

文章编号 1672-6634(2023)05-0099-12

DOI 10.19728/j.issn1672-6634.2023040001

虚拟筛选结合活性验证发现新型 SHP2 催化位点抑制剂

黎娅琳,任吉霞

(聊城大学 生命科学学院,山东 聊城 252059)

摘要 原癌蛋白 SHP2 是一种蛋白酪氨酸磷酸酶,其在多种恶性血液病和实体瘤中异常活跃表达。所以,研发靶向 SHP2 的抗肿瘤小分子抑制剂已成为当前研究的热点。通过基于 SHP2 抑制剂和非抑制剂的贝叶斯分类模型,药效团模型,及分子对接的组合虚拟筛选技术联合酶水平活性验证,发现 3 个新型 SHP2 催化位点抑制剂。因蛋白酪氨酸磷酸酶催化位点结构高度保守,靶向催化位点抑制剂一般选择性较差。检测了活性最高化合物 cpd2 在 10 $\mu\text{mol/L}$ 浓度下对 SHP2, SHP1, PTP-1B 和 DUSP22 的抑制作用,检测结果提示该化合物具有一定潜在选择性。通过分析化合物 cpd2 与药效团的匹配结果,及分子对接技术预测的 cpd2 与 SHP2 催化活性位点间相互作用,阐明该化合物抑制 SHP2 催化活性的作用机制。所得结果为发现更多高活性和高选择性 SHP2 催化位点抑制剂奠定了新的基础。

关键词 肿瘤;SHP2 抑制剂;虚拟筛选;作用机制

中图分类号 R730

文献标识码 A

开放科学(资源服务)标识码(OSID)



Virtual Screening Combined with Activity Verification Revealed Novel Inhibitors of SHP2 Catalytic Sites

LI Yalin, REN Jixia

(School of Life Sciences, Liaocheng University, Liaocheng 252059, China)

Abstract Abstract Proto-oncoprotein SHP2 is a protein tyrosine phosphatase that is hyperactively expressed in a variety of hematologic malignancies and solid tumors. Therefore, development of anti-tumor small molecule inhibitors targeting SHP2 has become the focus of current research. In this study, three novel inhibitors of SHP2 catalytic sites were identified through a combination of virtual screening techniques based on the Bayesian classification model of SHP2 inhibitors and non-inhibitors, pharmacophores model, and molecular docking combined with enzyme level activity verification. Due to the highly conserved structure of the catalytic site of protein tyrosine phosphatase, the selectivity of the targeted catalytic site compounds was poor. By detecting the inhibitory effect of the highest active compound cpd2 on SHP-1, TPP-1B and DUSP22, it was found that this compound had a potential degree of selectivity. By analyzing the matching results between cpd2 and pharmacophores and the interaction between the active compound cpd2 and

收稿日期:2023-04-02

基金项目:国家自然科学基金青年基金项目(81502963)资助

通讯作者:任吉霞,女,汉族,博士,副教授,研究方向:生信和药物设计, E-mail:renjixia@lcu.edu.cn.

the catalytic active site of SHP2 predicted by molecular docking technology, the mechanism of inhibiting the catalytic activity of SHP2 by cpd2 was clarified. The results of this study lay a new foundation for the discovery of more highly active and selective SHP2 active site inhibitors.

Key words tumor; SHP2 inhibitor; virtual screening; mechanism of action

1 引言

SHP2(SH2 domain-containing protein-tyrosinephosphatase-2)由 *PTPN11* 基因编码,属于非受体蛋白酪氨酸磷酸酶。完整的 SHP2 蛋白晶体结构显示^[1],在本底状态下,SHP2 通过 N-SH2 域和催化位点 PTP 域间形成分子内相互作用而处于一种自抑制的失活构象;在生长因子、细胞因子等外界环境因子的刺激下,SHP2 的失活状态被打破,变成有活性的开放构象。研究表明,SHP2 参与多种信号通路,SHP2 的活化是 Ras/ERK1/2 通路完全激活所必需的,此外,SHP2 活化 PI3K/AKT,JAK/STAT,JNK 和 NF- κ B 信号通路,进而调控细胞增殖、分化、细胞周期维持和迁移等生理功能^[2]。SHP2 还通过其 N-SH2 结构域结合免疫检查点蛋白 PD-1 的磷酸酪氨酸基序(pTyr motif),参与调节 T 细胞的活性^[3]。综上 SHP2 是细胞中多种信号通路的重要信号调节分子,同时 SHP2 还是控制细胞因子产生及免疫细胞反应的重要调控因子。

所以,当 *PTPN11* 基因发生突变而表达异常时,将导致细胞增殖分化异常,引起多种疾病的发生发展。近几年大量遗传学和临床研究表明,在多种恶性血液病和实体瘤中都存在 SHP2 突变活化或过表达^[4]。例如,生殖系获得功能性 *PTPN11* 基因突变会导致 Noonan 综合征(40%~50%),该疾病是一种可增加恶性肿瘤风险的常染色体显性遗传发育障碍症^[5]。SHP2 高活化突变出现于幼年慢性骨髓单核细胞性白血病(JMML,35%),急性 B 淋巴细胞白血病(7%)和急性髓性白血病(4%)^[6,7]。白血病相关 *PTPN11* E76K 突变会促进骨髓增殖性肿瘤的发生和发展^[8]。SHP2 过表达在多种实体瘤中非常常见,例如:在大约 55%原发性乳腺癌中 SHP2 过表达,并会导致预后不良^[9]。最近 Hana Algül 课题组发现 KRAS 突变驱动的胰腺导管腺癌和非小细胞肺癌依赖于 SHP2^[10]。据 Rene Bernards 课题组报道 SHP2 是 KRAS 突变型非小细胞肺癌所必需的^[11]。除此之外,近期研究表明 SHP2 与肿瘤免疫调节密切相关,例如,Zhao 等研究发现^[12],抑制 SHP2 活性可激活抗肿瘤免疫力,并可以与阻断 PD-1 起协同作用,这为肿瘤免疫治疗提供了新的思路。以上研究均提示 SHP2 是一个抗肿瘤新靶标。靶向抑制 SHP2 可双管齐下,既减缓癌细胞生长,同时也调节免疫功能以激活其抗肿瘤作用。

综上,研发 SHP2 小分子抑制剂,对靶向 SHP2 的抗肿瘤治疗及其分子机制研究都具有十分重要的意义。目前已经有一些 SHP2 小分子抑制剂被报道,这些小分子抑制剂主要分为两大类:催化位点抑制剂和变构位点抑制剂^[13,14]。张仲寅课题组开发了多个 SHP2 催化位点抑制剂^[15-17],最近,国内学者王文龙课题组^[18]和王润玲课题组^[19,20]也发现多个新型 SHP2 催化位点抑制剂。除此,还有 4 款小分子变构抑制剂进入临床研究^[21,22]。

计算机辅助药物设计(computer aided drug design,CADD)技术已在药物研发中得到广泛应用,据估计 CADD 使得新药研发周期缩短了 0.9 年,研发费用降低 1.3 亿美元^[23]。根据药物作用靶点三维结构是否已知,CADD 大概可以分为基于配体的药物设计(如药效团模型)和基于结构(如分子对接)的药物设计。研究证明三维药效团模型和分子对接技术的联合应用可以缩短大型化合物数据库虚拟筛选的时间,并可提高命中率^[24]。虚拟筛选获得化合物的活性基本上在 $\mu\text{mol/L}$ 水平,有少量的化合物达到 nmol/L 水平,这一活性水平的化合物可以作为先导化合物进行结构改造,以获得活性高的类药性化合物^[25]。计算机辅助药物分子设计作为创新药物研究的新方法和新技术,已经成为一种实用化工具,加入到了创新药物研究的工作流程中,成为现代药物研发的核心技术之一。

