3833-3835

Rao, S. N., Head, M. S., Kulkarni, A., & LaLonde, J. M. (2007). Validation studies of the site-directed docking program LibDock. *Journal of chemical information and modeling, 47*(6), 2159-2171

Ruckert, M. T., de Andrade, P. V., Santos, V. S., & Silveira, V. S. (2019). Protein tyrosine phosphatases: promising targets in pancreatic ductal adenocarcinoma. *Cellular and Molecular Life Sciences*, 1-22

Sadeghi-Kaji, S., Shareghi, B., Saboury, A. A., & Farhadian, S. (2020). Spermine as a porcine pancreatic elastase activator: spectroscopic and molecular simulation studies. *Journal of Biomolecular Structure and Dynamics, 38*(1), 78-88. https://doi.org/10.1080/07391102.2019.1568306

Saiz-Urra, L., Cabrera, M. A., & Froeyen, M. (2011). Exploring the conformational changes of the ATP binding site of gyrase B from Escherichia coli complexed with different established inhibitors by using molecular dynamics simulation: protein-ligand interactions in the light

of the alanine scanning and free energy decomposition methods. *J Mol Graph Model, 29*(5), 726-739. https://doi.org/ 10.1016/j.jmgm.2010.12.005

Sarkar, R. (2018). Stiffening of flexible SUMO1 protein upon peptide-binding: Analysis with anisotropic network model. *Mathematical biosciences, 295*, 67-72

Sarver, P., Acker, M., Bagdanoff, J. T., Chen, Z., Chen, Y.-N., Chan, H., . . . Hao, H. (2019). 6-Amino-3-Methylpyrimidinones as Potent, Selective, and Orally Efficacious SHP2 Inhibitors. *Journal of medicinal chemistry, 62*(4), 1793-1802

Scott, D. E., Coyne, A. G., Hudson, S. A., & Abell, C. (2012). Fragment-based approaches in drug discovery and chemical biology. *Biochemistry, 51*(25), 4990-5003

Sierra, M. L., Beneton, V., Boullay, A.-B., Boyer, T., Brewster, A. G., Donche, F., . . . Grillot, D. A. (2007). Substituted 2-[(4-aminomethyl) phenoxy]-2-methylpropionic acid PPARα agonists. 1. Discovery of a novel series of potent HDLc raising agents. *Journal of medicinal chemistry, 50*(4), 685-695

Smith, M. D., Rao, J. S., Segelken, E., & Cruz, L. (2015). Force-field induced bias in the structure of Aβ21–30: A comparison of OPLS, AMBER, CHARMM, and GROMOS force fields. *Journal of chemical information and modeling, 55*(12), 2587-2595

Sun, X., Ren, Y., Gunawan, S., Teng, P., Chen, Z., Lawrence, H., . . . Wu, J. (2018). Selective inhibition of leukemia-associated SHP2 E69K mutant by the allosteric SHP2 inhibitor SHP099. *Leukemia, 32*(5), 1246

Sun, X., Ren, Y., Gunawan, S., Teng, P., Chen, Z., Lawrence, H. R., . . . Wu, J. (2018). Mutation selectivity of the allosteric SHP2 inhibitor SHP099: AACR.

Susnow, R. G., & Dixon, S. L. (2003). Use of robust classification techniques for the prediction of human cytochrome P450 2D6 inhibition. *J Chem Inf Comput Sci, 43*(4), 1308-1315. https://doi.org/10.1021/ci030283p

Taute, S., Böhnke, P., Sprissler, J., Buchholz, S., Hufbauer, M., Akgül, B., & Steger, G. (2019). The Protein Tyrosine Phosphatase H1 PTPH1 Supports Proliferation of Keratinocytes and is a Target of the Human Papillomavirus Type 8 E6 Oncogene. *Cells, 8*(3), 244

Tiperciuc, B., Zaharia, V., Colosi, I., Moldovan, C., Crişan, O., Pirnau, A., . . . Oniga, O. (2012). Synthesis and Evaluation of Antimicrobial Activity of Some New Hetaryl-Azoles Derivatives Obtained from 2-Aryl-4-methylthiazol-5-carbohydrazides and Isonicotinic

Acid Hydrazide. *Journal of Heterocyclic Chemistry, 49*(6), 1407-1414

Uusitalo, J. J., Ingólfsson, H. I., Marrink, S. J., & Faustino, I. (2017). Martini coarse-grained force field: extension to RNA. *Biophysical journal, 113*(2), 246-256

Van Huijsduijnen, R. H., Bombrun, A., & Swinnen, D. (2002). Selecting protein tyrosine phosphatases as drug targets. *Drug discovery today, 7*(19), 1013-1019

Verma, S., & Sharma, S. (2018). Protein tyrosine phosphatase as potential therapeutic target in various disorders. *Current molecular pharmacology, 11*(3), 191-202

