Journal of Computer-Aided Molecular Design
https://doi.org/10.1007/s10822-019-00213-z

Structure based design of selective SHP2 inhibitors by De novo design, synthesis and biological evaluation

Wen-Shan Liu[1] · Wen-Yan Jin[1] · Liang Zhou[1] · Xing-Hua Lu[1] · Wei-Ya Li[1] · Ying Ma[1] · Run-Ling Wang[1]

Received: 27 December 2018 / Accepted: 8 July 2019

Abstract

SHP2 phosphatase, encoded by the PTPN11 gene, is a non-receptor PTP, which plays an important role in growth factor, cytokine, integrin, hormone signaling pathways, and regulates cellular responses, such as proliferation, differentiation, adhesion migration and apoptosis. Many studies have reported that upregulation of SHP2 expression is closely related to human cancer, such as breast cancer, liver cancer and gastric cancer. Hence, SHP2 has become a promising target for cancer immunotherapy. In this paper, we reported the identification of compound 1 as SHP2 inhibitor. Fragment-based ligand design, De novo design, ADMET and Molecular docking were performed to explore potential selective SHP2 allosteric inhibitors based on SHP836. The results of docking studies indicated that the selected compounds had higher selective SHP2 inhibition than existing inhibitors. Compound 1 was found to have a novel selectivity against SHP2 with an in vitro enzyme activity IC_{50} value of 9.97 μM. Fluorescence titration experiment confirmed that compound 1 directly bound to SHP2. Furthermore, the results of binding free energies demonstrated that electrostatic energy was the primary factor in elucidating the mechanism of SHP2 inhibition. Dynamic cross correlation studies also supported the results of docking and molecular dynamics simulation. This series of analyses provided important structural features for designing new selective SHP2 inhibitors as potential drugs and promising candidates for pre-clinical pharmacological investigations.

Keywords SHP2 · Selective allosteric inhibitors · De novo · Molecular dynamics simulation

Abbreviations

JMML	Juvenile myelomonocytic leukemia
MS	Myelodysplastic syndrome
AML	Acute myeloid leukemia
TCPTP	T cell protein-tyrosine phosphatase
PTP1B	Protein tyrosine phosphatase 1B
SHP1	SH2 domain-containing phosphatase 1
H bond	Hydrogen bond
ADMET	Absorption, distribution, metabolism, excretion, and toxicity
MCSS	Multi-copy simultaneous search
FBDD	Fragment based drug design
HIA	Human intestinal absorption
BBB	Blood–brain barrier
PPB	Aqueous solubility plasma protein binding
PME	Particle mesh Ewald
MM-PBSA	Molecular mechanics Poisson Boltzmann surface area
LCPO	Linear combination of pairwise overlaps
DCC	Dynamic cross correlation
MD	Molecular dynamics
HTVS	High throughput virtual screening
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation

Wen-Shan Liu and Wen-Yan Jin contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10822-019-00213-z) contains supplementary material, which is available to authorized users.

✉ Ying Ma
maying@tmu.edu.cn

✉ Run-Ling Wang
wangrunling@tmu.edu.cn

[1] Tianjin Key Laboratory on Technologies Enabling Development of Clinical Therapeutics and Diagnostics (Theranostics), School of Pharmacy, Tianjin Medical University, No. 22 Qixiangtai Rd, Heping Dist, Tianjin 300070, People's Republic of China

Published online: 12 July 2019

Introduction

SHP2 phosphatase, encoded by the PTPN11 gene, is a non-receptor PTP, which is composed of two N-terminal Src homology 2 (SH2) domains, one PTP domain, and one C-terminal tail [1]. In cells, SHP2 plays a role in the cytoplasm downstream of multiple receptor-tyrosine kinases and is involved in many cancer cell signaling cascades (e.g., RAS–ERK, PI3K–AKT, JAK–STAT) [2]. RAS was dephosphorylated by SHP2 and then associated with effector protein RAF to activate downstream proliferative RAS/ERK/MAPK signaling. Furthermore, germline or somatic mutations in PTPN11 that cause hyperactivation of SHP2 have been identified in Noonan syndrome (50%), juvenile myelomonocytic leukemia (JMML, 35%), myelodysplastic syndrome (MS, 10%), B cell acute lymphoblastic leukemia (7%), acute myeloid leukemia (AML, 4%) [3–7], and solid tumors, including colon cancer, melanoma, lung adenocarcinoma, and hepatocellular carcinoma. Thus, SHP2 became a promising target for cancer immunotherapy.

Many SHP2 inhibitors have been identified so far. For example, NSC-87877 was the first PTP inhibitor, inhibiting PTP domain of SHP2 in cell cultures without a detectable off-target effect, which cross-inhibited SHP1 in vitro [8]. Other SHP2 inhibitors, such as II-B08, cefsulodin, Fumosorinone, had also been identified and characterized [2, 3, 9]. However, the discovery of SHP2 inhibitors has been hampered for two main aspects. Firstly, SHP2 has low selectivity for other PTPs (such as SHP1, PTP1B, and TCPTP) due to the highly conserved of PTP domain (the overall sequence of SHP1 and SHP2 in their PTP domain has 60% identity and about 75% similarity) [5, 10, 11]; Secondly, it is difficult to identify cell permeable compounds. Thus, SHP2 inhibitors have not yet progressed to the clinic [12]. The identification of SHP2 allosteric sites (a tunnel-like region formed between the C–SH2, N–SH2, and PTP domains, but not the PTP catalytic site) represented a potentially solution to the problem of druggability. By binding to this allosteric site comprised of all three domains (N–SH2, C–SH2, and PTP domain), SHP836, one of SHP2 allosteric inhibitor, stabilized the active conformation of SHP2, in which the catalytic site was blocked and no longer accessible to substrate [10]. SHP836 had no inhibition against PTP domain of SHP2 and had moderate inhibitory activity ($IC_{50} = 12$ μM) against full length SHP2. SHP099, a highly selective, active, potent SHP2 allosteric inhibitor, was further optimized by SHP836, which was identified as an allosteric modulator that stabilized the auto-inhibited conformation of SHP2 [13]. There are many methods to discover novel drugs, one of which is an economical and time-saving method, namely computer-aided drug design [14]. In order to discover novel and selective SHP2 allosteric inhibitors, we used De novo design to generate scaffolds. Then, ADMET prediction was carried out to screen for good quality compounds. The selected compounds were then docked into binding pockets of SHP2 and SHP1, respectively. In vitro phosphatase activity and cell proliferation assays were performed to verify biological activity. Finally, Molecular dynamic simulation, binding free energy calculation and dynamic cross correlation study were used to explore the affinity and selectivity of SHP2. Thus, the findings obtained could provide useful insights for developing new and powerful SHP2 selective allosteric inhibitors against cancer and provide lead compounds.

Materials and methods

Our calculations were carried out on Dell Precision TM T5500 computer with Discovery Studio v 3.5 software package.

Protein preparation

The crystal structures of SHP1 (PDB ID: 3PS5) and SHP2 (PDB ID: 5EHR) were downloaded from the protein data bank (PDB) and used for molecular simulation studies [15, 16]. The "clean protein" [17] protocol in Discovery Studio v 3.5 was used to prepare the protein for correcting the missing atoms and residues, incorrect atom order in amino acids, deleting alternate conformations and water, and protonating all the residues at a specific pH conditions.

Fragment-based ligand design

Fragment-based ligand design approach, such as the multicopy simultaneous search (MCSS) methodology, had proven to be a useful tool to search novel therapeutic compounds that bind to pre-specified targets of known structure. MCSS offered a variety of advantages over traditional high-throughput screening methods, and had been applied successfully to design novel hit targeting challenging targets. In our study,MCSS algorithm was used for fragment-based drug design (FBDD) [18]. The major steps in the procedure are as follows: distribute fragment replicas in the protein search sphere; perform a CHARMm minimization; prune replicas to remove fragments converged to the same minima. According to the top MCSS_score poses, new scaffold libraries (Scaffolds A and B) were built to De novo design.

De novo design

Ludi, one of the most widely used re-design algorithm, was used to discover new potentially active compounds, which

could save researchers a lot of time [19]. It was a powerful design tool, which allowed users to simulate screening prior to experimental analysis, and allowed the transformation of existing compounds. The Ludi algorithm worked in three steps. Firstly, it calculated interaction sites, which were discrete positions in space that was suitable to form hydrogen bonds or fill hydrophobic pockets. Then, the molecular fragments were fitted onto the interaction sites. Finally, some or all of the fitted fragments was connected to a single molecule.

In our study, De novo design method [20] by Ludi algorithm was used to design novel inhibitors. Firstly, De novo library generation protocol was used to generate fragments library for De novo design. Secondly, De novo receptor protocol in Discovery Studio v3.5 was used to define the binding site of a receptor. Scaffold A and scaffold B were derived from computational Fragment-based drug design in allosteric sites. Thirdly, the De novo link protocol used "Ludi algorithm" to add linkers between the defined scaffolds A and scaffold B in the binding site of a receptor. The placed linker parts were scored using Ludi energy estimation, and then the final scaffolds C library was established. Accordingly, each scaffold C was made up of scaffold A and scaffold B and linkers, respectively. Fourthly, the protocol of De novo evolution could develop whole molecule in the binding site environment of the receptor based on the scaffold C. The Ludi algorithm was used to add appropriate fragments to scaffold C, and then produced a collection of whole molecules with higher scores. The evolution mode was set to full evolution, which allowed the scaffold C to link up to a maximum of three fragments. Finally, 24 top-ranked molecules were selected for the following ADMET analysis.

Lipinski's filter and ADMET analysis

During virtual screening in early drug discovery, the ADMET descriptors of DS v3.5 could be used for estimating crucial physicochemical and biological properties for large numbers of candidate drug compounds. The ADMET (absorption, distribution, metabolism, excretion, and toxicity) properties of each compounds were calculated to assess the good pharmacokinetics of a drug in the human body using the 'ADMET Descriptors' calculation of DS v3.5. Some important ADMET descriptors were calculated, including human intestinal absorption (HIA) [21], blood–brain barrier (BBB) [22], aqueous solubility plasma protein binding (PPB) [23] and cytochrome P450 CYP 2D6 inhibition hepatotoxicity [24]. Hence, attrition could be reduced in drug discovery and development by filtering the ADMET characteristics of the gained compounds. The property analyses for carcinogenicity, aerobic biodegradability, developmental toxicity potentials, AMES mutagenicity, and ocular and skin irritancy [25] were considered to evaluate the toxicity of the compounds in the Toxicity Prediction (TOPKAT) module. Promising compounds thus were obtained for the further studies.