本研究使用 DS 软件构建 SHP2 催化位点抑制剂和非抑制剂的贝叶斯分类模型,抑制剂的 3D-QSAR 药效团模型,利用贝叶斯分类模型对大型商业数据库进行筛选,预测阳性化合物再由药效团进行筛选,通过两轮筛选的化合物利用分子对接技术对 SHP2 催化活性位点,通过分析对接结果选出潜在活性化合物,再从公司购买化合物实体,进行酶水平抑制活性检测,最终得到 3 个具有一定潜在选择性的 SHP2 新型抑制

剂。并通过分析活性最高化合物和药效团匹配和分子对接结果,阐明化合物起抑制作用的机制。

2 材料和方法

2.1 搜集 SHP2 催化位点小分子抑制剂

从 BindingDB 数据库下载已报导的 SHP2 小分子抑制剂的 SDF 文件^[26],利用 Discovery Studio 3.1(DS 3.1)打开文件,根据文件中提供的每个化合物的信息,删除非活性位点抑制剂。利用 DS 3.1 的“Prepare Ligands”模块对剩余 SHP2 催化位点小分子抑制剂进行去重处理;将去重处理后的所有 SHP2 小分子抑制剂按照活性 IC_{50} 进行排序,规定 $IC_{50} < 20 \mu\text{mol/L}$ 为 SHP2 的抑制剂,添加标注为 1, $IC_{50} > 20 \mu\text{mol/L}$ 为非抑制剂,添加标注为 -1。

2.2 构建 SHP2 催化位点小分子抑制和非抑制剂二分类贝叶斯模型

利用 DS 3.1 的“Generate Training and Test Data”模块,构建训练集和测试集,将注有标记的 SHP2 催化活性位点抑制剂作为输入化合物,“Split Method”参数设置为“Random”,“Training Set Percentage”参数设置为 80。利用 DS 3.1 中“Create Bayesian Model”模块,以生成的训练集和测试集化合物作为输入,构建贝叶斯模型,设置“Calculate Properties”时,先确定所用分子指纹的类型,然后将确定的分子指纹和常见的几种分子描述符组合进行模型构建。为确定关键分子描述符,采取逐一排查法,最终确定分子指纹和分子描述符的最优组合。

2.3 构建 SHP2 催化位点小分子抑制剂 3D-QSAR 药效团

2.3.1 训练集化合物的选取。药效团建模过程中,训练集分子的挑选对模型的构建具有非常重要的作用。利用 DS3.1 构建 3D-QSAR 药效团模型对训练集分子有如下要求^[27]:(1)18-25 个结构多样性的化合物,为保证结果具有统计学意义,化合物分子数目不得低于 15 个。(2)化合物活性跨度至少为 4 个数量级,每个数量级至少有 3 个化合物。(3)结构相似的化合物活性至少相差一个数量级,有相似活性的化合物结构必须不同。(4)所有化合物提供的信息不得冗余。根据这 4 条要求,从搜集的 SHP2 催化位点抑制剂中选择 16 个化合物构成训练集。

2.3.2 3D-QSAR 药效团的构建。利用 DS 3.1 中的“3D QSAR Pharmacophore Generation”模块以 16 个训练集化合物作为输入化合物,两个高活性化合物 1($IC_{50}:0.2 \mu\text{mol/L}$)和 2($IC_{50}:0.45 \mu\text{mol/L}$)用作建模过程中的模板分子,其对应的参数“Principal”及“MaxOmitFeat”分别设为 2 和 0,其余化合物的这两个参数都设为 1。

所有 SHP2 催化活性位点抑制剂利用 DS 3.1 的“Prepare Ligands”模块进行能量优化,使其初始构象处于能量较低的状态。在“3D QSAR Pharmacophore Generation”模块中“Conformation Generation”参数设为 Best,并将能量阈值设为“20 kcal/mol”,然后利用 Poling 算法对每个化合物产生 255 个构象,这些构象涵盖了所有可能与 SHP2 催化结构域结合的化合物的活性构象。“Maximum Pharmacophores”参数设置为 5,“Minimum Interfeature Distance”设为 1.5,利用剩余 SHP2 抑制剂作为测试集化合物,“Fitting Method”设置为 flexible,“Fisher Validation”设置为 95%。除此,构建药效团所需其它参数采用默认参数。

通过对所有 SHP2 小分子抑制剂的观察和相关文献报道,选择以下药效特征元素作为初始药效特征:氢键受体(Hydrogen-bond acceptor, A),氢键供体(Hydrogen-bond donor, D),疏水特征(Hydrophobic feature, H),芳香环特征(Ring aromatic feature, R)和电离后呈负电荷基团(NEG_ionizable, N)。

2.3.3 3D-QSAR 药效团的评价。通常通过分析计算结果中给出的 cost 值和其它相关统计学参数对 3D-QSAR 药效团模型进行初评,对通过初评的药效团模型再利用测试集和交叉验证法进行进一步的评价。参数修改后运算给出最好的 5 个药效团模型及其相应参数,这些参数主要有:Fixed cost,理想模型的 cost 值;Null cost,无意义模型的 cost 值,即模型预测活性和实验活性完全不相关时的 cost 值;Total cost,对应于每个不同的药效团模型其值不同;相关系数(correlation coefficient),代表利用药效团模型预测的训练集分子活性值和实际实验活性值的相关度;Config 值,代表药效团模型空间构象的复杂程度;还有均方根偏差值(RMSD value)等。一个好的药效团模型,应当具有较高的相关系数,其 Total cost 值应该接近 Fixed cost 值

远离 Null cost 值, Config 值不大于 17, 均方根偏差越小越好。交叉验证法可以进一步评价通过初评药效团模型的可信度(confidence level), 将置信度设为 95%, 那么程序将产生 19 个随机表单, 每个表单中化合物的结构和活性数据是随机混乱匹配的, 但是表单中的其它参数与构建初始药效团模型的参数值完全相同。通过分析随机产生的 19 个训练集构建的 95 个药效团模型的统计学参数可评价构建的定量药效团模型的可信性。在测试集验证方法中, 将除训练集之外的 66 个化合物组成独立测试集, 运用构建的定量药效团模型对测试集化合物进行活性预测并与其实验活性做回归分析, 以验证药效团模型的预测能力。

2.4 分子对接研究

本研究中的所有对接研究均使用 GOLD 5.1 来完成^[28]。SHP2 催化结构域的晶体结构来源于蛋白质数据库(PDB ID: 4RDD), 因为其在所有的 SHP2 催化结构域晶体结构中分辨率最高(0.16 nm), 所以被作为对接研究中小分子化合物的受体。通过使用 DS 3.1 将氢原子添加到蛋白质, 然后分配 ChARMm 力场。结合位点被定义为一个球体, 其覆盖以配体为中心 1 nm 范围内的受体氨基酸残基, 这个球体足够覆盖活性位点中配体化合物的结合区域。通过将 SHP2 催化结构域共结晶的抑制剂对接到受体的活性部位, 预先优化了评分函数和对接参数。

2.5 新型 SHP2 催化位点抑制剂的虚拟筛选

使用构建好的贝叶斯分类模型通过 DS 3.1 中“Calculate Molecular Properties”模块对大型商业化合物数据库里的化合物进行预测, 利用预测为阳性的化合物, 构建药效团筛选数据库。使用 DS 3.1 中的“Search 3D Database”模块, 利用建立的药效团模型对构建好的化合物数据库进行 3D 结构匹配, 选择匹配值(Fit Value)超过一定阈值的化合物进行分子对接研究, 通过对对接结果的分析, 选择购买数个化合物进行后续的研究。