Votano, J. R., Parham, M., Hall, L. M., Hall, L. H., Kier, L. B., Oloff, S., & Tropsha, A. (2006). QSAR modeling of human serum protein binding with several modeling techniques utilizing structure-information representation. *J Med Chem, 49*(24), 7169-7181. https://doi.org/10.1021/jm051245v

W Sousa da Silva, A., & Vranken, W. F. (2012). ACPYPE-AnteChamber PYthon Parser interfacE. *BMC Research Notes, 5*(1)

Wang, L., Deng, Y., Wu, Y., Kim, B., LeBard, D. N., Wandschneider, D., . . . Abel, R. (2016). Accurate modeling of scaffold hopping transformations in drug discovery. *Journal of chemical theory and computation, 13*(1), 42-54

Wesson, L., & Eisenberg, D. (1992). Atomic solvation parameters applied to molecular dynamics of proteins in solution. *Protein Sci, 1*(2), 227-235. https://doi.org/ 10.1002/pro.5560010204

Wu, J.-w., Yin, L., Liu, Y.-q., Zhang, H., Xie, Y.-f., Wang, R.-l., & Zhao, G.-l. (2019). Synthesis, biological evaluation and 3D-QSAR studies of 1, 2, 4-triazole-5-substituted carboxylic acid bioisosteres as uric acid transporter 1 (URAT1) inhibitors for the treatment of hyperuricemia associated with gout. *Bioorganic & medicinal chemistry letters, 29*(3), 383-388

Wu, J.-W., Zhang, H., Duan, Y.-Q., Dong, W.-L., Cheng, X.-C., Wang, S.-Q., & Wang, R.-L. (2014). Design novel inhibitors for treating cancer by targeting Cdc25B catalytic domain with de novo design. *Combinatorial chemistry & high throughput screening, 17*(10), 837-847

Wu, J., Sun, Y., Zhou, H., Ma, Y., & Wang, R. (2019). Design, synthesis, biological evaluation and molecular dynamics simulation studies of (R)-5-methylthiazolidin-4-One derivatives as

megakaryocyte protein tyrosine phosphatase 2 (PTP-MEG2) inhibitors for the treatment of type 2 diabetes. *Journal of Biomolecular Structure and Dynamics*, 1-10. https://doi.org/10.1080/07391102.2019.1654410

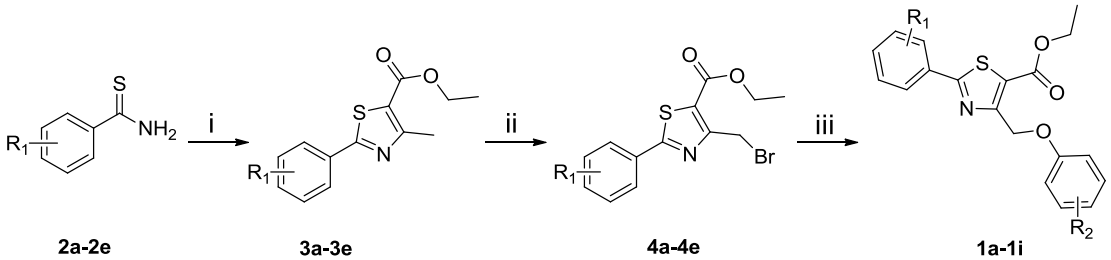
Xia, X., Maliski, E. G., Gallant, P., & Rogers, D. (2004). Classification of kinase inhibitors using a Bayesian model. *J Med Chem*, *47*(18), 4463-4470. https://doi.org/10.1021/jm0303195

Xie, J., Si, X., Gu, S., Wang, M., Shen, J., Li, H., . . . Liu, C. (2017). Allosteric inhibitors of SHP2 with therapeutic potential for cancer treatment. *Journal of medicinal chemistry*, *60*(24), 10205-10219

Zeng, L.-F., Zhang, R.-Y., Yu, Z.-H., Li, S., Wu, L., Gunawan, A. M., . . . Chan, R. J. (2014). Therapeutic potential of targeting the oncogenic SHP2 phosphatase. *Journal of medicinal chemistry*, *57*(15), 6594-6609

Zhang, H., He, X., Ni, D., Mou, L., Chen, X., & Lu, S. (2020). How does the novel T315L mutation of breakpoint cluster region-abelson (BCR-ABL) kinase confer resistance to ponatinib: a comparative molecular dynamics simulation study. *Journal of Biomolecular Structure and Dynamics*, *38*(1), 89-100. https://doi.org/10.1080/07391102.2019.1567390

Zhang, X., He, Y., Liu, S., Yu, Z., Jiang, Z.-X., Yang, Z., . . . Gunawan, A. M. (2010). Salicylic acid based small molecule inhibitor for the oncogenic Src homology-2 domain containing protein tyrosine phosphatase-2 (SHP2). *Journal of medicinal chemistry*, *53*(6), 2482-2493



