

## 基于药效团模型及虚拟筛选方法发现 EphB4 全新抑制剂

邵鑫<sup>1</sup>, 衣倩颖<sup>2</sup>, 杨春蕾<sup>1\*</sup>(四川大学<sup>1</sup>生命科学学院;<sup>2</sup>化学工程学院, 成都 610065)

**摘要** 采用计算机模拟手段, 通过建立 3D-QSAR 模型、虚拟筛选及分子对接方法发现恶性肿瘤治疗靶标 EphB4 潜在的抑制剂。首先, 通过 Catalyst/HypoGen 算法建立药效团模型。其中最好的模型 Hypo1 具有最高的科雷尔值 (Correl 值): 0.96, 最低的 RMS 值: 0.89, 与固定消耗值 (fixed cost): 89.20 最接近的总消耗值 (total cost): 101.26, 和最高的  $\Delta$  消耗值 ( $\Delta$ cost 值): 89.14。随后, Hypo1 经过测试集验证及 Fischer 随机验证, 并用于筛选化合物数据库。然后利用类药性筛选及分子对接手段进一步减少分子数量。最终, 根据预测活性分析、对接得分值及结合模式分析, 得到 23 个具有全新骨架的化合物作为 EphB4 的潜在抑制剂可用于后续研究。

**关键词** EphB4; 3D-QSAR 模型; 虚拟筛选; 分子对接; Catalyst/HypoGen 算法; 肿瘤靶标

**中图分类号** R914.2 **文献标志码** A **文章编号** 1000-5048(2016)01-0038-10

doi:10.11665/j.issn.1000-5048.20160105

**引用本文** 邵鑫, 衣倩颖, 杨春蕾. 基于药效团模型及虚拟筛选方法发现 EphB4 全新抑制剂[J]. 中国药科大学学报, 2016, 47(1): 38–47.  
**Cite this article as:** SHAO Xin, YI Qianying, YANG Chunlei. Discovery of novel EphB4 inhibitors based on pharmacophore modeling and virtual screening techniques[J]. *J China Pharm Univ*, 2016, 47(1): 38–47.

## Discovery of novel EphB4 inhibitors based on pharmacophore modeling and virtual screening techniques

SHAO Xin<sup>1</sup>, YI Qianying<sup>2</sup>, YANG Chunlei<sup>1\*</sup><sup>1</sup>College of Life Sciences; <sup>2</sup>College of Chemical Engineering, Sichuan University, Chengdu 610065, China

**Abstract** The objective of this paper was to discover new potent inhibitors against EphB4 for cancer therapy via computer-aided drug design strategies including 3D-QSAR modeling, virtual screening and molecular docking. The first step was to generate pharmacophore models based on Catalyst/HypoGen algorithm. The best model, Hypo1, had the highest Correl value (0.96), the lowest RMS value (0.89), the closest total cost (101.26) to fixed cost (89.20), and the best  $\Delta$ cost (89.14). Subsequently, Hypo1 was subjected to test set validation and Fischer's randomization verification and was then used as a 3D query to screen database. In order to further narrow the number of hits, drug-likeness screening and molecular docking techniques were applied. Finally, 23 novel molecules with diverse scaffolds were selected as possible candidates against EphB4 for further studies based on predicted activity analysis, docking scores, and binding mode analysis.

**Key words** EphB4; 3D-QSAR models; virtual screening; molecular docking; Catalyst/HypoGen algorithm; tumor target

This study was supported by the Science and Technology Pillar Program of Sichuan Province (No. 2011SZ0116)

Eph 受体是已知最大的酪氨酸蛋白激酶受体家族中的一个亚族, 在很多肿瘤的表达和转移中发挥作用, EphB4 是其中一个重要的成员<sup>[1]</sup>。研究

表明, EphB4 在人胶质瘤、子宫内膜癌、结肠癌、胃癌、乳腺癌等中过表达<sup>[2]</sup>。因此寻求 EphB4 分子的抑制剂对于肿瘤的发生和恶性进展具有重要的