2.6 酶水平化合物活性检测

本研究中酶水平活性检测选用欧洲 Eurofins Discovery 公司的 Discovery Services 模块中的 PhosphataseProfiler™ 完成, 该方法使用荧光法测试磷酸酶活性。参考文献^[29,30]并作适当调整, 以 6,8-二氟-4-甲基-7-羟基香豆素磷酸酯(DiFMUP)作为底物, SHP2 催化其水解产生 6,8-二氟-4-甲基-7-羟基香豆素(DiFMU), 通过 Thermo Scientific Verioskan Flash 多功能读数仪, 以 358 nm 为激发波长来检测 450 nm 处的荧光值, 确定 SHP2 的酶活。以钒酸钠($\text{Na}_3\text{VO}_4 \cdot 12\text{H}_2\text{O}$)为阳性抑制剂, 它能够氧化蛋白酪氨酸磷酸酶的保守催化活性位点半胱氨酸, 是一种 PTPs 的通用抑制剂。除了感兴趣的待测化合物被正钒酸钠取代, 阳性对照孔包含了反应的所有组分以及控制溶剂效应的 DMSO; 除了化合物, 空白孔包含了反应的所有成分。各种不同酶的活性抑制实验具体操作如下。

2.6.1 SHP2 活性检测。使用在大肠杆菌中表达的人重组磷酸酶 PTPN11(SHP-2)。将试验化合物与 45 ng/mL 酶在 37 °C、pH 7.2 的改良 HEPES 缓冲液中预孵育 15 min。通过添加 100 $\mu\text{mol/L}$ DiFMUP 60 min 来引发反应。在 358 nm/450 nm 下用荧光光谱法读取所形成的 DiFMU 的量的测定。化合物在 10 $\mu\text{mol/L}$ 下进行筛选。

2.6.2 SHP1 活性检测。使用在大肠杆菌中表达的人重组蛋白磷酸酶 SHP1。将试验化合物与 2 U/mL 酶在 37 °C、pH 7.2 的改良 HEPES 缓冲液中预孵育 15 min。通过加入 100 $\mu\text{mol/L}$ DiFMUP 30 min 来引发反应。在 358 nm/450 nm 下用荧光光谱法读取所形成的 DiFMU 的量的测定。化合物在 10 $\mu\text{mol/L}$ 下进行筛选。

2.6.3 PTP1B 活性检测。使用在大肠杆菌中表达的人重组蛋白磷酸酶 PTP1B。将受试化合物与 170 $\mu\text{U/mL}$ 酶在 37 °C、pH 7.2 的改良 HEPES 缓冲液中预孵育 15 min。通过加入 10 μM DiFMUP 60 min 来引发反应。在 358 nm/450 nm 下用荧光光谱法读取所形成的 DiFMU 的量的测定。化合物在 10 $\mu\text{mol/L}$ 下进行筛选。

2.6.4 DUSP22 检测。使用在大肠杆菌细胞中表达的人重组蛋白磷酸酶 DUSP22。将受试化合物与 2 $\mu\text{g/mL}$ 酶在 37 °C、pH 7.2 的改良 HEPES 缓冲液中预孵育 15 min。通过加入 10 $\mu\text{mol/L}$ DiFMUP 30 min 来引发反应。在 358 nm/450 nm 下用荧光光谱法读取所形成的 DiFMU 的量的测定。化合物在 10 $\mu\text{mol/L}$ 下进行筛选。

3 结果和讨论

3.1 SHP2 催化位点小分子抑制和非抑制剂二分类贝叶斯模型

总共搜集到 376 个 SHP2 催化位点抑制剂,其中 220 个化合物活性(IC_{50})高于 $20 \mu\text{mol/L}$,在模型构建中作为非活性化合物,标注为-1;剩余 157 个化合物活性(IC_{50})低于 $20 \mu\text{mol/L}$,在构建模型时作为活性抑制剂,标注为 1。训练集化合物占比 80%,剩余 20%为测试集化合物,经程序随机分配后,训练集总共 301 个化合物,其中 178 个非活性化合物,123 个活性化合物;测试集总共 75 个化合物,由 42 个非活性化合物和 33 个活性化合物组成。

模型构建时,分子指纹 LCFC16 与分子描述符 Num_H_Donors, Molecular_Fractional Polar Surface Area 为最优组合,给出最好的二分类结果。最优贝叶斯分类模型 SHP2_BDM 评价数据见表 1,ROC Score, SE, SP 和 Concordance 数值越接近 1,说明模型的二分类能力越优。从表 1 中的数据可见已构建模型是一个优秀的贝叶斯分类模型,可用于后续虚拟筛选。

表 1 SHP2 催化位点抑制剂和非抑制剂贝叶斯分类模型的验证结果

项目	Model Name	ROC Score	ROC Rating	TP	FN	FP	TN	SE	SP	Concordance
训练集 5 重交叉验证结果	SHP2_BDM	0.893	Good	114	9	18	160	0.927	0.899	0.910
测试集验证结果	SHP2_BDM	0.928	Excellent	30	3	7	35	0.909	0.833	0.867

注:TP, true positive; TN, true negative; FP, false positive; FN, false negative; SE: sensitivity, $SE = TP/(TP + FN)$; SP: specificity, $SP = TN/(TN + FP)$; Concordance: overall accuracy, $Concordance = (TP + TN)/(TP + TN + FP + FN)$ 。