2a-2e **3a-3e** **4a-4e** **1a-1i**

1a-4a	R_1=2-bromol	R_2=4-chloro	2d-4d, 1f	R_1=4-fluoro	R_2=4-methoxy
1b-4b	R_1=4-methoxy	R_2=4-methoxy	1g	R_1=4-fluoro	R_2=4-chloro
1c	R_1=4-methoxy	R_2=4-methyl	2e-4e, 1h	R_1=4-bromol	R_2=4-methyl
1d	R_1=4-methoxy	R_2=4-chloro	1i	R_1=4-bromol	R_2=4-chloro
2c-4c, 1e	R_1=3-phenoxy	R_2=4-methoxy			

Scheme 1 Synthetic routes to desired products **1a-1i**

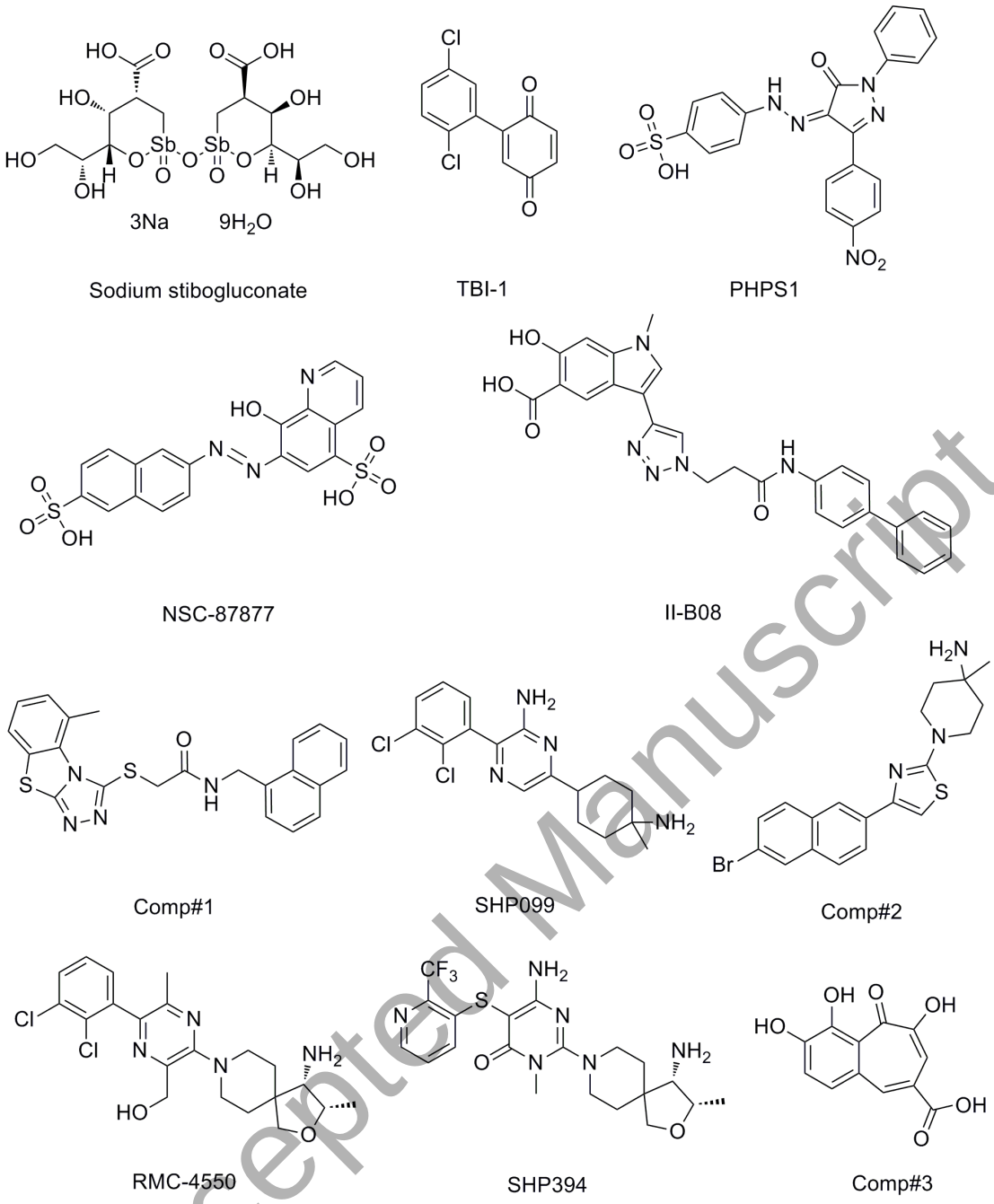
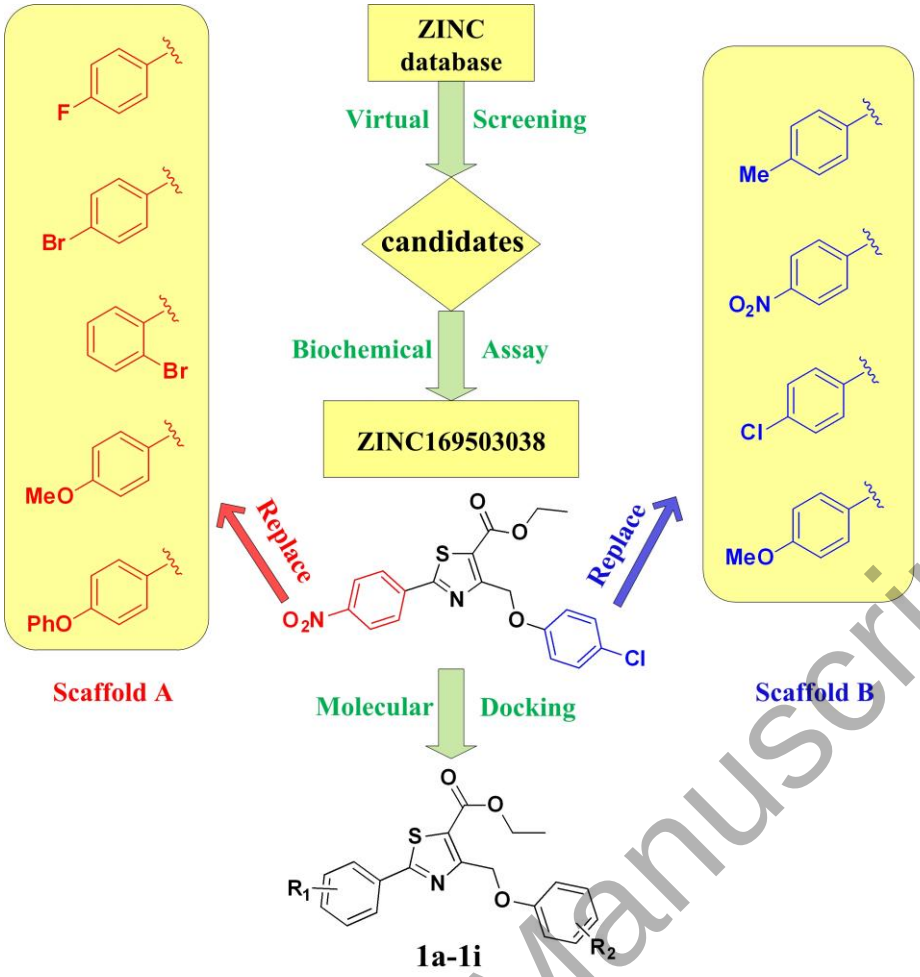