Flexible docking

Molecular docking was one of the most frequently used methods in drug design, which was used to investigate interaction patterns of target protein with its inhibitors [26]. In the process of docking, the conformations of the protein side chain and ligand molecular were flexible.

During the process of core docking, the 1st step, protein was prepared by the clean protein protocol in DS v3.5. The 2nd step was to calculate protein conformation using ChiFlex [27] (CHARMm) by changing the side chain conformation. The 3rd step, the preparation ligand protocol was used to prepare ligand for pharmacophore generation, including removing duplicates, enumerating isomers tautomers, generating 3D conformation, and generating possible states by ionization at target pH 7.0 ± 2.0. The 4th step was to define the binding site. There were two ways to define the binding pocket: the first was to define binding site from key residue in the structure of the receptor; the second was to calculate a binding site from a selected ligand. In our study, the bind site was defined applying the second method. In the docking process, the residues Arg111, Phe113 Glu250, Leu254, Glu257, Pro491, and Glu495 were used to generate active site for SHP2 and Arg109, Glu247, Ser250, Gln254, Gln485 and Gln489 for SHP1. The 5th step was to optimize the selected protein side chains using the ChiRotor [27] in the presence of a rigid ligand. In 6th step, ligands were flexibly docked to the binding sphere of diverse receptor conformations by the LibDock program. The last step was to optimize the final ligand using CDOCKER program [28]. Prior to the docking analysis, the docking model was validated by the re-dock method. When all steps were finished, the compounds were docked into the receptor pockets.

Chemistry

All the reagents were purchased from commercial suppliers and were used without further purification unless otherwise indicated. 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). All column chromatography was performed using 200–300 mesh silica gel. ^{1}H-NMR and ^{13}C-NMR spectra were taken on a Bruker Avance 300-MHz NMR Spectrometer at 300 K with TMS as the internal standard, and $CDCl_3$ and DMSO-d_6 were used as solvents, the values of the chemical shifts (δ) were expressed in parts per million (ppm), and coupling constants (J) were expressed in hertz (Hz). MS spectra

were recorded on an Agilent 1100 LC/MSD (ESI) Mass Spectrum.

SHP2 kinase assay and fluorescence titrations assay

Glutathione S-transferase (GST) fusion proteins of purified SHP2 were used as the enzyme, and a peptide sequence EFpYAEVGRSPPDPAK (H–Glu–Phe–pTyr–Ala–Glu–Val–Gly–Arg–Ser–Pro–Pro–Asp–Pro–Ala–Lys) was used as the substrate. The Malachite Green Phosphate Assay Kit was based on quantification of the complex formed between Malachite Green, molybdate, and free orthophosphate. The assay determined free phosphate generated by dephosphorylation of the substrate using the Malachite Green Phosphate Assay Kit. Firstly, 0.4 mg/mL GST–SHP2 protein, 0.06 mM substrate and assay buffer (25 mM Tris–HCl, 50 mM NaCl, 5 mM DTT, and 2.5 mM EDTA, pH 7.4) with the tested compounds at various concentrations were added in 96well plates at room temperature for 30 min. Secondly, 25 µL NaOH (4 M) was added to sample, and heated at 100 °C for 30 min and cooled to room temperature. Then, 25 µL HCl (4 M) was added to the sample. Finally, 20 µL of Malachite Green working reagent was added, and OD620 was measured after 30 min at room temperature. The tested compounds were used to determine its IC_{50} value against SHP2. The results were analyzed by GraphPad Prism software.

In order to validate that compound 1 (comp#1) bound directly to the SHP2 protein, fluorescence quenching assay was used to examine whether the binding of the inhibitor altered the fluorescence of the SHP2 protein. Purified SHP2 His-fusion protein was diluted into reaction buffer with pH 7.4. Titrations were performed by increasing the compound 1 concentrations while maintaining the SHP2 protein concentration at 0.4 mg/mL. Contributions from background fluorescence of the inhibitor were explained by subtracting the fluorescence of individual inhibitors from the protein–inhibitor solution. The excitation wavelength was 295 nm, and fluorescence was monitored from 360 to 500 nm. All reported fluorescence intensities were relative values and were not corrected for wavelength variations in detector response.

Cell culture and cell viability assay

Mouse B lymphocyte BaF3 from the American Type Culture Collection (ATCC; Manassas, VA, USA) was cultured in RPMI 1640 (Gibco) medium supplemented with 10% fetal bovine serum (FBS; Origin South America sterile filtered), 1% penicillin–streptomycin, and 2 ng/mL murine IL-3 at 37 °C in a humidified atmosphere of 5% CO_2.

Cell growth was measured by 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxy methoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTS) assay. Briefly, cells were seeded in 96-well plates at a density of 1.2×105 cells/mL and cultured for 48 h. 1–3 h before culture termination, 20 µL of MTS reagent was added to the wells. The absorbance density was read on a 96-well plate reader (NEST, Nest Biotechnology) at wavelength 490 nm. IC_{50} values were calculated as the concentration of the drug required to obtain 50% of maximal inhibition in cell viability.

Molecular dynamics

The docked structure was served as a starting structure for MD simulations using Gromac4.5.5 [29]. GROMOS96 43a1force field was used for the protein [30]. The complex was solvated in a rectangular box of spc216 water, with a minimum distance of 1.0 Å between the protein and the box. Sodium ions were added to the system by random replacement of water molecules to neutralize the system [31]. The particle mesh Ewald (PME) method was used to handle long-range coulomb interactions [32]. Before starting the dynamics, we could confirm there were no atoms in a collision, no conflict of chemical bonds, no irrelevant three-dimension. Then, the steepest descent method was used to carry out energy minimization. The whole system was carried out NVT equilibration at constant temperature of 300 K for 100 ps, and then was equilibrated with NPT with constant pressure of 1 atm for 100 ps. LINCS algorithm [32] was used to keep the bonds constrained. A production run for 10 ns was performed using NPT ensemble at 300 K and 1.0 atm pressure with a time step of 2 fs. Coordinate trajectories were recorded every 2 ps for the whole MD runs.

Binding free energy calculation

Binding free energies for all complex systems were calculated by using the molecular mechanics Poisson Boltzmann surface area (MM-PBSA) method [33]. The free energy of binding ΔG_{bind} was calculated as:

$$\Delta E_{MM} = \Delta E_{internal} + \Delta E_{vdw} + \Delta E_{ele} \quad (1)$$

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

$$\Delta G_{sol} = \Delta G_{nonpol} + \Delta G_{pol} \quad (3)$$

Here, ΔE_{MM}, ΔG_{sol} and $T\Delta S$ were equivalent to the changes of the gas phase MM energy, the solvation free energy and the conformational entropy on binding. ΔE_{MM} was standard molecular mechanics term including $\Delta E_{internal}$ (bond, angel and dihedral energies) which would be cancelled when we used a single trajectory approach to reduce the noise [34], van der Waals interaction ΔE_{vdw}

and electrostatic ΔE_{ele} energies. The nonpolar solvation free energy ΔG_{nonpol} was calculated from the area (SASA) using the method of linear combination of pairwise overlaps (LCPO) (ΔG_{nonpol} ¼ = 0.0072 × ΔSASA) [35]. The SASA here was determined with probe radii of 1.4. The electrostatic free energy of solvation ΔG_{pol} was calculated by the generalized Born method (igb ¼ 5) developed by Onufriev et al. [36]. TΔS was the conformational entropy change calculated by normal mode analysis on a set of conformational snapshots taken from MD simulation [37, 38].

A total of 50 snapshots that were extracted during equilibrium phase between 10 and 50 ns were employed for the calculation using g-mmpbsa tool of Gromacs 4.5.5.

Dynamic cross-correlation

The dynamic cross correlation (DCC) analysis was a popular method for analyzing the trajectories of molecular dynamics (MD) simulations. The DCC between the *i*th and *j*th atoms was defined by the following equation:

$$\mathrm{DCC}(i,j) = \frac{\langle \Delta \mathbf{r}i(t) \cdot \Delta \mathbf{r}j(\mathrm{t}) \rangle \mathrm{t}}{\sqrt{\left\langle \left\| \Delta \mathbf{r}_{\mathrm{i}}(\mathrm{t}) \right\|^2 \right\rangle_{\mathrm{t}}} \sqrt{\left\langle \left\| \Delta \mathbf{r}_{\mathrm{j}}(\mathrm{t}) \right\|^2 \right\rangle_{\mathrm{t}}}}, \tag{4}$$

where $\mathbf{r}_i(\mathrm{t})$ denoted the vector of the *i*th atom's coordinates as a function of time t, $<.>_{\mathrm{t}}$ meant the time ensemble average and $\Delta \mathbf{r}i(t) = \mathbf{r}i(t) - \langle \mathbf{r}i(t) \rangle_{\mathrm{t}}$ [39]. The cross-correlated displacement of Cα atoms across MD simulations could provide critical information about involvement of specific domains in ligand binding. Thus dynamic cross-correlation maps (DCCM) were generated to denote dynamic cross-correlated displacements of Cα atoms through MD simulation for both SHP2 and SHP1. The complete frames were superimposed against the original starting structure of each system prior to the implementation of the analysis.

Results and discussion

Fragment-based ligand and design De novo design

It was a big challenge to identify an appropriate scaffold in developing effective small molecule inhibitors. Most SHP2 inhibitors stemmed from either the modified derivatives of pre-existing inhibitor scaffolds or the isolation of new scaffolds by high throughput virtual screening (HTVS). In our study, we had designed a series of novel scaffolds C by utilizing De novo link protocol. It could be seen from the Fig. 1 that the Scaffold C was divided into three parts, Scaffold A, Scaffold B, and the linker. Scaffold A and Scaffold B were derived from computational Fragment-based drug design in allosteric sites, while the linker stemmed from fragment library by the De novo link protocol. Based on the Scaffolds C, a total of 130 top-ranked "Ludi_3_SCORE" molecules were selected for the following ADMET procedure.