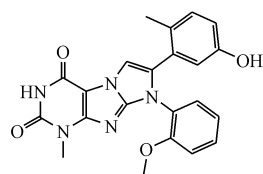
临床意义。

近年来,不断有新的 EphB4 受体蛋白与抑制剂结合的晶体结构照片被解析出来,对了解受体结构和与配体的作用模式起到至关重要的作用,极大促进了基于结构的 EphB4 抑制剂设计。计算机辅助药物设计(CADD)在现代药物设计中起到至关重要的作用,然而采用 CADD 方法设计及筛选 EphB4 抑制剂的研究较少。本研究将根据该蛋白与其现有抑制剂的三维构效关系,构建药效团模型,进而用于筛选数据库并预测抑制活性。在此过程中,采用分子对接手段进一步筛选化合物并研究受体-配体的结合模式,从而获得潜在的 EphB4 抑制剂。

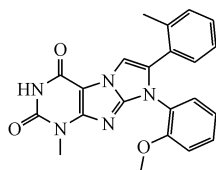
## 1 材料和方法

### 1.1 训练集及测试集分子的准备

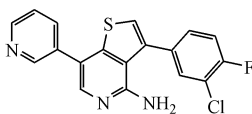
采用 Discovery Studio Client (DS) v2.5.0 软件构建药效团模型及虚拟筛选。首先,考虑到抑制剂骨架多样性及活性至少横跨 4 个数量级 ( $IC_{50}$ : 0.001 ~ 50 nmol/L) 的原则,从文献[3-9]中搜集采用同一方法测得生物活性的共 57 个 EphB4 的抑制剂分子。其中 25 个分子作为训练集用于药效团模型的构建,其 2D 结构如图 1 所示;另外 32 个分子作为测试集用于药效团模型的验证,其 2D 结构略。然后,用 Chem&Bio 3D 12.0 软件画出这些分子的三维结构。



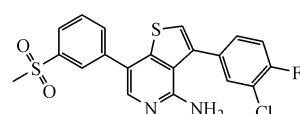
1 ( $IC_{50}$ =0.002  $\mu$ mol/L)



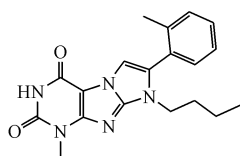
2 ( $IC_{50}$ =0.047  $\mu$ mol/L)



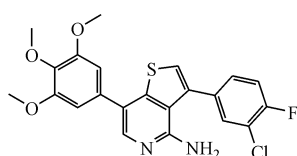
3 ( $IC_{50}$ =0.051  $\mu$ mol/L)



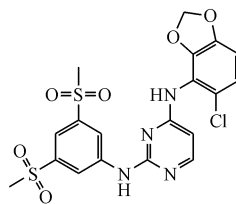
4 ( $IC_{50}$ =0.059  $\mu$ mol/L)



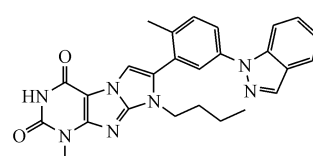
5 ( $IC_{50}$ =0.082  $\mu$ mol/L)



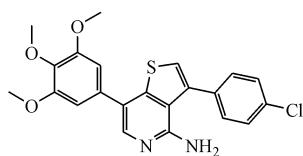
6 ( $IC_{50}$ =0.110  $\mu$ mol/L)



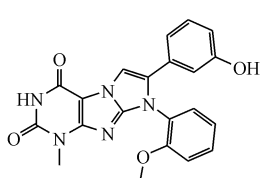
7 ( $IC_{50}$ =0.114  $\mu$ mol/L)



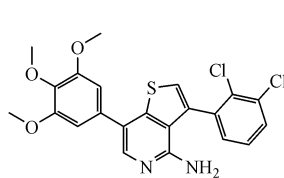
8 ( $IC_{50}$ =0.133  $\mu$ mol/L)



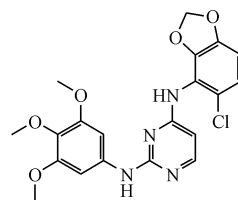
9 ( $IC_{50}$ =0.190  $\mu$ mol/L)



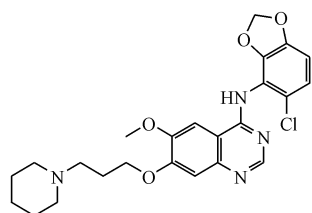
10 ( $IC_{50}$ =0.213  $\mu$ mol/L)