Figure 1 Structures of selected SHP2 inhibitors



1a	R_1=2-bromol	R_2=4-chloro	1f	R_1=4-fluoro	R_2=4-methoxy
1b	R_1=4-methoxy	R_2=4-methoxy	1g	R_1=4-fluoro	R_2=4-chloro
1c	R_1=4-methoxy	R_2=4-methyl	1h	R_1=4-bromol	R_2=4-methyl
1d	R_1=4-methoxy	R_2=4-chloro	1i	R_1=4-bromol	R_2=4-chloro
1e	R_1=3-phenoxy	R_2=4-methoxy			

Figure 2 The compound design flowchart

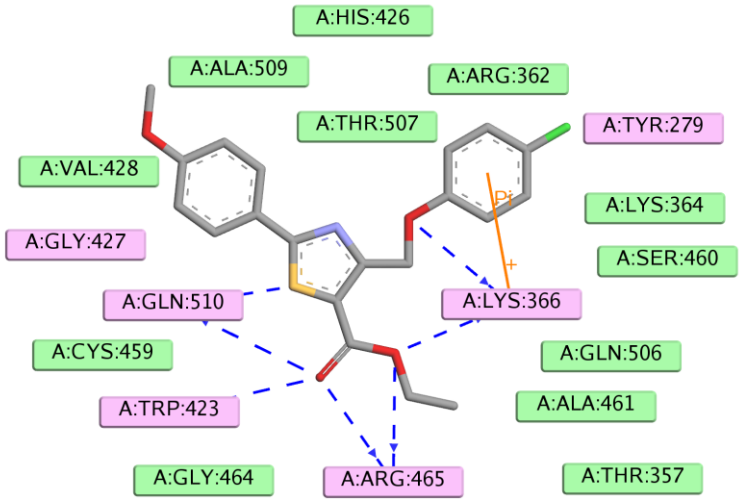


Figure 3 The two-dimensional (2D) diagram of compound **1d**-3O5X interaction. Residues involved in van der waals interactions are represented by green boxes. Residues involved in hydrogen-bond, charge or polar interactions are represented by pink boxes. Hydrogen-bond interaction with amino acid side chain is represented by a blue dashed arrow directed towards the electron donor. Pi-interactions are represented by an orange line with symbols indicating the interaction.