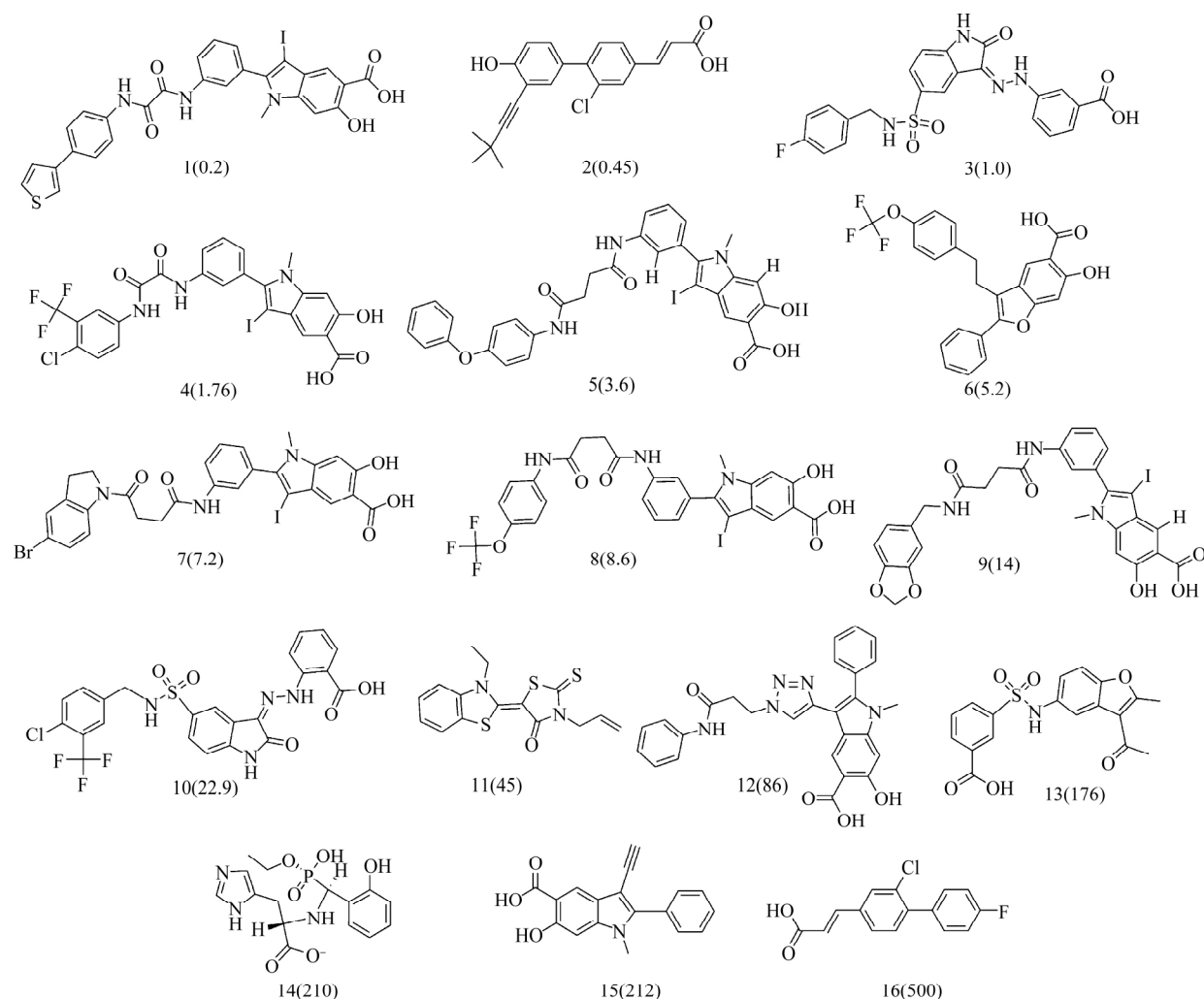


图 1 构建药效团模型的训练集分子化学结构及其生物活性(IC_{50} : $\mu\text{mol/L}$)

3.2 3D-QSAR 药效团模型的构建及评价

3.2.1 药效团模型的构建。利用选出的 16 个训练集小分子化合物(图 1),经过运算得到 5 个最佳药效团模型,其参数见表 2,从表 2 数据可见,P1-P5 的 Cost 差值分布在 53.014~63.656 之间。就药效团有意义的显著性而言,真正重要的是任何药效团的 Total cost 与 Null cost 之间的差异大小,如果二者差异为 40~60,那么代表有 75%~90%的可能性活性数据和化学结构是真正相关的。Total cost 和 Null cost 差值越大代表构建的药效团预测化合物活性的能力越高。

表 2 排名前五的 SHP2 催化位点小分子抑制剂定量药效团模型的统计学参数

Pharmacophore No.	Total cost	Cost diff. ^a	RMSD(Å)	Features ^b
P1	69.784 90	63.456 0	0.842 20	A ₁ H ₁ N ₁ R ₁ V ₄
P2	79.358 10	53.883 0	1.396 00	A ₂ H ₂ N ₁ V ₃
P3	79.479 10	53.762 0	1.352 41	A ₁ H ₂ N ₁ R ₁ V ₀
P4	79.760 90	53.480 0	1.376 70	A ₁ H ₂ N ₁ V ₀
P5	80.227 50	53.014 0	1.417 52	A ₁ D ₁ H ₁ N ₁ V ₄

注:a: (Null cost-total cost), Null cost = 133.241, Fixed cost = 63.527 6, configuration cost = 15.953 9; b: A, D, H, N 和 V 分别代表氢键受体、氢键供体、疏水特征、负电荷特征和排除体积(排斥球)。

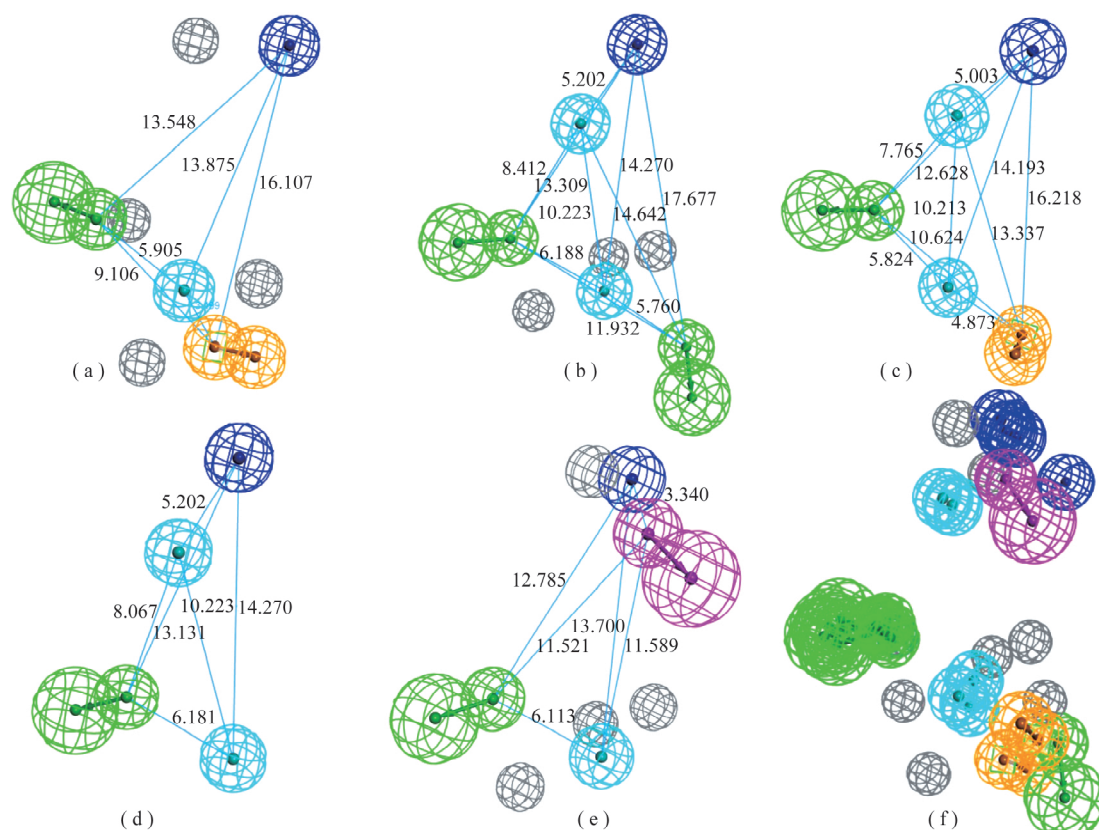


图 2 DS 产生的 SHP2 催化位点抑制剂的定量药效团模型(a~e)药效团 P1-P5 药效特征元素的空间排布及其几何参数;
(f) 5 个药效团的叠合结果(绿色:氢键受体;玫红色:氢键供体;蓝色:疏水特征;黄色:芳香环;灰色:排斥球)

图 2(a~e)显示了构建的 5 个定量药效团模型,包括药效特征元素种类和药效团特征元素三维空间以及距离限制,图 2(f)显示了 5 个药效团的叠合结果,从图 2 可以看出 5 个药效团的药效特征的空间排布和各个特征元素之间的距离是相差无几的。但 5 个药效团的药效特征的种类和数目却有较大差别。5 个药效团都含有空间分布类似的 1 个电离负电荷基团特征,1 个氢键受体和 1 个疏水特征。P2、P3 和 P4 还各含有一个位置相似的疏水基团特征;P1 和 P3 还各含有一个位置类似的芳香环特征;P2 在与 P1 和 P3 芳香环特征位置相似的空间位置有一个氢键受体;P5 还在一个靠近电离负电荷特征的位置有个氢键供体;除此,P1、P2

和 P5 还含有 3~4 个不等的排斥球,即化合物不能占据的空间。经过分析 5 个药效团含有的药效特征种类、数目和空间排布,本研究倾向选择 P3 和 P5 进行后续的活性化合物筛选。

3.2.2 药效团模型的验证。好的药效团模型不仅能正确预测训练集化合物的分子活性,还应当具有正确预测训练集之外其它分子生物活性的能力。因此,本研究使用 66 个训练集分子之外的 SHP2 催化位点抑制剂组成测试集,用于测试药效团正确预测外部分子活性的能力。图 3 显示了 5 个药效团模型对 16 个训练集化合物和 66 个测试集化合物预测活性值与实验活性值的回归分析结果。由图可见,5 个药效团模型对训练集化合物的活性预测是非常准确的,化合物预测活性值和实验活性值间的相关系数在 0.923 3~0.952 4。通过对 5 个药效团模型对测试集化合物预测活性值和实验活性值进行回归分析,得出药效团模型准确预测训练集之外化合物活性值的能力,从图 3 可见,5 个药效团模型对 66 个测试集化合物的活性预测值和实验活性值之间的相关系数在 0.822 8~0.909 1,其中 P3 和 P5 的预测准确度最高,相关系数分别为 0.899 8 和 0.909 1。通过对化合物活性预测值和实验值相关系数的分析,本研究倾向选择 P3 和 P5 进行后续的虚拟筛选,以发现新型 SHP2 催化位点抑制剂。