ADMET analysis

ADMET properties of all designed compounds were predicted and compared using DS v3.5 based on ADMET properties of standard compounds. Prediction of ADMET properties was used to sort out those compounds that already followed Lipinski′s rule and showed good predicted activity. DS v3.5 defined the prediction level for all ADMET properties. For aqueous solubility level: 0 (extremely low); 1 (very low, but possible), 2 (low), 3 (good); for ADMET plasma protein binding (PPB) level: 0 meant that the binding level was below 90%, 1 meant that the binding level was above 90%, and 2 meant that the binding level was above 95%; for blood–brain barrier (BBB) level: 0 (very good), 1 (good), 2 (moderate), 3 (poor), 4 (undefined); for ADMET Cytochrome P450 2D6 (CYP2D6) Probability level: Less than 0.5 meant that it was unlikely to inhibit CYP2D6 (Non-CYP2D6 inhibitors), and above 0.5 meant that it was likely to inhibit CYP2D6 enzyme (CYP2D6 Inhibitors); for hepatotoxic prediction, true represented non-hepatotoxic and for human intestinal absorption (HIA) level: 0 (good), 1 (moderate), 2 (poor), 3 (very poor). The results of toxicity prediction were listed in Table S1, including Ames Mutagenicity, Rat Oral LD_{50}, Aerobic Biodegradability, and NTP Carcinogenicity Call. According to the results of toxicity prediction, 86 molecules were relatively safe by passing ocular irritancy, mutagenicity, and skin sensitization tests and were found to be aerobically biodegradable. The toxicity and ADMET prediction of top 10 compounds (Fig. 2) were listed in Table S1 and S2. Finally, 50 promising compounds within the reasonable ranges were obtained for the further studies.

Flexible docking

To estimate the applicability of flexible docking in our study, the original crystal structure of ligand-SHP2 from the PDB was re-docked into the X-ray structure of the receptor. The root mean square deviation (RMSD) between the docked poses and their actual X-ray poses in the crystal structure were 0.4461 Å for SHP2 complex [40]. 50 compounds were docked into crystal structures of SHP1 and SHP2 by using the flexible docking, respectively.

The two-dimensional (2D) diagram of SHP2–SHP099 interaction was shown in Fig. 3. It could be seen from Fig. 3a that the SHP099 basically occupied same binding pocket of SHP2. The H-bonding networks played a role in stabilizing the conformation of the SHP2, which was vitally important for receptor-binding and activation. The H-bond

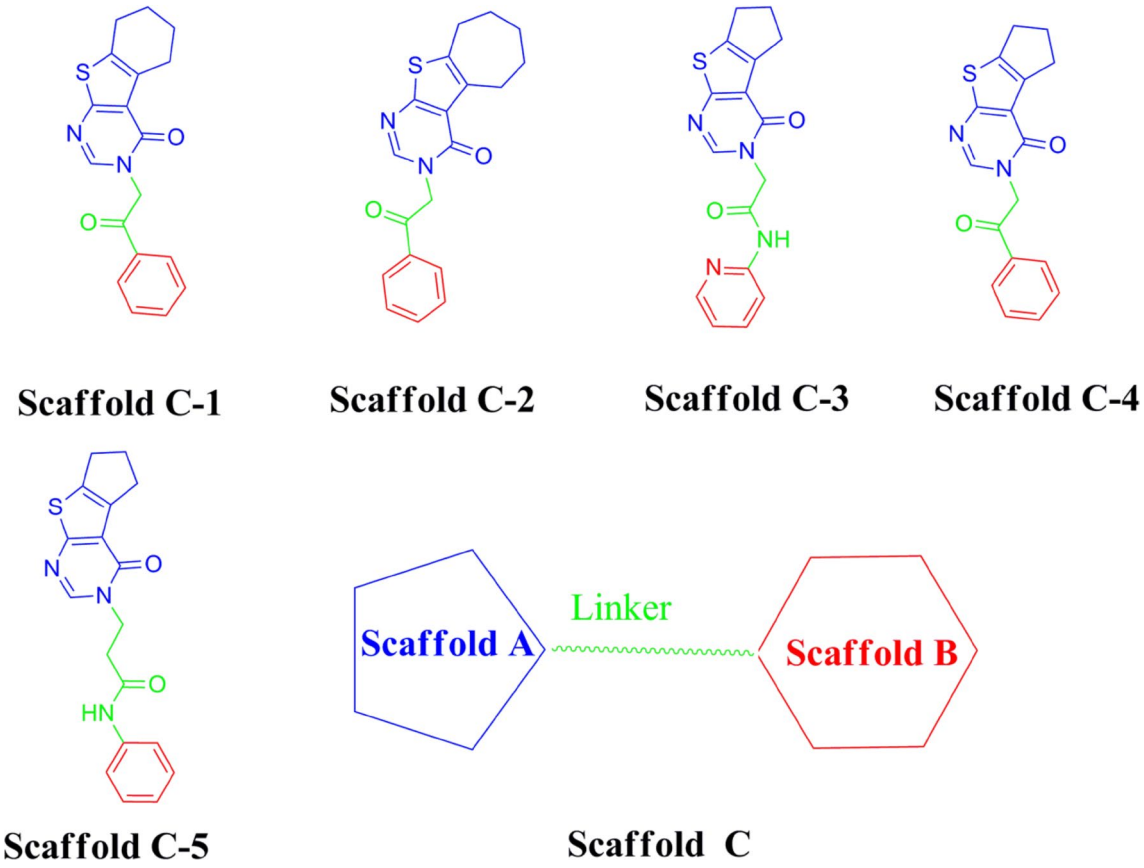


Fig. 1 The illustration to show the chemical structures of the top five scaffolds. The Scaffold A, Scaffold B and the linker are colored in blue, red and green, respectively

interactions were formed by three key residues of SHP2, including Arg111, Phe113 and Glu250. In addition, SHP099 formed hydrophobic interactions with residues His114 and Lys492 of the side chains. Figure 3b displayed the binding pattern of the ligand SHP099 and SHP2 at the active site. For SHP099, it formed H-bond interactions with residues Arg111–N2 and Phe113–N22 located on the linker between the N–SH2 and C–SH2 domains and Glu250–N7 from the PTP domain. The phenyl ring of SHP099 formed a cation–Pi interaction with residue Arg111.

The CDOCKER_ENERGY was an important index for evaluating the binding affinities. The docking score of top 10 compounds was listed in Table 1, the compound 1 (CDOCKER_ENERGY = 46.7889) had the higher docking score than SHP099 (CDOCKER_ENERGY = 40.9327). The selected 10 compounds with better binding affinities were shown in Fig. 2.

General procedure for the synthesis of compounds

Compound 1a (20 mmol), isopropyl cyanoacetate (22 mmol), sulfur powder (28 mmol), and 100 mL ethanol were added to the flask and stirred for 10 min. Then, morpholine (11 mmol) was added dropwise to the system, and the mixture was heated to 80 °C under N_2 for 24 h. The reaction mixture was poured into ice water (50 mL) for quenching. Then, it was extracted with CH_2Cl_2 (3 × 50 mL), and the organic phases were combined, washed with 5% brine (2 × 50 mL) and dried over anhydrous Na_2SO_4. The organic phase after drying was removed in vacuo and the residue obtained was purified by chromatography to gain compound 2a. The same method gained compounds 2b, 2c.

Compound 2a (10 mmol) and 30 mL formamide were added to the flask and heated to 210 °C under N_2 for 12 h. After the temperature of the system was lowered to room temperature, 30 mL ice water was added to the system, stirred for 2 h, and then suction filtration. The filter cake was washed with water (2 × 20 mL) and ethanol (2 × 10 mL), and the filter cake was dried to gain the compound 3a. The same method gained compounds 3b, 3c

Compound 2a (10 mmol), acetonitrile (200 mmol) and hydrogen chloride in 1,4-dioxane solution (50 mL, 2 mol/L) were added to the flask. The mixture was stirred at 60 °C for 12 h, and the reaction was completed by TLC analysis. After cooling to room temperature, the solvent was removed under reduced pressure, and pH was adjusted to 9 with

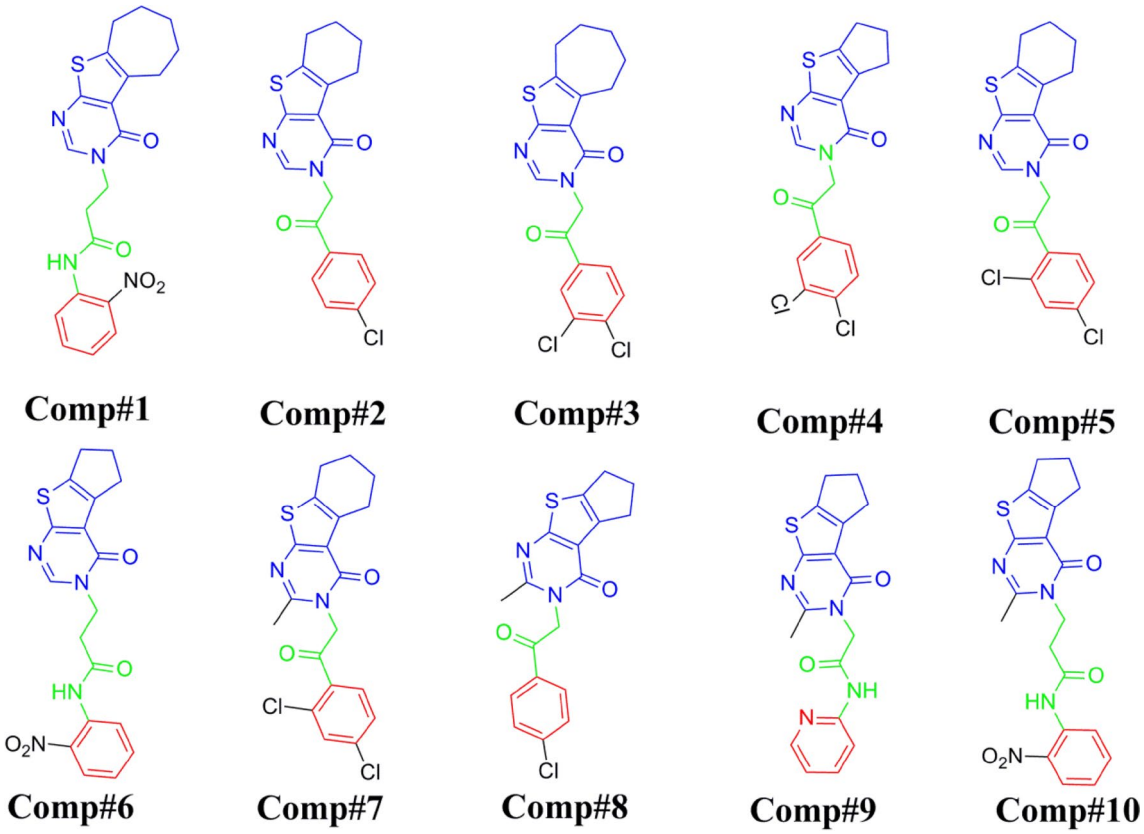


Fig. 2 The illustration to show the chemical structures of the top 10 compounds. The Scaffold C and the fragments generated by the protocol of De novo evolution are colored in black