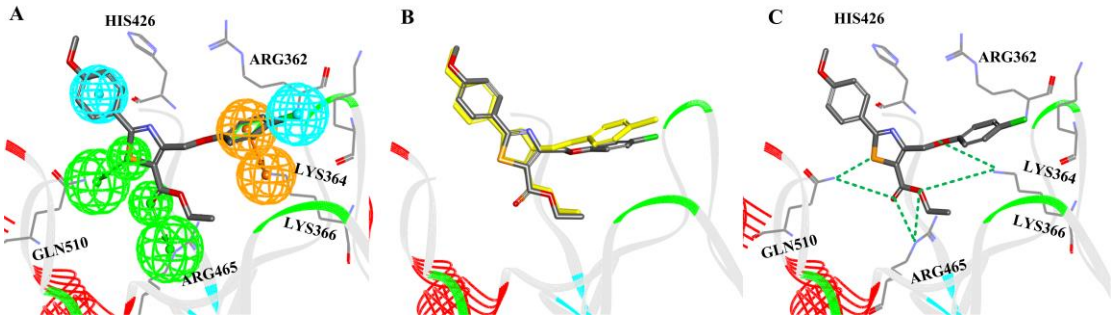


Figure 4 (A) Comparision and superimpose of Hypo-1-SHP2 in active site of SHP2; (B) The superimposition of the docked pose of compound **1d** (gray) to SHP2 to the conformation generated (yellow) in Hypo-1-SHP2; (C**)** the binding model of compound **1d** (gray) with SHP2, respectively. The features are colored coded with green, hydrogen-bond acceptor; brown, Ring aromatic; cyan, hydrophobic.

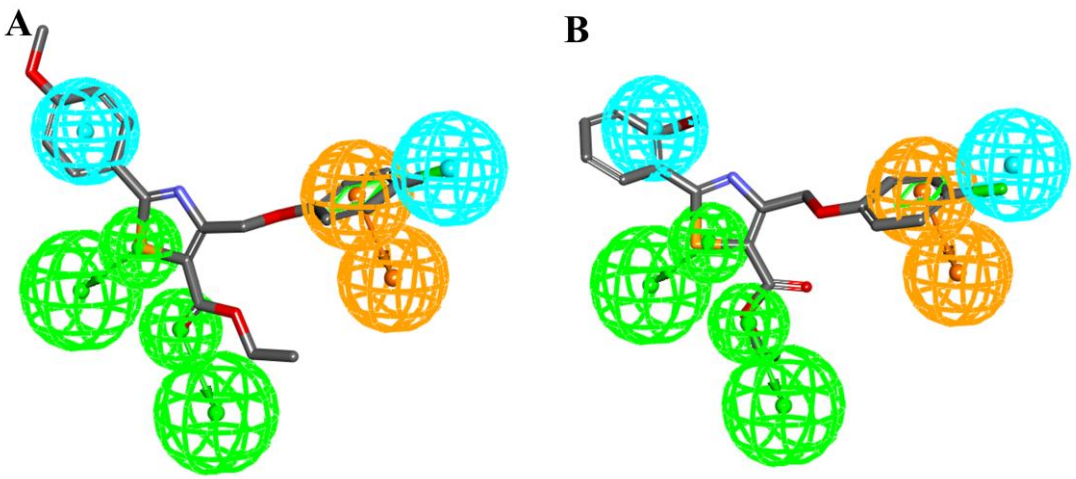


Figure 5 Alignment of hypotheses to training set compounds. (A) most active compound **1d** of SHP2(IC_{50} = 0.99 μM) and (B) least active compound 1a of SHP2 (IC_{50}=35.56 μM).The features are colored coded with green, hydrogen-bond acceptor; brown, Ring aromatic; cyan, hydrophobic.