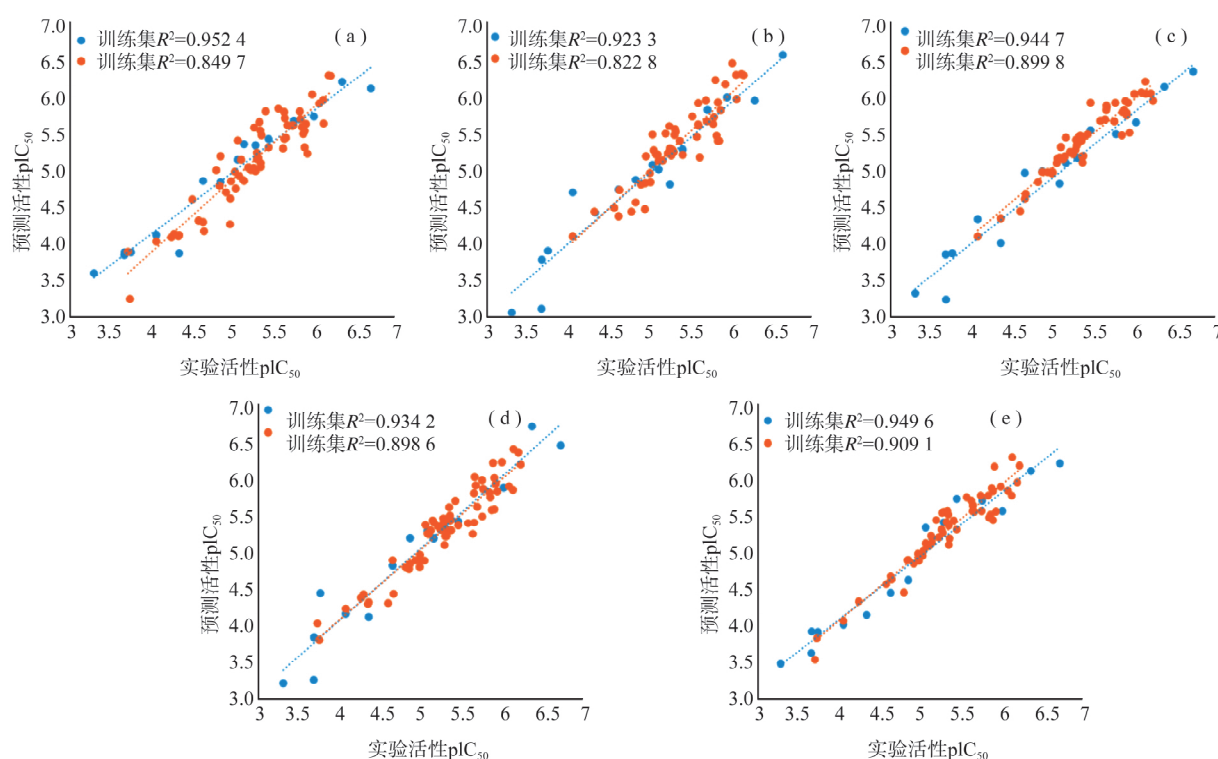


图 3 5 个药效团模型对训练集和测试集化合物活性预测值和实验值回归分析结果

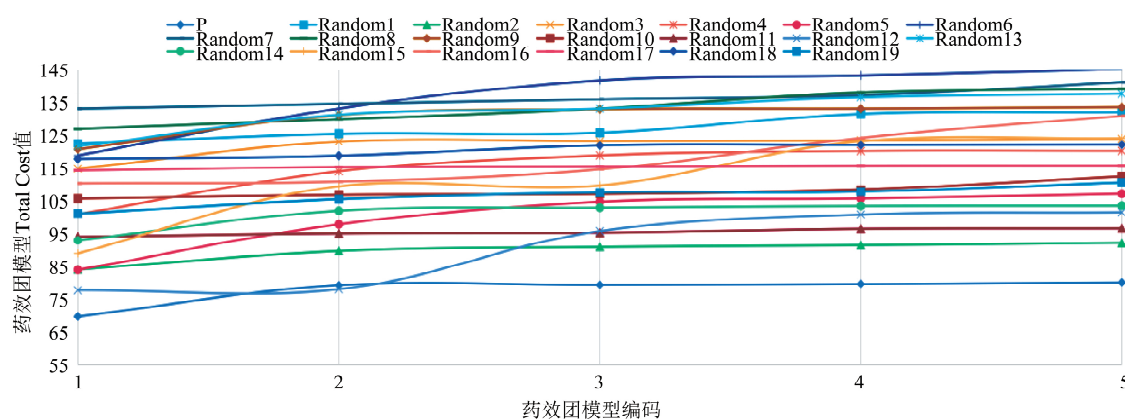


图 4 原始表单和 19 个随机表单产生的药效团的 total cost 值的区别

除了测试集方法,本研究采用 Fisher 验证法对药效团进行随机交叉验证分析。在本次验证中,置信水

平设为 95%, 程序将训练集化合物结构和活性值随机混乱组合, 产生 19 个输入表单, 这些表单参数与构建药效团模型时使用的参数设置完全相同, 程序将产生 95 个药效团型。图 4 显示了新建的 95 个药效团模型和原始表单产生的 5 个药效团模型的 Total cost 值, 比较发现原始表单产生的 5 个药效团模型除药效团 P2 外的 Total cost 值比随机表单产生的 95 个药效团的 Total cost 值低的多, 这说明构建的药效团模型 P1、P3、P4 和 P5 具有很强的可信性。

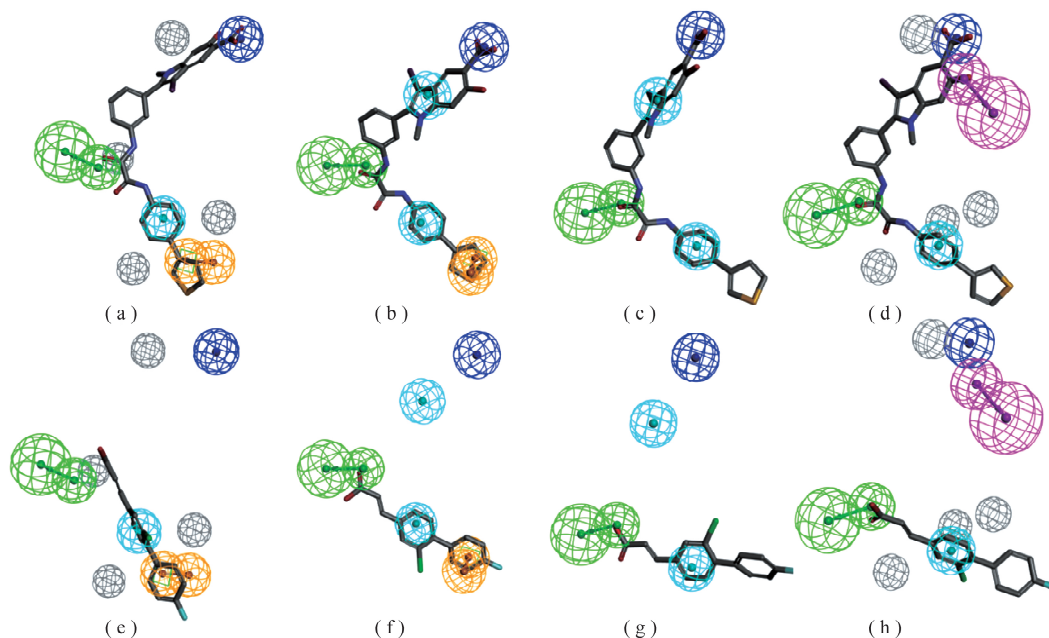


图 5 药效团和训练集活性最高的化合物 1 (IC_{50} : $0.2 \mu\text{mol/L}$) 和活性最低化合物 16 (IC_{50} : $500 \mu\text{mol/L}$) 的匹配效果图
(a~d) P1、P3、P4 和 P5 与活性最高的化合物 1 的匹配图; (e~h) P1、P3、P4 和 P5 与活性最低的化合物 16 的匹配图
(绿色: 氢键受体; 玫红色: 氢键供体; 蓝色: 疏水特征; 黄色: 芳香环特征; 灰色: 排斥球)