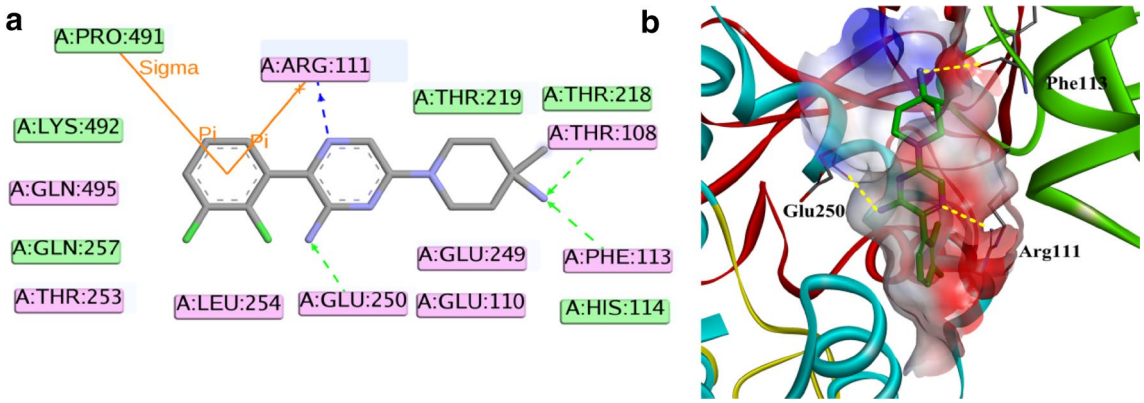


Fig. 3 **a** The ligand–protein interaction diagram of SHP099 and SHP2 (PDB ID: 5EHR). **b** The molecular surface is shown around the binding site of SHP2 (PDB ID: 5EHR). Hydrogen-bond interactions with main-amino acid residues are represented by a green dashed arrow directed towards pointing to the electron donor. Hydrogen-bond interactions with side-amino acid residues are represented by a blue dashed arrow directed towards pointing to the electron donor. Pi interactions are represented by an orange line. Residues involved in hydrogen-bond, charge or polar interactions are represented by pink rectangles. Residues involved in van der Waals interaction are represented by green rectangle

$NaHCO_3$ solution in water, and then extracted with CH_2Cl_2 (2 × 50 mL). The combined organic phases were washed with water (2 × 50 mL), dried over anhydrous Na_2SO_4, filtered, and concentrated in vacuo. Finally, the residue obtained was purified by chromatography to gain the compound 3d. The same method gained compounds 3e, 3f.

Table 1 The CDOKER_ENERGY between ligands and receptors

Compounds	SHP2-CDOKER_ENERGY (kcal/mol)	SHP1-CDOKER_ENERGY (kcal/mol)
SHP099	40.9327	20.3562
1	46.7889	18.5326
2	43.4768	17.6859
3	42.9263	17.513
4	41.6485	15.3245
5	41.516	13.2356
6	38.256	13.0125
7	37.655	12.9586
8	37.253	11.2564
9	35.153	11.1325
10	34.826	10.0248

Compound 4 g (10 mmol), pyridine (30 mmol), 60 mL CH_2Cl_2 were added to the flask and stirred for 5 min. Then, Compound 5i (11 mmol) was added dropwise under ice bath, and stirring was continued for 10 min after the addition was completed, the ice bath was removed and the reaction was continued at room temperature for 1 h. The reaction was completed by TLC analysis. The reaction mixture was poured into water (60 mL) and extracted with CH_2Cl_2 (3 × 50 mL). The combined organic layers were washed with 5% brine (2 × 50 mL) and dried over anhydrous Na_2SO_4. The organic phase after drying was evaporated in vacuo, and the obtained residue was purified by column chromatography to gain product 6k. The same method gained compounds 6l, 6m, 6n.

Compound 3a (4 mmol), Compound 6n (4.4 mmol), $CsCO_3$ (8 mmol) and 20 mL DMF were added to the flask, and the mixture was stirred for 8 h at 80 °C. TLC showed the reaction was completed. After the system was cooled to room temperature, the reaction mixture was poured into water (30 mL) and extracted with CH_2Cl_2 (30 mL × 3). The combined organic layers were washed with 5% brine and dried over anhydrous Na_2SO_4. The organic phase after drying was evaporated in vacuo, and the obtained residue was purified by chromatography to gain the final compound. By this method shown in Fig. 4, compounds 1–10 were obtained.

SHP2 activity assay and fluorescence titrations

Initially, the inhibition of 10 compounds against SHP2 was evaluated by SHP2 activity assay. In those compounds, compound 1 showed an inhibition rate larger than 50% at the concentration of 2 μM. Then, all compounds were tested for half-maximal inhibitory concentrations (IC_{50}) values. All compounds showed IC_{50} values more than 10 μM except compound 1. The compound 1 showed inhibition activity with an IC_{50} value of 9.97 μM (Fig. 5a, Table 2).

As shown in Fig. 5b, compound 1 exhibited strong quenching of SHP2 fluorescence in a dose-dependent manner. These fluorescence quenching experiments implied that compound 1 bound directly to SHP2, thereby attenuating its activity on the substrates.

Cell viability assay

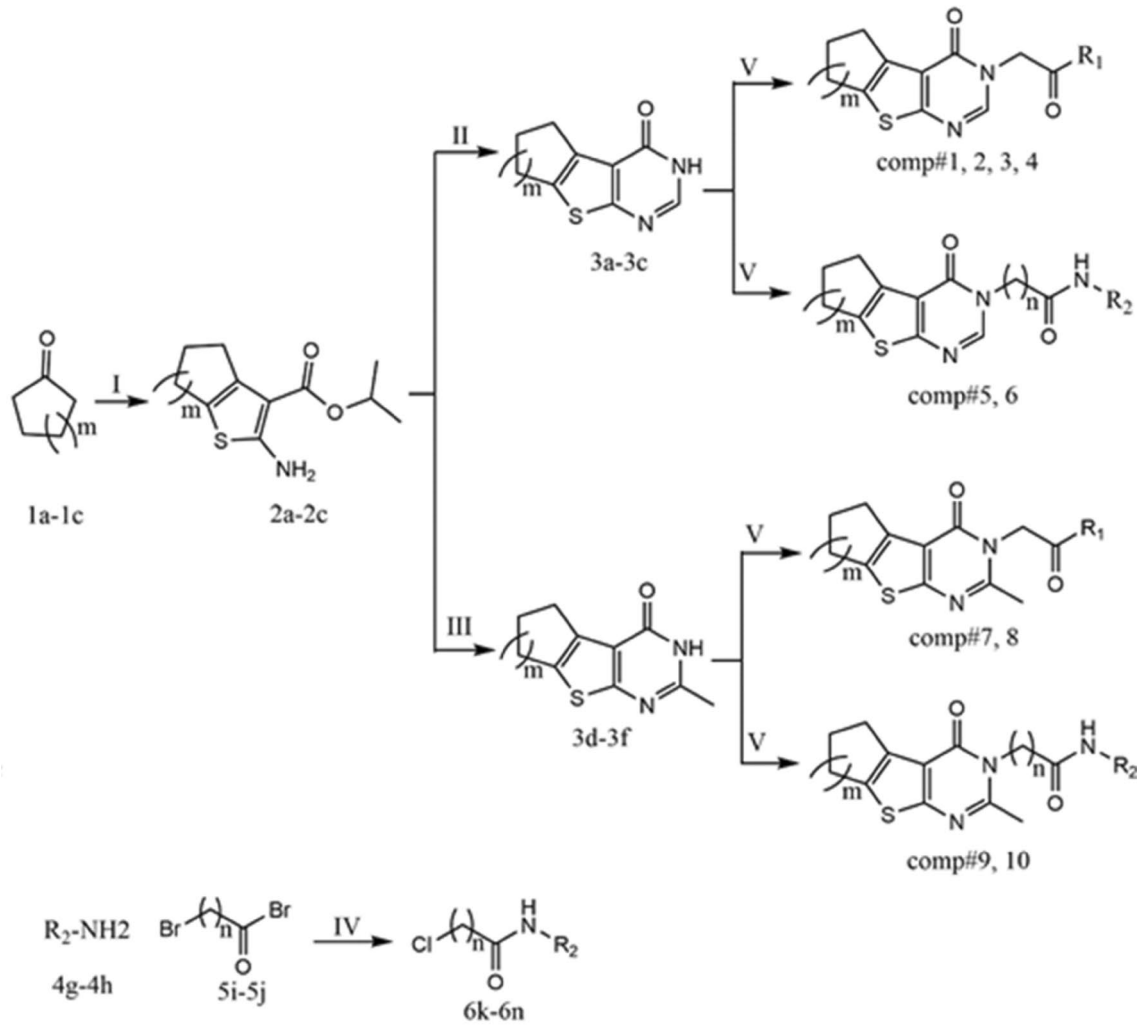
The MTS assay was conducted on BaF3 cell lines. Firstly, the cytotoxicity of 10 selected compounds was estimated at 20 μM concentration in vitro. Secondly, these compounds were further evaluated to determine their IC_{50} values at concentrations of 0.1, 1.0, 5.0, 10 and 20 μM. In the end, compound 1 showed IC_{50} value of 10.73 μM on BaF3 cell line. IC_{50} values were calculated using Graph-Pad Prism 7, and the IC_{50} values of ten compounds were attached in Table 3.

Molecular dynamics trajectory analysis

Molecular dynamics can provide useful information for characterizing the internal motions of bio-macromolecules over time. For comparison study, the 50 ns molecular dynamics simulations were performed for the complexes formed by SHP2 (5EHR) and SHP1 (3PS5) proteins with ligand SHP836 by gromacs 4.5.5.