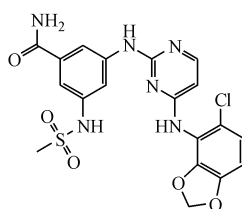
11 ( $IC_{50}$ =0.230  $\mu$ mol/L)



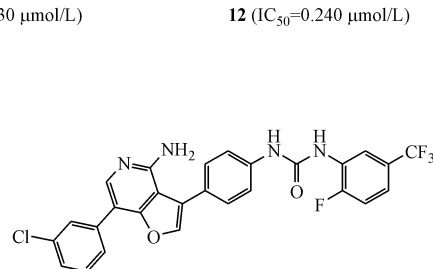
12 ( $IC_{50}$ =0.240  $\mu$ mol/L)



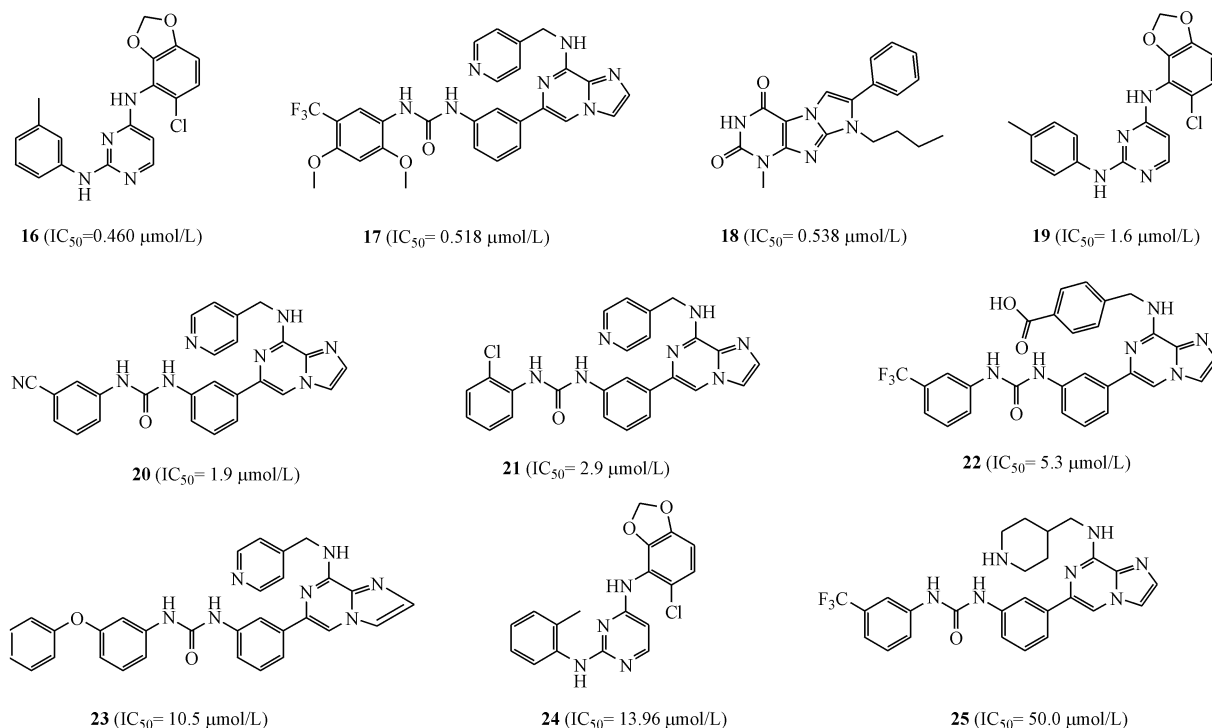
13 ( $IC_{50}$ =0.350  $\mu$ mol/L)



14 ( $IC_{50}$ =0.370  $\mu$ mol/L)



15 ( $IC_{50}$ =0.389  $\mu$ mol/L)