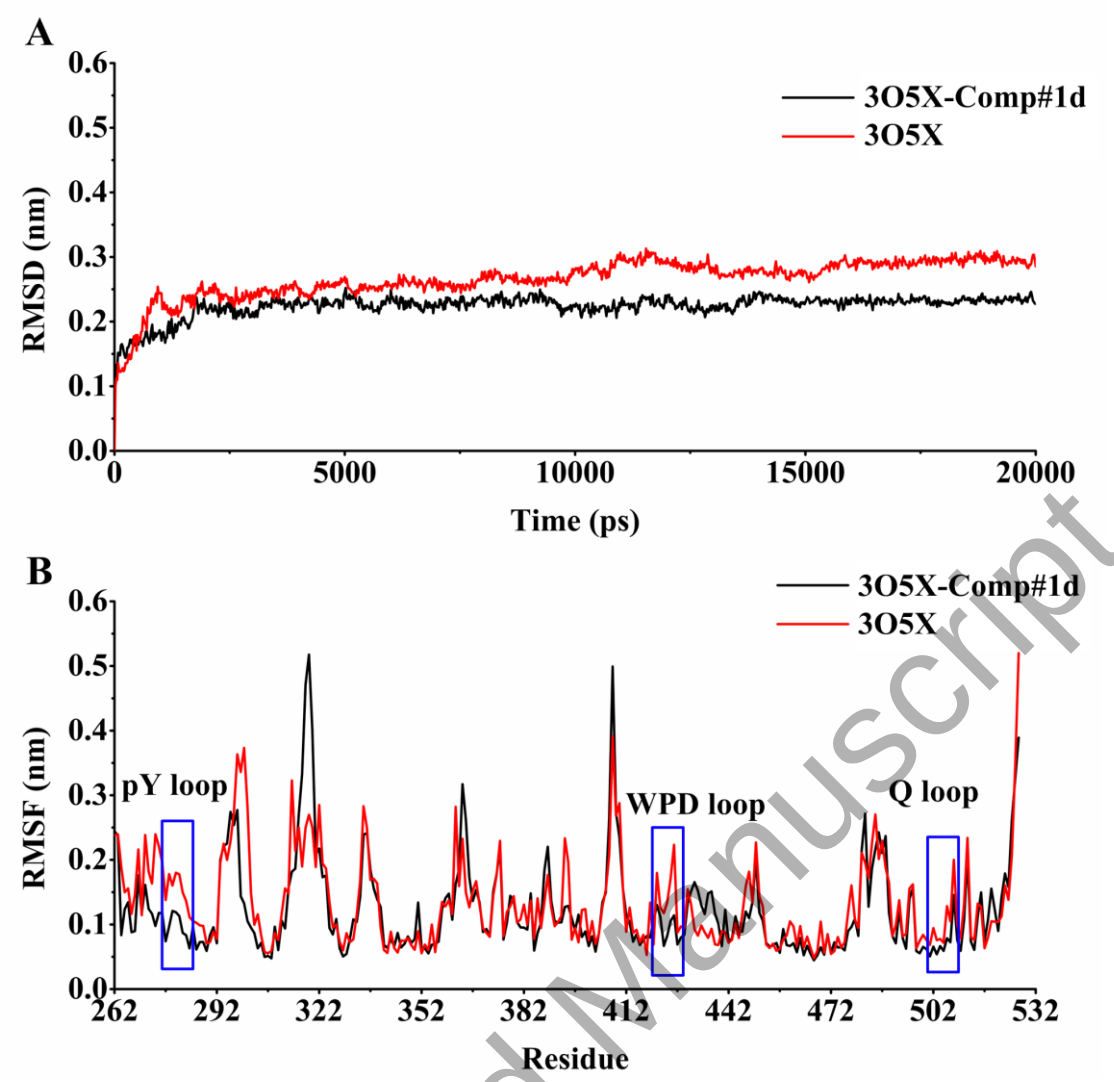


Figure 6 Illustration to show the outcomes of molecular dynamics simulations for SHP2 with ligand compound **1d** and without ligand systems. (A) The RMSD (root mean square deviation) values of all backbone atoms for SHP2 with ligand compound **1d** and without ligand systems, respectively. (B) The RMSF (root mean square fluctuation) of the side-chain atoms for SHP2 with ligand compound **1d** and without ligand systems, respectively. The red line indicates the outcome for the system of the receptor alone without any ligand, the black line for that of the receptor with the ligand compound **1d**.