The RMSD versus the simulation time was considered as a significant criterion to evaluate the stability of dynamic behavior. Initially, the RMSD values of both complexes increased quickly, which was caused by the relaxation and repulsion of the complexes after dissolution of the solution. As we could see from Figs. 6a and 7a, all the characters concerned reached the simulation equilibrium within the 3 ns. During the simulation, the RMSD values of SHP2 with ligand SHP836 and compound 1 were found to be relatively stable about 3.9 Å and 3.7 Å, respectively, while the RMSD values of SHP1 with ligands SHP836 and compound 1 were found to be relatively stable about 0.5 Å and 4.8 Å, respectively. The RMSD curves of SHP2–comp#1 and SHP1–comp#1 were more stable than that of SHP2–SHP836, SHP2-without ligand, SHP1–SHP836, and SHP1-without ligand systems.

The root mean square fluctuation (RMSF) was another useful method to study the stability of systems, reflecting the mobility of certain amino acid residues around their average positions. The greater the fluctuation of the RMSF, the more unstable the residues, and the smaller the fluctuation of the RMSF, the more stable the residues. The RMSF fluctuation of the amino acids in the SHP2–comp#1 complex were close to that of the SHP2–SHP836 complex, indicating that the compound 1 had a similar stability function in SHP2 as SHP836 did (Fig. 6b). As Fig. 6b demonstrated



1a, 2a, 3a, 3d: m = 1
1b, 2b, 3b, 3e: m = 2
1c, 2c, 3c, 3f: m = 3
n = 1, 2
4g: R_2 = 2-Py
4h: R_2 = 2-$PhNO_2$
5i: n = 1
5j: n = 2
6k: n = 1, R_2 = 2-Py
6l: n = 1, R_2 = 2-$PhNO_2$
6m: n = 2, R_2 = 2-Py
6n: n = 2, R_2 = 2-$PhNO_2$

comp#1: m = 2, R_1 = 2,4-di-Cl-Ph
comp#2: m = 2, R_1 = 4-Cl-Ph
comp#3: m = 3, R_1 = 3,4-di-Cl-Ph
comp#4: m = 1, R_1 = 3,4-di-Cl-Ph
comp#5: m = 3, n = 2, R_2 = 2-$PhNO_2$
comp#6: m = 1, n = 2, R_2 = 2-$PhNO_2$
comp#7: m = 2, R_1 = 2,4-di-Cl-Ph
comp#8: m = 1, R_1 = 4-Cl-Ph
comp#9: m = 1, n = 1, R_2 = 2-Py
comp#10: m = 1, n = 2, R_2 = 2-$PhNO_2$

Fig. 4 Synthesis of the Compounds [a]Reagents and conditions: (I) Isopropyl Cyanoacetate, Cyclopentanone, Sulphur Powder, Morpholine, ethanol, N_2, 80 °C, 24 h; (II) Formamid, N_2, 210 °C, 12 h; (III) MeCN, HCl, 60 °C, 12 h; (IV) Pyridine, CH_2Cl_2, 0 °C ~ 25 °C, 1.5 h; (V) $CsCO_3$, acetone, 80 °C, 8 h

the RMSF of key residues Arg111–S8, Phe113–Cl25, and Glu250–O14 in the SHP2–comp#1 complex fluctuated lower than SHP2–SHP836 complex and SHP2 system, reflecting that compound 1 could form stronger hydrogen bond (H bond) interactions with these key residues than SHP836. For SHP1, the RMSF fluctuations of the

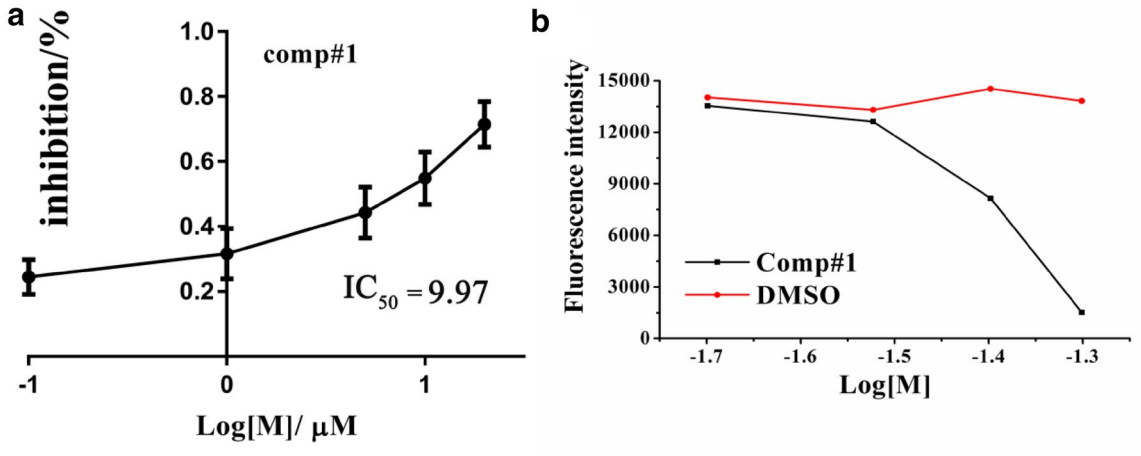


Fig. 5 **a** Inhibition of SHP2 after treated with compound 1 at concentrations of 0.1, 1, 5, 10, 20 μM. **b** Fluorescence titration of SHP2 was performed by increasing the concentrations of compound 1 while maintaining the SHP2 protein concentration (1.0 μg). The fluorescence is plotted against the log concentration μmol/L (log [M]) for the compound. Experiments are repeated three times. Similar results are obtained in each. Data shown are the mean ± SEM of triplicates from one representative experiment

Table 2 The IC_{50} values of 10 compounds

Compound IC_{50} (μM)	1	2	3	4	5	6	7	8	9	10
SHP2	9.97	12.33	39.43	> 100	> 100	28.44	> 100	> 100	76.7	> 100
SHP1	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100

Table 3 The IC_{50} values of ten compounds

Compounds	1	2	3	4	5	6	7	8	9	10
IC_{50} (μM)	10.73	11.59	37.49	>100	>100	16.29	>100	>100	30.92	48.55

residues in the SHP1–comp#1 complex were similar to the SHP1–SHP836 complex, indicating that the compound 1 had a similar inhibitory function as SHP836 (Fig. 7b). Figure 7b also revealed that the RMSF values of key residues in the SHP2–comp#1 complex were lower than that of SHP1–comp#1, SHP1–SHP836 and SHP1-without ligand systems, indicating that the compound 1 formed stronger interactions with the residues than SHP836. Furthermore, compared to the complexes of SHP1, the complexes of SHP2 were more stable.

To understand the specific interactions between protein systems, the binding free energy was calculated using MM/PBSA consisting of four terms, including the van der waals interaction energy, the electrostatic energy, the polar solvation free energy, and the non-polar solvation free energy. For all systems, each snapshot structures were extracted during equilibrium phase in the 10–50 ns trajectory to calculate the binding free energy.

The binding free energies and detailed contributions of the four energy components obtained from the MM/PBSA calculation of the protein–ligand complexes were listed in Table 4. It was clear that SHP2–comp#1 complex (− 1603.625 J/mol) showed more negative binding free energy than SHP2–SHP836 (− 256.936 kJ/mol), SHP1–SHP836 (2.411 kJ/mol) and SHP1–comp#1 complexes (− 194.569 kJ/mol), which supported the result of MD simulations, indicating that SHP2–comp#1 complex was more stable. Comparison of the binding energy components provided a better understanding of the relative contributions of different components to the overall binding free energies. It was apparent that electrostatic interactions, Van der Waals and non-polar solvation energy had a negative contribution to the total interaction energy, and only polar solvation energy had a positive contribution to the total free binding energy, indicating that electrostatic interactions, Van der Waals and non-polar solvation energy together contributed to the stability of the SHP2–comp#1, SHP2–SHP836, SHP1–comp#1, and SHP1–SHP836 complexes. For negative contribution, electrostatic interactions offered greater contributions than Van der Waals interactions and the nonpolar free energy in the SHP2–comp#1 system, while for SHP2–SHP836 system, electrostatic

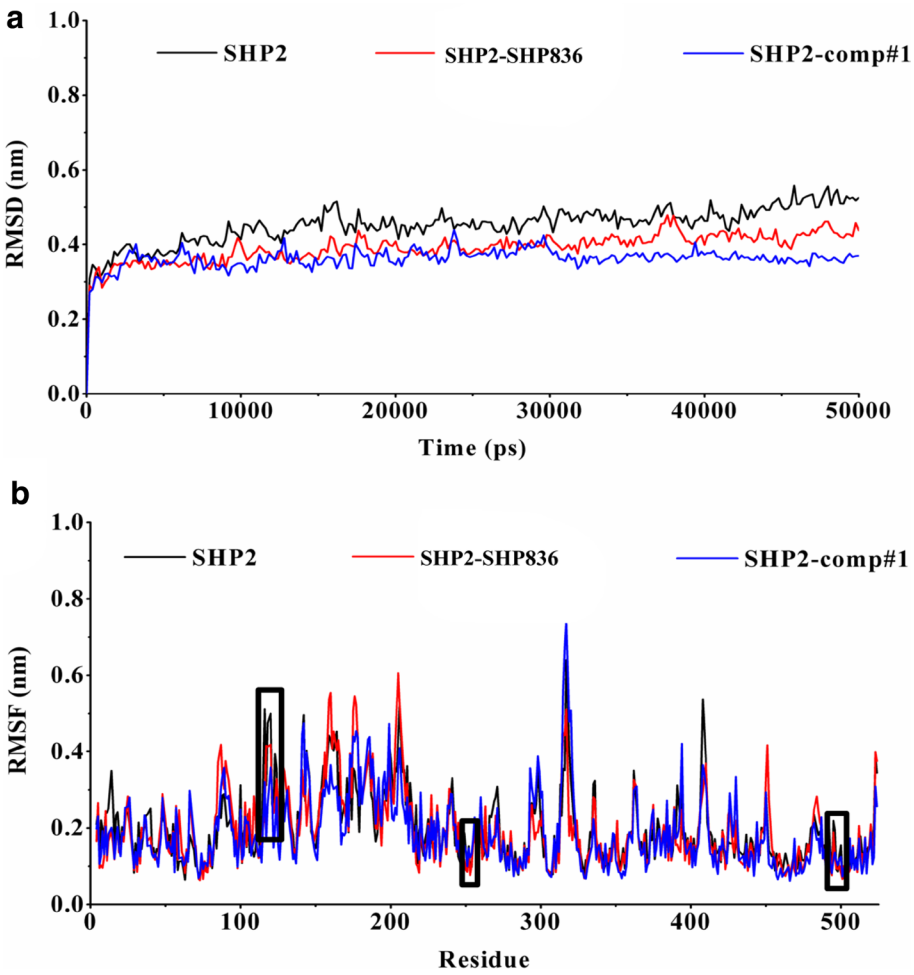


Fig. 6 Analysis of molecular dynamics simulations. **a** The RMSD of all backbone atoms for the receptor SHP2. **b** The RMSF of the side-chain atoms for the receptor SHP2. The black line indicates the outcome for the system of the receptor without any ligand; the red line indicates the outcome for the system of the receptor with ligand SHP836; and the blue line indicates the outcome for the system of the receptor with compound 1. The black rectangular box indicates the fluctuation of key residues

interactions possessed the largest proportion among the three components.