**Figure 1** 2D structures of the 25 training set molecules and their  $IC_{50}$

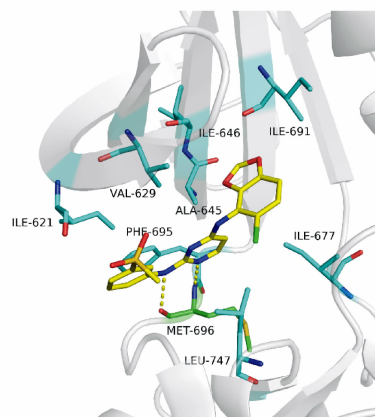
### 1.2 药效团模型的建立

首先,将画好的 52 个抑制剂分子用 CHARMM 力场做能量优化,添加不确定值 (uncertain) 为 2, 并利用特征映射 (Feature Mapping) 模块预测可能的药效学元素。另外结合蛋白与抑制剂复合物进一步确定可能的药效学元素。图 2 为 PDB 编号为 2VWY 的蛋白与抑制剂复合物的晶体结构,从图中可以明显的看到该抑制剂的嘧啶环上的 N 作为氢键受体与 MET696 的氨基上的 H 形成一个氢键;该抑制剂上的亚氨基作为氢键供体与 MET696 的羧基上的 O 形成另一个氢键;该抑制剂包含多个芳环,并暴露于 EphB4 活性空腔中大量的疏水氨基酸(图 2 蓝色棍棒)中。由此可推断 EphB4 抑制剂包含的可能的药效学元素有氢键供体 (HBD)、氢键受体 (HBA)、疏水作用 (HY) 及芳环中心 (RA)。然后,用构象形成 (Generate Conformations) 模块及 BEST 方法产生分子构象。随后,根据 Feature Mapping 结果以及蛋白-抑制剂复合物分析结果, HBD、HBA、HY 及 RA 被选作可能的药效学元素,基于 HypoGen 算法中分子结构-活性关系,用 25 个训练集分子产生共 10 个药效团模型。其中得分最高的 Hypo1 被认为是最好的药效团模型。

### 1.3 药效团模型的验证

利用测试集验证及 Fischer 随机验证方法<sup>[10]</sup>对所建立的药效团模型进行验证。

测试集方法是评估药效团质量好坏的最常用的方法之一。采用配体药效映射 (Ligand Pharmacophore Mapping) 模块及 Flexible/BEST 方法验证 Hypo1 的可信性。随后产生一系列数值包括预测  $IC_{50}$ 、误差 (error) 值、拟合值 (fit value) 等,均可作为 Hypo1 质量的评估标准。



**Figure 2** Binding mode of co-crystallized ligand and EphB4 (PDB code: 2VWY). The protein and ligand are represented in grey cartoon and yellow stick. Cyan and green sticks represent the key amide acid residues

本研究又利用 Fischer 随机验证方法进一步评估所建立药效团的合理性。该方法重新洗牌所有训练集分子的活性数据并随机运行 19 次独立的 HypoGen 算法以获得 95% 的可信度。如果之前的 10 个药效团模型的 total cost 值均低于新产生的药效团模型的 total cost 值,则证明本研究建立的药效团是合理可行的。

#### 1.4 虚拟筛选数据库

首先,将经过验证的药效团模型 Hypo1 用于虚拟筛公共数据库:天然产物数据库(Natural products database),Ibscreen 和 Asinex(共 970 991 个分子)。采用 Ligand Pharmacophore Mapping 模块,参数设置选用弹性(Flexible)拟合方法及 FAST 的构象产生方法,最大忽略数为 0。考虑到分子的预测活性,本研究只保留 Fit value 大于 9 的分子。

其次,采用 ADMET 性质<sup>[11-12]</sup>对上一步筛选出来的所有苗头化合物做进一步筛选。ADMET 是对药物的吸收、分别、代谢、排泄和毒性进行全面的研究,药物早期 ADMET 性质评价可以减少药物毒性、不良反应的发生,从而显著提高药物研发的成功率。本研究采用 DS 中 ADMET Descriptors 模块进行计算,具体筛选标准为:①肠道吸收率:0~1;②水溶解度:-6~0;③血-脑脊液屏障通透性:2~4;④细胞色素 P450 2D6 抑制性:0;⑤肝毒性:0。

然后,采用里宾斯基五规则对 Ibscreen 和 Asinex 库中经过前两步筛选得到的分子做进一步筛选以提高小分子的口服利用度。先采用 DS 中计算分子特性(Calculate Molecular Properties)模块计算出分子的特征值。再通过目筛保留符合如下标准的分子:①相对分子质量(MW)<500;②氢键供体(HBD)<5;③氢键受体(HBA)<10;④疏水性(AlogP)<5;⑤可旋转键<10。