为了进一步分析药效团模型 P1、P3、P4 和 P5 的有效性, 将这 4 个药效团模型与训练集活性最高和活性最低的 2 个化合物进行匹配叠合。图 5 显示了定量药效团模型 P1、P3、P4 和 P5 与训练集中活性最高的化合物 1 ($IC_{50} = 0.2 \mu\text{mol/L}$) 和活性最低化合物 16 ($IC_{50} = 500 \mu\text{mol/L}$) 的匹配情况, 从图 5(a-d) 可见, 各药效团和化合物 1 匹配得非常完美, 其每个药效特征元素均与化合物 1 中的相应化学功能基团匹配, 而在图 5(e-h) 中药效团模型只与化合物 16 有部分匹配。因此, 这些结果在一定程度上反映了各药效团的有效性。

根据药效团模型对训练集和测试集化合物的活性预测能力, Fisher 验证法分析, 药效特征元素的组成和空间排布, 及药效团和训练集化合物的匹配情况, 最终选择药效团 P3 和 P5 进行后续的虚拟筛选。

3.3 分子对接参数和评分函数的确定

由于分子对接参数和评分函数对最终对接结果具有很大影响, 因此在进行基于分子对接的虚拟筛选之前需要确定最优对接参数和评分函数。本研究选择 SHP2 催化结构域结合 1 个头孢磺啶衍生物的晶体复合物结构 (PDB ID: 4RDD) 作为对接研究的受体, 因为在所有已结晶 SHP2 催化结构域与催化位点抑制剂的复合物结构中, 4RDD 具有高分辨率 (0.16 nm)。为了确定最佳的对接参数和评分函数, 将已经与 SHP2 催化位点共结晶的 4 种活性化合物重新对接到 SHP2 的活性位点, 调整对接

表 3 SHP2 催化位点抑制剂的结晶构象和对接构象之间 RMSD 值

Ligand ID ^a	PDB entry	RMSD/nm
3LU	4RDD	0.186 95
JZG	3O5X	0.630 73
B2B	3MOW	0.735 35
2WT	4PVG	0.803 07

注: a: PDB 数据库中配体化合物的 ID 编号。

参数和评分函数, 直到小分子化合物对接构象尽可能与原来的结晶构象重合。经过不断测试, 对接参数设置如: Fitness Function 选用 Chemscore, GA Parameters 选用 GOLD Default, Generate Diverse Solution 设为 True, Early Termination 设为 False, 其它参数保留软件默认参数。选用参数和评分函数的组合, 使配体的

对接构象与它们在晶体结构中的相应结合构象之间有最小均方根偏差(RMSD)值。表3列出了计算的RMSD值。4个化合物中3LU具有最小RMSD值0.186 95 nm,此值可接受。其它3个化合物的RMSD值都高于0.5 nm,这是因为这3个化合物本身柔性较大,与SHP2催化位点结合不够紧密,化合物三分之二的结构暴露于溶剂中与受体蛋白无任何相互作用,故对接构象与原结晶构象差别较大。基于化合物3LU的对接结果及对接软件GOLD5.1在其它研究中的可靠性^[29],表明GOLD软件在一定程度上能够重现配体的正确结合构象,可用于本研究中的分子对接研究。

3.4 新型 SHP2 催化位点抑制剂的虚拟筛选

本研究采用3种计算机辅助药物发现技术,即贝叶斯分类模型,药效团模型和分子对接技术,进行多步骤的组合虚拟筛选以发现潜在新型SHP2活性位点抑制剂。虚拟筛选工作流程如图6所示。本研究所用3种计算机辅助药物发现技术中,基于贝叶斯分类模型的虚拟筛选速度最快,所以首先利用贝叶斯分类模型对包含129 087个化合物的大型商业化合物数据库进行SHP2催化位点抑制剂的预测;把预测为阳性的48 507个化合物构建为可用药效团模型进行筛选的3D数据库,利用比分子对接速度快很多的药效团模型进行3D数据库筛选;将筛选化合物与药效团匹配值高于4(FitValue大于4.0)的158个化合物进行最后的分子对接筛选,根据对接化合物的打分值高低和观察对接构象与活性位点关键氨基酸残基的相互作用,最终选取了8个化合物进行购买和后续的酶水平活性验证。图7显示8个购买化合物的分子结构,表4列出8个化合物的筛选参数。

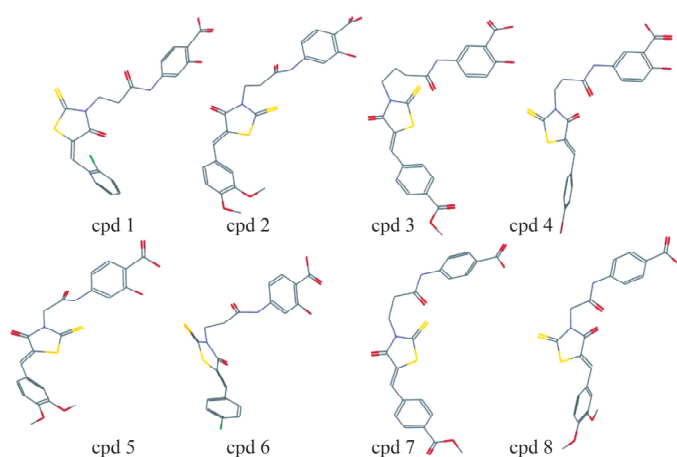


图7 8个命中化合物的化学结构

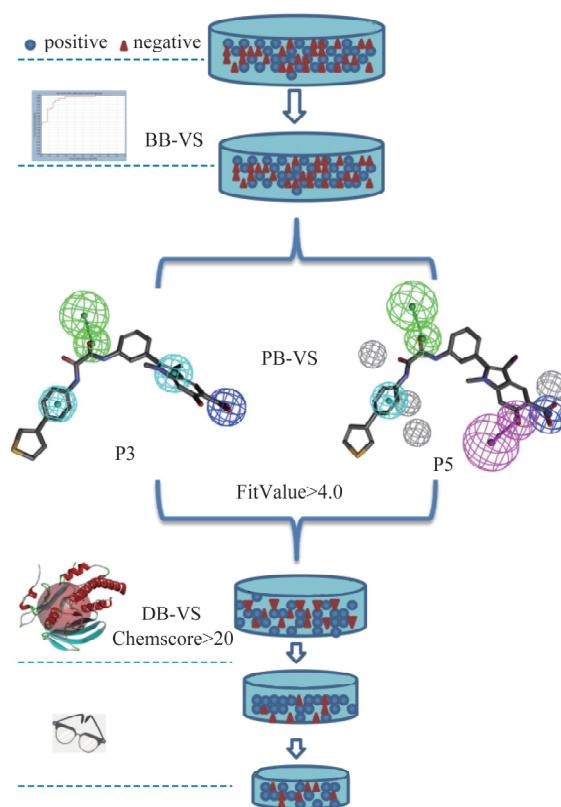


图6 本研究的虚拟筛选流程

表4 8个命中化合物的筛选参数,包括药效团匹配值、3D-QSAR药效团模型预测的活性值和对接打分

Compound Name	P3		P5		Docking score
	Fit Value	Pre.Ac.	Fit Value	Pre.Ac.	
cpd1	5.620 96	2.053 24	6.348 26	0.113 643	28.772 4
cpd2	5.572 29	2.296 76	6.175 98	0.168 973	25.525 3
cpd3	5.533 63	2.510 53	6.122 95	0.190 921	27.672 4
cpd4	5.456 69	2.997 19	5.497 72	0.805 527	22.762 8
cpd5	5.375 81	3.610 70	5.206 38	1.575 510	29.294 1
cpd6	5.354 68	3.790 71	4.955 71	2.806 000	23.100 8
cpd7	5.342 57	3.897 86	4.682 09	5.268 800	30.138 8
cpd8	4.526 13	25.542 80	4.105 80	19.860 800	23.978 6