To further elucidate the effects of key residues on the binding free energies of each complex, the decomposition in terms of each residue was performed (Fig. 8). In general, a residue would be considered as more important if its decomposed energy was more negative. The key residues were observed in the significantly different regions of SHP2 (Thr108–His114, Leu216–Thr219, Glu249–Gln257, and Asp489–Gln495) and SHP1 (Pro212–Thr216, Glu247–Gln254, and Asp483–Gln489), which was in accordance with the docking results in Fig. 3. The major residues of SHP2–comp#1 and SHP2–SHP836 complexes contributed to the binding free energy varied in the range from − 206.436 to 39.614 kJ/mol and − 207.494 to 33.302 kJ/mol, respectively, while the major residues of SHP1–comp#1 and SHP1–SHP836 complexes contributed to the binding free energy varied in the range from − 109.584 to 124.2927 kJ/mol and − 363.4859 to 226.8984 kJ/mol, respectively.

For SHP2–SHP836 complex, the key residues were Thr108, Glu110, Arg111, Phe113, His114, Thr218, Thr219, Glu249, Glu250, Thr253, Leu254, Gln257, Pro491, Lys492, Gln495, while some other residues (Leu216, Asn217, Asp489) played an important role in SHP2–comp#1 system. For SHP1–SHP836 complex, the key residues were Arg109, Lys127, Tyr214, Thr216, Glu246, Ser250, Leu251, Gln254, Asp483, Gln485, Lys486, Gln489, while some other residues (Ala215, Glu235) also played an important role in SHP2–comp#1 complex. The SHP2–comp#1 complex (Fig. 7), the major favorable contributions to the binding free energy came from Glu110 (− 58.8895 kJ/mol), Phe113 (− 9.1562 kJ/mol), His114 (− 19.2202 kJ/mol), Thr218 (− 4.0998 kJ/mol), Thr219 (− 5.0751 kJ/mol), Glu249 (− 180.881 kJ/mol), Glu250 (− 206.436 kJ/mol), Thr253 (− 9.2609 kJ/mol), Leu254 (− 3.2948 kJ/mol) and Pro491 (− 2.6899 kJ/mol). Among those residues, Glu250 gave the most negative decomposed energy than other residues, suggesting the stronger interactions with SHP2. Additionally, residues Pro491and Thr108 formed the hydrophobic active site and made favorable van der Waals/hydrophobic contributions to the total energy. However, Arg111, Thr218, and Lys492 showed positive decomposed energy, which indicated their unfavorable contributions. Compared to SHP2–SHP836 complex, the binding energy of residues

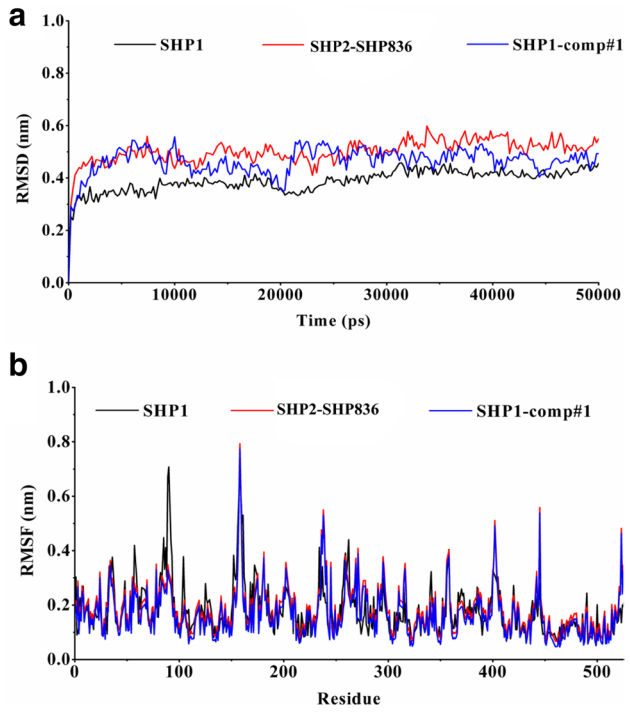


Fig. 7 Analysis of molecular dynamics simulations. **a** The RMSD of all backbone atoms for the receptor SHP1. **b** The RMSF of the side-chain atoms for the receptor SHP1. The black line indicates the outcome for the system of the receptor alone without any ligand; the red line indicates the outcome for the system of the receptor with ligand SHP836; and the blue line indicates the outcome for the system of the receptor with compound 1

Glu249 and Glu110 of SHP2–comp#1 complex showed remarkable decrease, with the binding energy changing from − 180.881 kJ/mol to − 204.725 kJ/mol, and − 58.8895 kJ/mol to − 65.5473 kJ/mol, respectively. The decrease of the binding free energy indicated that the interactions between these residues had been strengthened. One reason was that compound 1 bound to SHP2 increased the electronic interactions between the residues and SHP2 protein. As shown in Fig. 1, the Scaffold A of compound 1 was pointed close to residues Arg111 and Pro491, which was beneficial to form more stable H bond and cation–Pi interactions. The Scaffold B and Scaffold C in compound 1 were located nearly to residues Leu254, Glu249, Thr218, and Thr219 which increased the opportunity to form H bond interactions. Another reason was that the compound 1 formed more hydrophobic interactions with residues Asp489.

For SHP1–comp#1 complex, Tyr214 (− 12.2148 kJ/mol), Thr216 (− 0.6229 kJ/mol), Ser250 (− 3.9602 kJ/mol), Leu251 (− 1.5779 kJ/mol), Gln254 (− 3.1456 kJ/mol) and Asp483 (− 53.1966 kJ/mol) were responsible for major favorable contributions to the binding free energy (Fig. 8). Major contribution to the van der Waals energies came from residues Leu251, Asp483 and Gln489, while the major contribution of electrostatic energy came from residues Arg109. The above findings provided suitable explanation for the different binding modes of compound 1 in the SHP2 and SHP1 proteins. Furthermore, compound 1 in both complexes exhibited similar hydrophobic interactions with nearby residues, especially three conserved residues (Ser109, Ser250 and Gln254 in SHP1 and Ser109, Glu250 and Leu250 in SHP2). Compared to SHP1–SHP836 and SHP1–comp#1 complexes, the binding energy of all residues of SHP2–comp#1 complex showed remarkable decrease.

The analysis of H-bond occupancy

The stability of the three-dimensional structure of a protein was decided by the subtle balance among all kinds of weak interactions, such as conjugation interactions, hydrophobic interactions, and H bonds, in which hydrogen bonds play the most important role in stabilizing the system.

The occupancies of H bond interactions in the 50 ns simulations of SHP2–SHP836, SHP2–comp#1, SHP1–SHP836 and SHP1–comp#1 complexes were calculated and listed in Table 5, respectively. The H bonds formed between Phe113(O–N22), Glu250(O–N7), and Arg111(NH1–N2) occupied 62%, 46%, and 2.8% of the whole simulation time in SHP2–SHP836 system, respectively, whereas the three H bonds formed between Phe113(O–Cl25), Glu250(O–O14), and Arg111(HH26–S7) occupied 97.2%, 87.6%, and 20% of the total simulation time in SHP2–comp#1 system, respectively. It was revealed that these H bonds formed between these residues were unstable in SHP2–comp#1 system. Apart from these stable H bonds formed in SHP2–comp#1 system, there were also two other H bonds formed around it. In SHP2–comp#1 system, the H bond formed between

Table 4 Binding free energies (kJ/mol) and its components between protein and ligand

Complex	Van der Waal energy	Electrostattic energy	Polar solvation energy	Non-polar energy	Binding energy
SHP2–SHP836	− 146.848	− 1314.676	1225.359	− 20.771	− 256.936
SHP2–comp#1	− 145.798	− 2741.876	1302.632	− 18.583	− 1603.625
SHP1–SHP836	− 176.150	− 936.630	1135.911	− 18.720	2.411
SHP1–comp#1	− 200.556	− 807.259	832.354	− 19.109	− 194.569

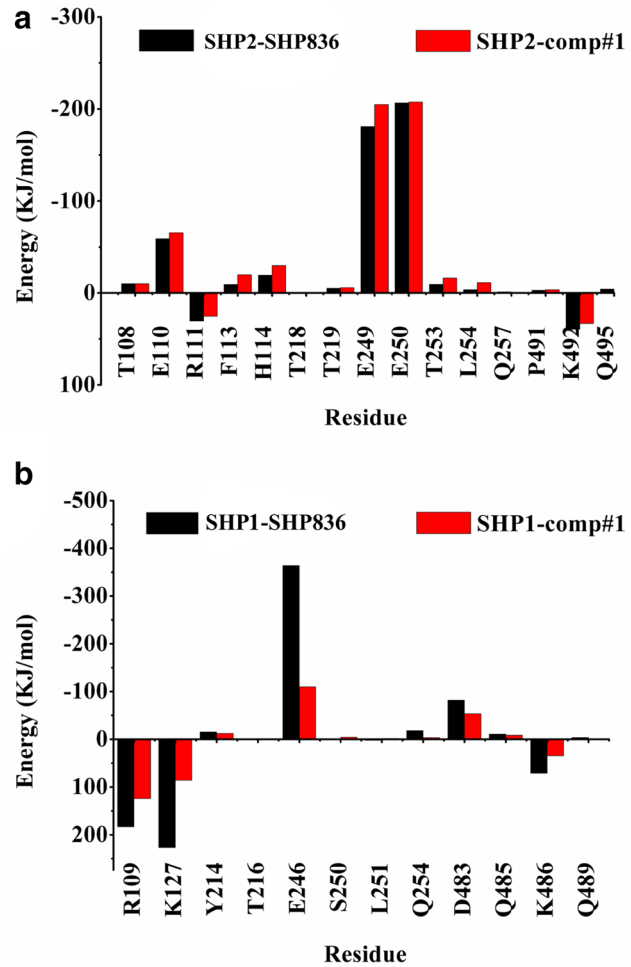


Fig. 8 **a** Interaction energies between SHP2 and SHP836 (black) and comp#1 (red), respectively. Black represents SHP2–SHP836 complex, red represents SHP2–comp#1 complex. **b** Interaction energies of SHP1 with SHP836 (black) and compound 1 (red), respectively. Black represents SHP1–SHP836 complex, red represents SHP1–comp#1 complex

Thr218 (H–O18) and compound 1 occupied 100% of the whole simulation time. The results were well in accordance with the analysis of flexible docking, revealing that the distances between Arg111 and N2 of SHP836 was moved out of the reasonable H bond distance (0.25–0.35 nm) and the distances between Phe113 (O–N22) and Glu250 (O–N7) were within the H bond distance in SHP2–SHP836 system. Since the distances of these H bonds (Phe113–O16, Glu250–O16, and Arg111–Br8, respectively) were remained within H bond distance throughout the whole simulation, indicating that compound 1 had higher binding energy than SHP836.