#### 1.5 分子对接

将 Natural products database 中经过药效团筛选和 ADMET 性质筛选的分子以及 Ibscreen 和 Asinex 库中经过药效团筛选、ADMET 性质筛选以及里宾斯基五规则筛选得到的分子采用 GOLD v5.0 软件<sup>[13]</sup>对接到 EphB4 蛋白的活性口袋中。对接前,从蛋白质结构数据库(www.rcsb.org)下载 EphB4 与抑制剂结合的晶体结构照片(PDB 编号:2VWY),去除蛋白的所有水分子并加全部氢。采用

50 次遗传算法(GA)及慢速 GA 搜索选项,以配体识别对接口袋,空腔半径为 6 Å,采用 gold\_serine\_protease\_VS 模板以及 GOLDScore 打分函数进行分子对接。其他参数采用默认值。为保证对接结果合理性,分子对接前,本研究先将晶体结构照片中的原始配体与蛋白做了重对接,计算其 RMSD 值并观察结合模式。若重对接结果合理,则采用相同的对接参数将经过上述药效团筛选及类药性筛选的分子用于分子对接。最后选取 GOLDScore 高且结合模式好的分子作为具有潜在 EphB4 抑制活性的候选化合物。

## 2 结果和讨论

### 2.1 药效团模型的建立

利用 25 个训练集分子(图 1)的结构-活性关系构建的 10 个药效团模型及其统计学数据列于表 1 中。消耗值(cost)分析是评估 3D-QSAR 药效团模型的最基本的手段。科雷尔值(Correl 值)表明药效团模型和结构-活性关系的相关性,其值越接近 1 越好。RMS 值不应超过 1。理想的总消耗值(total cost)应处于固定值(fixed cost)和空值(null cost)之间,并且应接近于前者并远离于后者;通常,如果 null cost 和 total cost 大于 60,表明有大于 90% 的可能性说明该模型具有好的相关性。构造值(Configuration)应小于 17.0。由表 1 可知,由 HypoGen 算法得到的 10 个药效团模型的  $\Delta$ cost 值均大于 60,且 Configuration 值为 15.50,表明这些模型的相关性较高。其中, Hypo1 具有最高的 Correl 值(0.96),最低的 RMS 值(0.89),与 fixed cost(89.20)最接近的 total cost(101.26),和最高的  $\Delta$ cost 值(89.14),因此被认为是最好的药效团模型。图 3-A 和 3-B 列出了 Hypo1 的 4 个药效学元素(2 个 HBA,1 个 HBD,1 个 HY)及其空间位置。

表 2 列出了 25 个训练集分子的实验  $IC_{50}$  及其由 Hypo1 计算得到的统计学数据。为评估 Hypo1 的预测能力,本研究将训练集分子按活性高低分为 3 组:高活性( $IC_{50} \leq 0.2 \mu\text{mol/L}$ , + + +),中等活性( $0.2 < IC_{50} \leq 2.0 \mu\text{mol/L}$ , + +),低活性( $IC_{50} > 2.0 \mu\text{mol/L}$ , +)。由表 2 可知,有 8 个高活性分子被准确预测,然而由 1 个高活性分子被误认为中等活性分子;有 7 个中等活性分子被准确预测,而有 2 个中等活性分子被误认为高活性分子,有 1 个被

误认为低活性分子;所有低活性分子被准确预测。因此表明 Hypo1 能较为准确的预测分子的活性值。Error 值是预测  $IC_{50}$  和实验  $IC_{50}$  的比值,其中“+”和“-”分别表明预测活性比实验活性高和低。本研究所有 Error 值的绝对值均小于3,表明 Hypo1 有能力准确预测分子的活性。图 2-C 和图 2-D 分别表示训练集中活性最好的化合物 1