3.5 新型 SHP2 催化位点抑制剂活性检测

化合物由 LifeChemical 化合物数据库国内代理商上海陶术购买(根据化合物售卖公司提供的材料显示,所有化合物纯度均大于等于 90%),每种化合物皆用 DMSO 溶解,精确配置成 10 mmol/L 的母液,每个待测化合物取 20 μ L 母液,寄往英国 Eurofins Discovery 公司完成酶水平活性检测。图 8 显示化合物在 10 μ mol/L 浓度下对 SHP2 活性的抑制率,化合物 cpd2、cpd4 和 cpd6 在浓度为 10 μ mol/L 时对 SHP2 催化活性具有明显抑制能力,抑制率分别为 79%,71%和 67%。为了进一步检测活性最高化合物抑制不同磷酸酶催化活性的选择性,检测了 cpd2 在 10 μ mol/L 浓度时,对 SHP1、PTP1B 和 DUSP22 催化活性的抑制。如图 9 所示,化合物 cpd2 在 10 μ mol/L 浓度时,对 SHP1、PTP1B 和 DUSP22 的酶活抑制率分别为 42%,41%和 34%,与同浓度 cpd2 对 SHP2 酶活抑制率 21%相比较,提示化合物 cpd2 具有一定潜在选择性。酶活检测表明,化合物 cpd2 是一个具有潜在选择性的 SHP2 催化位点抑制剂,具有进一步改构的价值。

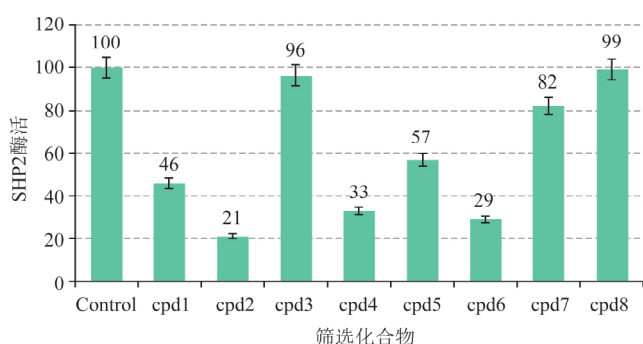


图 8 8 个筛选命中化合物抑制 SHP2 酶活性结果

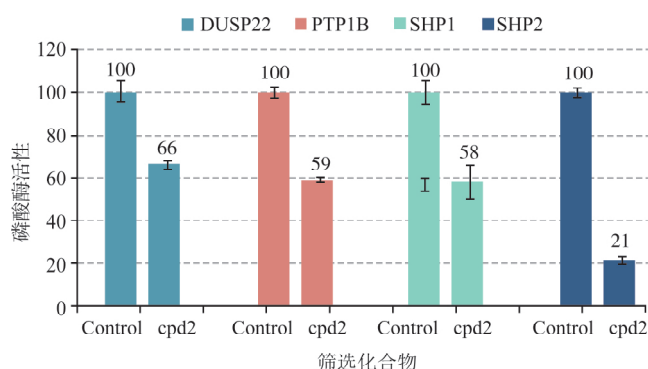


图 9 化合物 cpd2 酶水平抑制 SHP2、SHP1、PTP1B 和 DUSP22 活性比较

3.6 活性化合物 cpd2 与药效团模型的匹配及分子对接结果分析

为了更好地解释化合物 cpd2 结构和活性间的关系,分析了该化合物和药效团的匹配情况,及其与 SHP2 催化活性位点的对接结果,以阐明化合物 cpd2 为何会抑制 SHP2 催化活性。

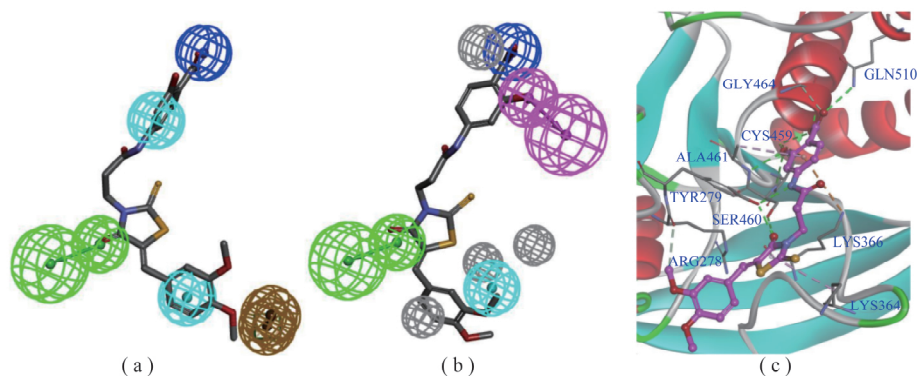


图 10 (a)和(b)分别是药效团 P3 和 P5 与化合物 cpd2 的匹配结果(绿色:氢键受体,玫红色:氢键供体,蓝色:疏水特征,黄色:芳香环特征,灰色:排斥球);(c)cpd2 在 SHP2 催化活性位点中可能的结合构象(绿色:氢键,浅粉:疏水相互作用,橙色:静电相互作用,灰绿:弱氢键)

从图 10(a)中可以看到,命中化合物 cpd2 与药效团模型 P3 匹配效果较好。具体地说,2 个疏水特征与 cpd2 的苯基匹配,解离负电荷特征与羧基匹配,氢键受体与硫氧代噻唑啉上的氧原子匹配,芳香环特征没有匹配到化合物 cpd2 的结构上。相比较,化合物 cpd2 与药效团模型 P5 匹配效果好于其与 P3 的匹配效果,如图 10(b)。具体来说,解离负电荷特征与羧基匹配,氢键受体与硫氧代噻唑啉上的氧原子匹配,氢键受体与羟基匹配,疏水特征与甲氧基匹配。图 10(c)显示了 cpd2 在 SHP2 催化活性位点中可能的结合构象,水杨

酸结合在 SHP2 底物磷酸化部位的结合位点,其中羧基与 GLN510 和 GLY464 形成氢键相互作用,羟基和 CYS459 形成氢键相互作用,苯环与 ALA461 形成疏水相互作用并与 LYS366 形成正电荷- π 相互作用;硫氧代噻唑啉上的氧原子与 SER460 形成氢键相互作用,硫氧代噻唑啉的五元环与 LYS364 形成疏水相互作用;邻二甲氧基苯上的一个甲氧基与 ARG278 形成弱氢键相互作用。化合物 cpd2 与药效团的匹配结果与其和 SHP2 催化活性位点结合形成的相互作用相吻合,说明药效团预测的 cpd2 活性构象和分子对接预测的活性构象基本吻合,化合物 cpd2 极可能是采取此种构象结合在 SHP2 催化位点抑制其催化活性。药效团匹配和分子对接的结果阐明了化合物 cpd2 抑制 SHP2 催化活性的作用机制。