Compared to SHP2–comp#1, the occupancy of H bond interactions of Gln485 were only 16% and 15.2% in SHP1–SHP836 and SHP1–comp#1 complexes, respectively, indicating that the binding energy of compound 1 to SHP2 was higher than that of compound 1 to SHP1. In conclusion, the H bond interactions between compound 1 and SHP2 were more stable than that of compound 1 and SHP1, and the occupancies of H bond interactions between compound 1 and SHP2 were higher, indicating that compound 1 had higher selectivity for SHP2.

Dynamic cross-correlation

DCC analysis was performed as described in previous section to identify the presence of correlated motions of Cα atoms within the backbone of the SHP2 and SHP1 systems obtained after MD simulations. The DCC results for the SHP2 and SHP1 systems were shown in Fig. 9. The DCC results clearly show that the active site residues of both the complexes demonstrated correlated motion which proves that was bound to the active sites of both SHP2 and SHP1. The results of DCC showed that only SHP2–comp#1 complex demonstrated the strongly correlated motion between residues, while SHP1–comp#1 complex did not show the obviously correlated motion between residues. This indicated that compound 1 was more suitable for binding to SHP2 but not to SHP1 and confirmed the selectivity of compound 1 to SHP2, which further supported the results of docking, MD simulation and binding energy calculations.

Table 5 The occupancies of hydrogen bond interactions in the 50 ns simulations of SHP2–SHP836, SHP2–comp#1, SHP1–SHP836 and SHP1–comp#1 complexes, respectively

Interaction type	Atom1	Atom2	Atom3	Occupancy (%) (SHP2–SHP836)	Occupancy (%) (SHP2–comp#1)	Occupancy (%) (SHP1–SHP836)	Occupancy (%) (SHP1–comp#1)
Hydrogen bond	Arg111(NH1–N2)	Arg111(HH26–S7)	Gln485	62	97.2	15.2	16
Hydrogen bond	Glu250(O–N7)	Phe113(O–Cl25)	0	46	87.6	0	0
Hydrogen bond	Phe113(O–N22)	Glu250(O–O14)	0	2.8	20	0	0
Hydrogen bond	0	Thr218 (H–O18)	0	0	100	0	0

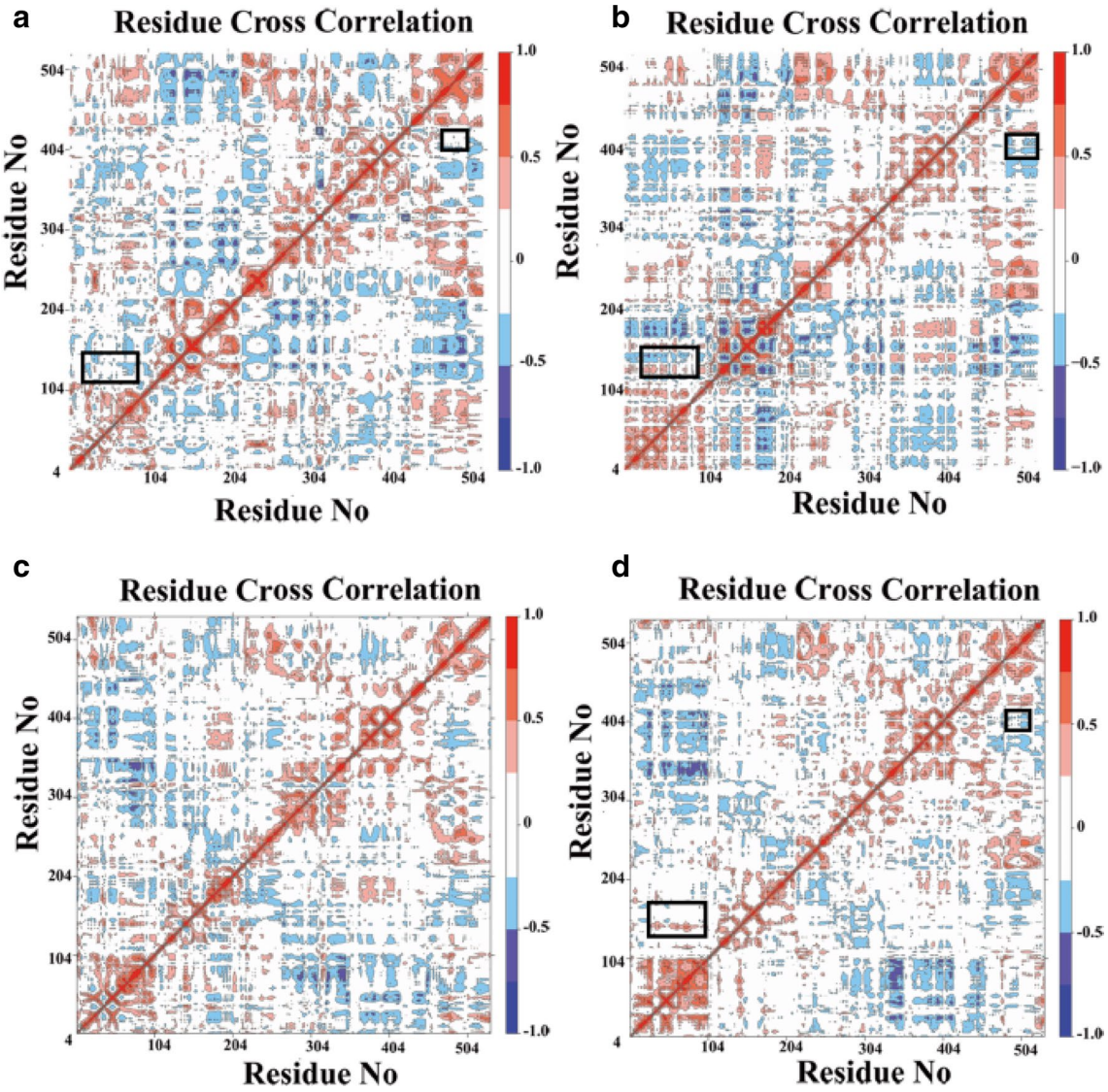


Fig. 9 Dynamic cross-correlation map (DCCM). **a** DCCM map for the SHP2–SHP836 complex shows positive and negative correlative motions between the residues. **b** DCCM map for SHP2–comp#1 complex shows positive and negative correlative motions between residues. **c** DCCM map for the SHP1–SHP836 complex showed positive and negative correlative motions between the residues. **d** DCCM map for SHP1–comp#1 showed positive and negative correlative motions between residues. Red represents positive correlations, and blue represents negative correlations

Conclusion

The aim of our study was to find new and highly selective allosteric inhibitors of SHP2. We utilized computer-aided drug design method to rapidly screen optimal ligands. From De novo design study, it was found that 130 compounds might be potential SHP2 inhibitors. It was validated by Lipinski's Rule of Five and ADMET prediction. Compound 1 was found to have a novel selectivity for SHP2 with an IC_{50} value of 9.97 μM in enzyme activity assay and 10.73 μM in cell proliferation assay. Fluorescence titration experiment confirmed that compound 1 directly bound to SHP2. Consequently, the identified compound could be regarded as a novel drug candidate or lead compound to inhibit SHP2. Finally, compound 1 was used as a representative for 50 ns molecular dynamic simulations to study the binding affinity of the protein–ligand complex and selectivity to SHP2. The compound 1 had a significantly higher affinity for SHP2 than SHP1, suggesting that it was a promising selective SHP2 inhibitor. The interactions of compound 1 with residues Phe113 and Glu250 in the SHP2 were identified as the main responsible for determining selectivity. Compared to the SHP836, the new inhibitor not only had the similar function of

inhibiting SHP2, but also could form a more stable conformation after binding to SHP2 protein.

Acknowledgements This study was supported by the Natural Science Foundations of China (Grant No. 81273361), the Natural Science Foundation of Tianjin (Grant No. 16JCZDJC32500) and the International (Regional) 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).

Compliance with ethical standards

Conflict of interest The authors report no conflicts of interest in this work.