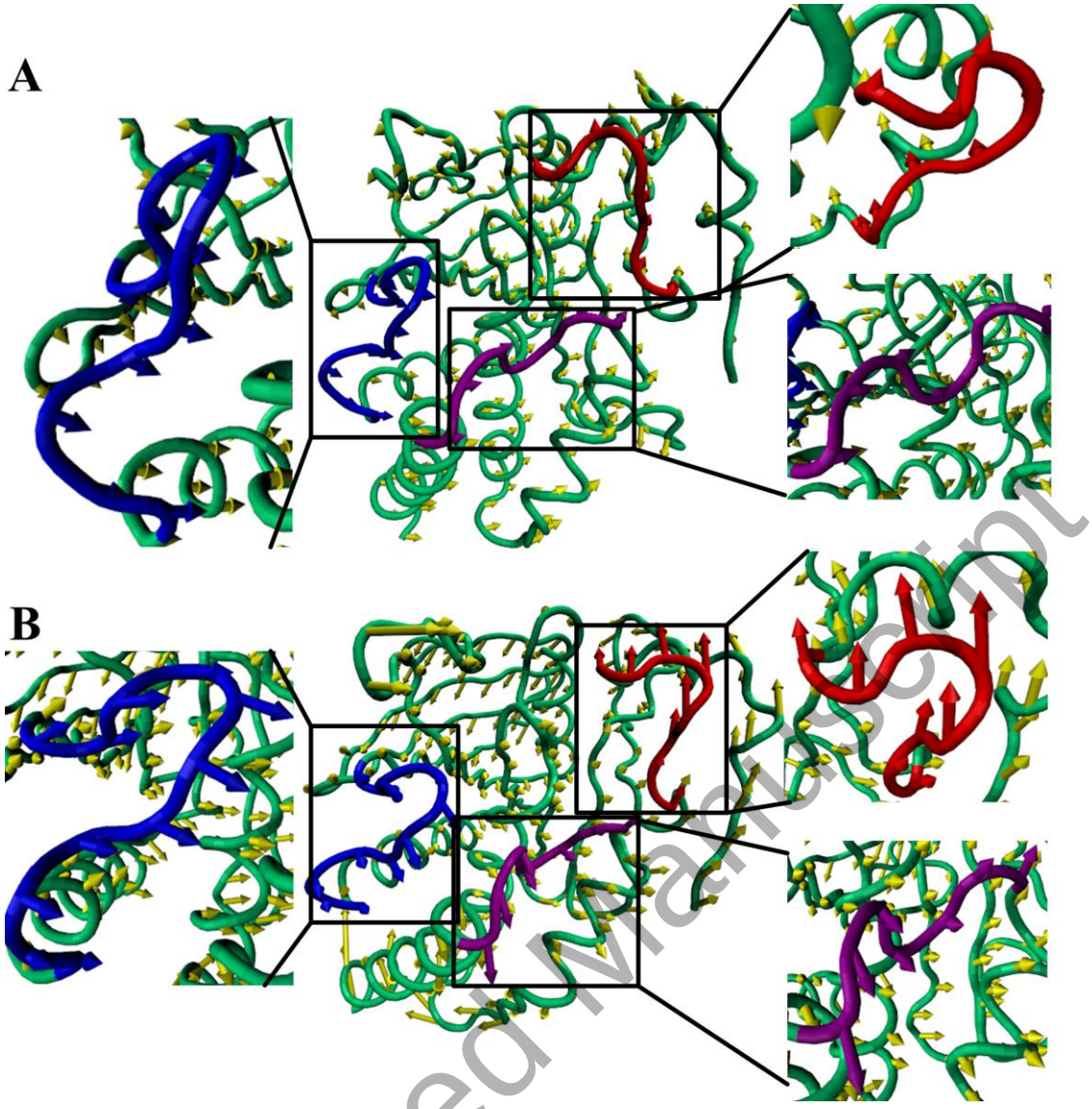


Figure 7 (A) The ANM analysis of SHP2 without ligand system system, (B) The ANM analysis of SHP2-compound **1d** system. The length of arrows is positively-correlated with motive magnitude, and the orientation of arrow indicates motive direction. The Q-loop (residues 501-507), WPD-loop (residues 421-431) and pTyr recognition loop (residues 271-284) are marked in purple, blue, and red, respectively. The areas (Q-loop, WPD-loop and pTyr recognition loop) are magnified and present on the right side of the picture.

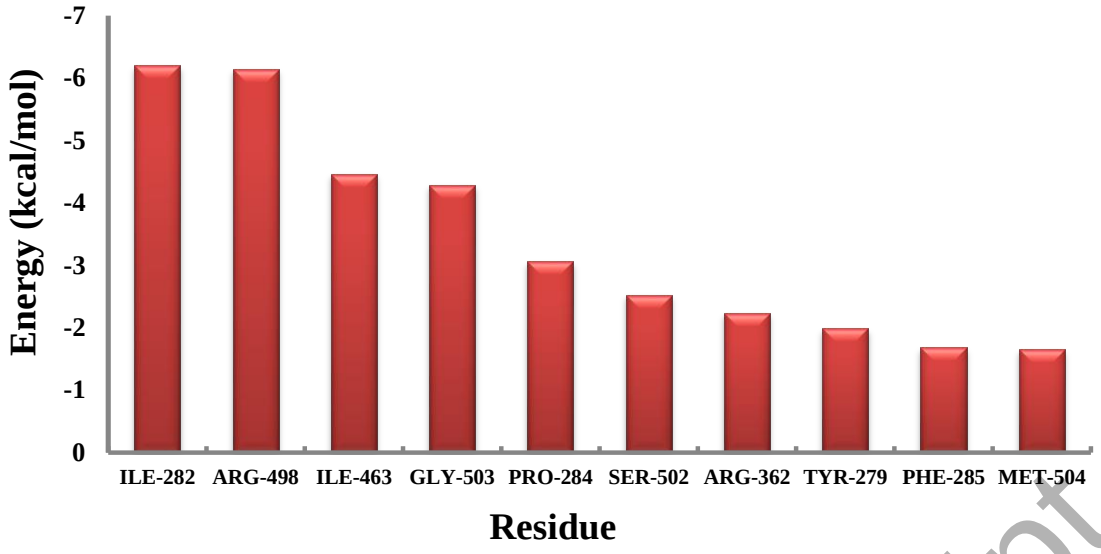


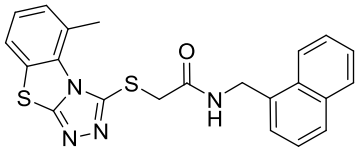
Figure 8 The top ten pairwise residue interaction energies of SHP2/compound **1d**

Table 1 the ADMET prediction for the 4-thiazolidone derivatives.