4 结论

本研究利用计算机辅助药物设计技术结合酶水平活性测试,发现了 3 个新型 SHP2 催化位点抑制剂,其中活性最高化合物 cpd2 具有一定的潜在选择性。通过分析 cpd2 与 SHP2 催化位点抑制剂 3D-QSAR 药效团的匹配结果以及 cpd2 与 SHP2 催化活性位点的对接结果,发现药效团预测的活性构象与分子对接预测的活性构象相当吻合,且 cpd2 与药效团关键药效特征的匹配与其和 SHP2 催化活性位点关键氨基酸残基的相互作用基本一致,所以阐明了 cpd2 抑制 SHP2 酪氨酸磷酸酶活性的作用机制。

参 考 文 献

- [1] BARFORD D, NEEL B G. Revealing mechanisms for SH2 domain mediated regulation of the protein tyrosine phosphatase SHP-2[J]. *Structure*, 1998, 6(3): 249-254.
- [2] FRANKSON R, YU Z H, BAI Y P, et al. Therapeutic targeting of oncogenic tyrosine phosphatases[J]. *Cancer Research*, 2017, 77(21): 5701-5705.
- [3] RAN H, TSUTSUMI R, ARAKI T, et al. Sticking it to cancer with molecular glue for SHP2[J]. *Cancer Cell*, 2016, 30(2): 194-196.
- [4] BENTIREZ-ALJ M, PAEZ J G, DAVID F S, et al. Activating mutations of the Noonan syndrome associated SHP2/PTPN1 gene in human solid tumors and adult acute myelogenous leukemia[J]. *Cancer Research*, 2004, 64(24): 8816.
- [5] TARTAGLIA M, MEHLER EL, GOLDBERG R, et al. Mutations in PTPN11, encoding the protein tyrosine phosphatase SHP-2, cause Noonan syndrome[J]. *Nature Genetics*, 2001, 29(4): 465-468.
- [6] TARTAGLIA M, NIEMEYER CM, FRAGALE A, et al. Somatic mutations in PTPN11 in juvenile myelomonocytic leukemia, myelodysplastic syndromes and acute myeloid leukemia[J]. *Nature Genetics*, 2003, 34(4): 464-464.
- [7] MULERO-NAVARRO S, SEVILLA A, ROMAN AC, et al. Myeloid dysregulation in a human induced pluripotent stem cell model of PTPN11-associated juvenile myelomonocytic leukemia[J]. *Cell Reports*, 2015, 13(3): 504-515.
- [8] DONG L, YU W M, ZHENG H, et al. Leukaemogenic effects of Ptpn11 activating mutations in the stem cell microenvironment[J]. *Nature*, 2016, 539: 304.
- [9] ACETO N, SAUSGRUBER N, BRINKHAUS H, et al. Tyrosine phosphatase SHP2 promotes breast cancer progression and maintains tumor-initiating cells via activation of key transcription factors and a positive feedback signaling loop[J]. *Nature Medicine*, 2012, 18(4): 529-537.
- [10] RUESS D A, HEYNEN G J, CIECIELSKI K J, et al. Mutant KRAS-driven cancers depend on PTPN11/SHP2 phosphatase[J]. *Nature Medicine*, 2018, 24(7): 954-960.
- [11] MAINARDI S, MULERO-SANCHEZ A, PRAHALLAD A, et al. SHP2 is required for growth of KRAS-mutant non-small-cell lung cancer in vivo[J]. *Nature Medicine*, 2018, 24(7): 961-967.
- [12] ZHAO M, GUO W, WU Y, et al. SHP2 inhibition triggers anti-tumor immunity and synergizes with PD-1 blockade[J]. *Acta Pharmaceutica Sinica B*, 2019, 9(2): 304-315.
- [13] 孔娇, 龙亚秋. 蛋白酪氨酸磷酸酶 SHP2 及其抑制剂的研究进展[J]. *药学进展*, 2019, 43(7): 517-526.
- [14] 张惠伦, 肖鹏, 张雪, 等. 基于酪氨酸磷酸酶 SHP2 变构抑制剂的肿瘤靶向治疗[J]. *中国肿瘤生物治疗杂志*, 2019, 26(2): 230-235.

- [15] HE R, YU Z H, ZHANG R Y, et al. Exploring the existing drug space for novel pTyr mimetic and SHP2 inhibitors[J]. ACS Medicinal Chemistry Letters, 2015, 6(7): 782-786.
- [16] ZHANG X, HE Y T, LIU S J, et al. Salicylic acid based small molecule inhibitor for the oncogenic src homology-2 domain containing protein tyrosine phosphatase-2 (SHP2)[J]. Journal of Medicinal Chemistry, 2010, 53(6): 2482-2493.
- [17] ZENG L F, ZHANG R Y, YU Z H, et al. Therapeutic potential of targeting the oncogenic SHP2 phosphatase[J]. Journal of Medicinal Chemistry, 2014, 57(15): 6594-6609.
- [18] Meng X D, GAO L X, WANG Z Y, et al. Synthesis and biological evaluation of 2,5-diaryl-1,3,4-oxadiazole derivatives as novel Src homology 2 domain-containing protein tyrosine phosphatase 2 (SHP2) inhibitors[J]. Bioorganic Chemistry, 2021, 116: 105384.
- [19] LIU W S, YANG B, WANG R R, et al. Design, synthesis and biological evaluation of pyridine derivatives as selective SHP2 inhibitors[J]. Bioorganic Chemistry, 2020, 100: 103875.
- [20] MA Y, LI W Y, SUN T, et al. Structure-based discovery of a specific SHP2 inhibitor with enhanced blood-brain barrier penetration from PubChem database[J]. Bioorganic Chemistry, 2022, 121: 105648.
- [21] MULLARD A. Phosphatases start shedding their stigma of undruggability[J]. Nature Reviews Drug Discovery, 2018, 17(12): 847-849.
- [22] SONG Z, WANG M J, GE Y, et al. Tyrosine phosphatase SHP2 inhibitors in tumor-targeted therapies[J]. Acta Pharmaceutica Sinica B, 2021, 11(1): 13-29.
- [23] Morgan S, Grootendorst P, Lexchin J, et al. The cost of drug development: a systematic review[J]. Health Policy, 2011, 100: 4-17.
- [24] YANG S Y. Pharmacophore modeling and applications in drug discovery: challenges and recent advances[J]. Drug Discovery Today, 2010, 15: 444-450.
- [25] 李洪林, 沈建华, 罗小民, 等. 虚拟筛选与新药发现[J]. 生命科学, 2005, 17: 125-131.
- [26] GILSON M K, LIY T Q, AITALUK M, et al. BindingDB in 2015: a public database for medicinal chemistry, computational chemistry and systems pharmacology[J]. Nucleic Acids Research, 2015, 44: D1045-D1053.
- [27] REN J X, LI L L, ZOU J, et al. Pharmacophore modeling and virtual screening for the discovery of new transforming growth factor- β type I receptor (ALK5) inhibitors[J]. European Journal of Medicinal Chemistry, 2009, 44(11): 4259-4265.
- [28] VERDONK M L, COLE J C, HARTSHORN M J, et al. Improved protein-ligand docking using GOLD[J]. Proteins, 2003, 52(4): 609-23.
- [29] WELTE S, BARINGHAUS K H, SCHMIDER W G, et al. 6,8-Difluoro-4-methylumbiliferyl phosphate, a fluorogenic substrate for protein tyrosine phosphatases[J]. Anal Biochem, 2005, 338(1): 32-38.
- [30] 邓腾, 杨莉娟, 罗国勇, 等. 蛋白酪氨酸磷酸酶 SHP2 抑制剂体外筛选模型的建立及应用[J]. 应用与环境生物学报, 2015, 21(4): 602-606.

(责任编辑: 刘庆松)