References

1. Hof P, Pluskey S, Dhe-Paganon S, Eck MJ, Shoelson SE (1998) Crystal structure of the tyrosine phosphatase SHP-2. Cell 92(4):441–450
2. Chen C, Cao M, Zhu S, Wang C, Liang F, Yan L, Luo D (2015) Discovery of a novel inhibitor of the protein tyrosine phosphatase Shp2. Sci Rep 5:17626. https://doi.org/10.1038/srep17626
3. Zhang X, He Y, Liu S, Yu Z, Jiang ZX, Yang Z, Dong Y, Nabinger SC, Wu L, Gunawan AM, Wang L, Chan RJ, Zhang ZY (2010) Salicylic acid based small molecule inhibitor for the oncogenic Src homology-2 domain containing protein tyrosine phosphatase-2 (SHP2). J Med Chem 53(6):2482–2493. https://doi.org/10.1021/jm901645u
4. Yang W, Wang J, Moore DC, Liang H, Dooner M, Wu Q, Terek R, Chen Q, Ehrlich MG, Quesenberry PJ, Neel BG (2013) Ptpn11 deletion in a novel progenitor causes metachondromatosis by inducing hedgehog signalling. Nature 499(7459):491–495. https://doi.org/10.1038/nature12396
5. Lawrence HR, Pireddu R, Chen L, Luo Y, Sung SS, Szymanski AM, Yip ML, Guida WC, Sebti SM, Wu J, Lawrence NJ (2008) Inhibitors of Src homology-2 domain containing protein tyrosine phosphatase-2 (Shp2) based on oxindole scaffolds. J Med Chem 51(16):4948–4956. https://doi.org/10.1021/jm8002526
6. Chan RJ, Feng GS (2007) PTPN11 is the first identified proto-oncogene that encodes a tyrosine phosphatase. Blood 109(3):862–867. https://doi.org/10.1182/blood-2006-07-028829
7. Bentires-Alj M, Paez JG, David FS, Keilhack H, Halmos B, Naoki K, Maris JM, Richardson A, Bardelli A, Sugarbaker DJ, Richards WG, Du J, Girard L, Minna JD, Loh ML, Fisher DE, Velculescu VE, Vogelstein B, Meyerson M, Sellers WR, Neel BG (2004) Activating mutations of the noonan syndrome-associated SHP2/PTPN11 gene in human solid tumors and adult acute myelogenous leukemia. Cancer Res 64(24):8816–8820. https://doi.org/10.1158/0008-5472.CAN-04-1923
8. Chen L, Sung SS, Yip ML, Lawrence HR, Ren Y, Guida WC, Sebti SM, Lawrence NJ, Wu J (2006) Discovery of a novel shp2 protein tyrosine phosphatase inhibitor. Mol Pharmacol 70(2):562–570. https://doi.org/10.1124/mol.106.025536
9. He R, Yu ZH, Zhang RY, Wu L, Gunawan AM, Lane BS, Shim JS, Zeng LF, He Y, Chen L, Wells CD, Liu JO, Zhang ZY (2015) Exploring the existing drug space for novel pTyr mimetic and SHP2 inhibitors. ACS Med Chem Lett 6(7):782–786. https://doi.org/10.1021/acsmedchemlett.5b00118
10. Fortanet JG, Chen CHT, Chen YNP, Chen ZL, Deng Z, Firestone B, Fekkes P, Fodor M, Fortin PD, Fridrich C, Grunenfelder D, Ho S, Kang ZB, Karki R, Kato M, Keen N, LaBonte LR, Larrow J, Lenoir F, Liu G, Liu SM, Lombardo F, Majumdar D, Meyer MJ, Palermo M, Perez L, Pu MY, Ramsey T, Sellers WR, Shultz MD, Stams T, Towler C, Wang P, Williams SL, Zhang JH, LaMarche MJ (2016) Allosteric inhibition of SHP2: identification of a potent, selective, and orally efficacious phosphatase inhibitor. J Med Chem 59(17):7773–7782. https://doi.org/10.1021/acs.jmedchem.6b00680
11. Chen C, Liang F, Chen B, Sun Z, Xue T, Yang R, Luo D (2017) Identification of demethylincisterol A3 as a selective inhibitor of protein tyrosine phosphatase Shp2. Eur J Pharmacol 795:124–133. https://doi.org/10.1016/j.ejphar.2016.12.012
12. Barr AJ (2010) Protein tyrosine phosphatases as drug targets: strategies and challenges of inhibitor development. Future Med Chem 2(10):1563–1576. https://doi.org/10.4155/Fmc.10.241
13. Chen YN, LaMarche MJ, Chan HM, Fekkes P, Garcia-Fortanet J, Acker MG, Antonakos B, Chen CH, Chen Z, Cooke VG, Dobson JR, Deng Z, Fei F, Firestone B, Fodor M, Fridrich C, Gao H, Grunenfelder D, Hao HX, Jacob J, Ho S, Hsiao K, Kang ZB, Karki R, Kato M, Larrow J, La Bonte LR, Lenoir F, Liu G, Liu S, Majumdar D, Meyer MJ, Palermo M, Perez L, Pu M, Price E, Quinn C, Shakya S, Shultz MD, Slisz J, Venkatesan K, Wang P, Warmuth M, Williams S, Yang G, Yuan J, Zhang JH, Zhu P, Ramsey T, Keen NJ, Sellers WR, Stams T, Fortin PD (2016) Allosteric inhibition of SHP2 phosphatase inhibits cancers driven by receptor tyrosine kinases. Nature 535(7610):148–152. https://doi.org/10.1038/nature18621
14. Veselovsky AV, Ivanov AS (2003) Strategy of computer-aided drug design. Curr Drug Targets Infect Disord 3(1):33–40
15. Wang W, Liu LJ, Song X, Mo Y, Komma C, Bellamy HD, Zhao ZJ, Zhou GW (2011) Crystal structure of human protein tyrosine phosphatase SHP-1 in the open conformation. J Cell Biochem 112(8):2062–2071. https://doi.org/10.1002/jcb.23125
16. Prlic A, Kalro T, Bhattacharya R, Christie C, Burley SK, Rose PW (2016) Integrating genomic information with protein sequence and 3D atomic level structure at the RCSB protein data bank. Bioinformatics 32(24):3833–3835. https://doi.org/10.1093/bioinformatics/btw547
17. Ma Y, Jin YY, Wang YL, Wang RL, Lu XH, Kong DX, Xu WR (2014) The discovery of a novel and selective inhibitor of PTP1B over TCPTP: 3D QSAR pharmacophore modeling, virtual screening, synthesis, and biological evaluation. Chem Biol Drug Des 83(6):697–709. https://doi.org/10.1111/cbdd.12283
18. Schubert CR, Stultz CM (2009) The multi-copy simultaneous search methodology: a fundamental tool for structure-based drug design. J Comput Aided Mol Des 23(8):475–489. https://doi.org/10.1007/s10822-009-9287-y
19. Bohm HJ (1992) The computer program LUDI: a new method for the De novo design of enzyme inhibitors. J Comput Aided Mol Des 6(1):61–78
20. Schneider G, Fechner U (2005) Computer-based De novo design of drug-like molecules. Nat Rev Drug Discov 4(8):649–663. https://doi.org/10.1038/nrd1799
21. Egan WJ, Merz KM, Baldwin JJ (2000) Prediction of drug absorption using multivariate statistics. J Med Chem 43(21):3867–3877. https://doi.org/10.1021/jm000292e
22. Egan WJ, Walters WP, Murcko MA (2002) Guiding molecules towards drug-likeness. Curr Opin Drug Discov Dev 5(4):540–549
23. Cheng A, Merz KM 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
24. 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

25. Xia XY, Maliski EG, 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
26. Koska J, Spassov VZ, Maynard AJ, Yan L, Austin N, Flook PK, Venkatachalam CM (2008) Fully automated molecular mechanics based induced fit protein-ligand docking method. J Chem Inf Model 48(10):1965–1973. https://doi.org/10.1021/ci800081s
27. Spassov VZ, Yan L, Flook PK (2007) The dominant role of side-chain backbone interactions in structural realization of amino acid code. ChiRotor: a side-chain prediction algorithm based on side-chain backbone interactions. Protein Sci 16(3):494–506. https://doi.org/10.1110/ps.062447107
28. Wu G, Robertson DH, Brooks CL 3rd, Vieth M (2003) Detailed analysis of grid-based molecular docking: a case study of CDOCKER-A CHARMm-based MD docking algorithm. J Comput Chem 24(13):1549–1562. https://doi.org/10.1002/jcc.10306
29. Abdolmaleki A, Ghasemi F, Ghasemi JB (2017) Computer-aided drug design to explore cyclodextrin therapeutics and biomedical applications. Chem Biol Drug Des 89(2):257–268. https://doi.org/10.1111/cbdd.12825
30. Fernandes CL, Sachett LG, Pol-Fachin L, Verli H (2010) GROMOS96 43a1 performance in predicting oligosaccharide conformational ensembles within glycoproteins. Carbohydr Res 345(5):663–671. https://doi.org/10.1016/j.carres.2009.12.018
31. Zang LL, Wang XJ, Li XB, Wang SQ, Xu WR, Xie XB, Cheng XC, Ma H, Wang RL (2014) SAHA-based novel HDAC inhibitor design by core hopping method. J Mol Gr Model 54:10–18. https://doi.org/10.1016/j.jmgm.2014.08.005
32. Hess B (2008) P-LINCS: a parallel linear constraint solver for molecular simulation. J Chem Theory Comput 4(1):116–122. https://doi.org/10.1021/ct700200b
33. Kumari R, Kumar R, Lynn A, Consort OSDD (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
34. Hou TJ, Wang JM, Li YY, Wang W (2011) Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized born surface area methods. II. The accuracy of ranking poses generated from docking. J Comput Chem 32(5):866–877. https://doi.org/10.1002/jcc.21666
35. Weiser J, Shenkin PS, Still WC (1999) Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). J Comput Chem 20(2):217–230. https://doi.org/10.1002/(Sici)1096-987x(19990130)20:2%3c217:Aid-Jcc4%3e3.0.Co;2-A
36. Onufriev A, Bashford D, Case DA (2004) Exploring protein native states and large-scale conformational changes with a modified generalized born model. Proteins Struct Funct Bioinform 55(2):383–394. https://doi.org/10.1002/prot.20033
37. Tama F, Sanejouand YH (2001) Conformational change of proteins arising from normal mode calculations. Protein Eng 14(1):1–6. https://doi.org/10.1093/Protein/14.1.1
38. Tang BW, Li BQ, Qian YQ, Ao MT, Guo KQ, Fang MJ, Wu Z (2017) The molecular mechanism of hPPAR alpha activation. RSC Adv 7(28):17193–17201. https://doi.org/10.1039/c6ra27740c
39. Kasahara K, Fukuda I, Nakamura H (2014) A novel approach of dynamic cross correlation analysis on molecular dynamics simulations and its application to Ets1 dimer-DNA complex. PLoS ONE 9(11):e112419. https://doi.org/10.1371/journal.pone.0112419
40. Cole JC, Murray CW, Nissink JWM, Taylor RD, Taylor R (2005) Comparing protein-ligand docking programs is difficult. Proteins Struct Funct Bioinform 60(3):325–332. https://doi.org/10.1002/prot.20497