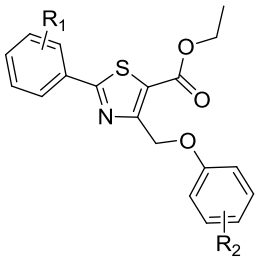
	ALogP[a]	Solubility-level[b]	BBB-Level[c]	CYP2D6 Prediction	Hepatotoxic# Prediction	Absorption -level[d]	PPB# Prediction
1a	5.947	4	1	False	False	1	True
1b	4.501	3	1	False	False	0	True
1c	5.004	4	1	False	False	0	True
1d	5.182	4	2	False	False	0	True
1e	6.078	3	1	False	False	1	True
1f	4.723	4	2	False	False	0	True
1g	5.404	3	0	False	False	0	True
1h	5.768	2	0	False	False	0	True
1i	5.947	3	0	False	False	1	True

a:AlogP, the logarithm of the partition coefficient between n-octanol and water; b:Aqueous solubility level: 0 (extremely low); 1 (very low, but possible); 2 (low); 3 (good);4(optimal);c: BBB level: 0 (very good); 1 (good); 2 (moderate); 3 (poor);4(undefined) ; d: Human intestinal absorption level: 0 (good); 1 (moderate); 2 (poor); 3 (very poor).

Table 2. Structure and SHP2 inhibitory activity of compounds **1a-1i** , ZINC169503038 and Comp#1



Comp#1



Compounds **1a-1i** and ZINC169503038

Compounds	R_1	R_2	MW	-CDOCKER ENERGY (Kcal/moL)	IC_{50} (μM)
Comp#1	-	-	418.53	45.36	11.38[a]
ZINC169503038	4-nitro	4-chloro	414.43	30.95	4.30
1a	2-bromol	4-chloro	452.75	31.64	35.56
1b	4-methoxy	4-methoxy	399.11	46.13	1.25
1c	4-methoxy	4-methyl	383.46	49.32	1.04
1d	4-methoxy	4-chloro	403.88	50.82	0.99
1e	3-phenoxy	4-methoxy	461.53	40.50	3.90
1f	4-fluoro	4-methoxy	387.42	38.18	10.84
1g	4-fluoro	4-chloro	391.84	33.17	11.48
1h	4-bromol	4-methyl	432.33	41.79	4.6
1i	4-bromol	4-chloro	452.75	38.62	5.52

[a]The reported IC_{50} value of lesinurad was 4.31 μM(Duan et al., 2014).

Table 3 Inhibitory activity of compound **1a-1i** on SHP1, TCPTP, MEG2, CDC25B and LAR

Compounds	SHP1 IC_{50} (μM)	TCPTP IC_{50} (μM)	MEG2 IC_{50} (μM)	CDC25B IC_{50} (μM)	LAR IC_{50} (μM)
1a	>100	>100	>100	>100	>100
1b	>100	>100	>100	>100	>100
1c	>100	34.94	>100	>100	>100
1d	>100	34.91	>100	>100	>100
1e	>100	>100	>100	>100	>100
1f	>100	36.65	>100	>100	>100
1g	>100	31.39	>100	>100	>100
1h	>100	26.83	>100	>100	>100
1i	>100	10.16	>100	>100	>100

Table 4. SHP2 inhibitors for HipHop pharmacophore modeling and the results of mapping features.

Compound	FitValue	HBA1	HBA2	RA	HYP1	HYP2	Pharmprint
1a	1.60	1	1	1	0	0	11100
1b	4.80	1	1	1	1	1	11111
1c	4.95	1	1	1	1	1	11111
1d	4.96	1	1	1	1	1	11111
1e	3.66	1	1	1	1	1	11111
1f	2.94	1	1	1	1	1	11111
1g	2.94	1	1	1	1	1	11111
1h	3.05	1	1	1	0	1	11101
1i	3.05	1	1	1	0	1	11101

Table 5 Binding free energies (kcal/mol) and its components between receptor and ligand, respectively.

complex	Van der Waal (Kcal/mol)	Electrostatic (Kcal/mol)	Polar solvation (Kcal/mol)	Non-polar solvation (Kcal/mol)	Binding energy (Kcal/mol)
SH2-compound 1d	-112.124	-7.688	55.803	-12.729	-76.737
SH2-compound 1a	-102.290	-5.317	52.251	-11.592	-66.949