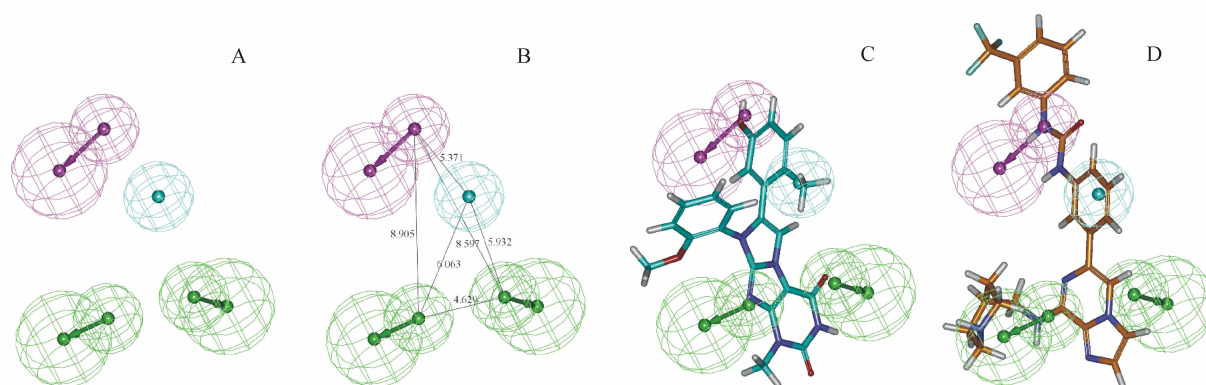
( $IC_{50} = 0.0016 \mu\text{mol/L}$ ) 和活性最差的化合物 25 ( $IC_{50} = 50.0 \mu\text{mol/L}$ ) 与 Hypo1 的匹配情况。很明显, Hypo1 的 4 个药效学元素均很好地与化合物 1 相匹配,然而仅有 3 个药效学元素与化合物 25 相匹配。另外,训练集分子的实验和预测活性的决定系数( $r^2$ ) 高达 0.91 (图 4)。因此,本研究认为 Hypo1 有能力预测训练集分子的活性。

**Table 1** Ten pharmacophore models and their statistical data by HypoGen run

Hypothesis	Correl value	RMS <sup>a</sup>	Total cost	$\Delta\text{cost}^b$	Chemical feature
1	0.96	0.89	101.26	89.14	HBA, HBA, HBD, HY
2	0.92	1.21	107.86	82.54	HBA, HBA, HBA, HY
3	0.92	1.18	109.40	81.00	HBA, HBA, HBD, HY
4	0.91	1.27	110.74	79.66	HBA, HBA, HBA, HY
5	0.90	1.32	111.47	78.93	HBA, HBA, HBA, HY
6	0.89	1.41	114.63	75.77	HBA, HBA, HBA, HY
7	0.89	1.42	114.78	75.62	HBA, HBA, HBA, HY
8	0.89	1.43	115.26	75.14	HBA, HBA, HY, HY, HY
9	0.87	1.53	119.13	71.17	HBA, HBA, HBD, HY
10	0.85	1.61	121.73	68.67	HBA, HY, HY, RA

Configuration value equals to 15.50, fixed cost equals to 89.20, and null cost equals to 190.40;

<sup>a</sup>RMS represents root mean square deviation; <sup>b</sup> $\Delta\text{cost}$  is the difference between null cost and total cost



**Figure 3** Pharmacophore model of EphB4 generated by HypoGen and training set molecules

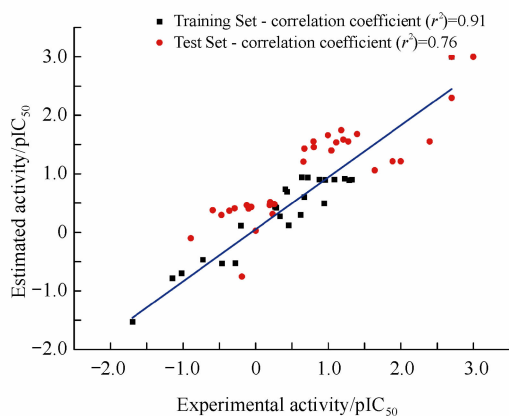
A: Best predictive model Hypo1; B: 3D spatial relationship and geometric parameters of Hypo1; C: Hypo1 aligning with the most active molecule 1 ( $IC_{50} = 0.0016 \mu\text{mol/L}$ ); D: Hypo1 aligning with the least active molecule 25 ( $IC_{50} = 50.0 \mu\text{mol/L}$ ). HBA, HBD and HY are color coded by green, magenta and cyan respectively

**Table 2** Statistical data of training set molecules based on Hypo1

Compd.	Experimental $IC_{50}/(\mu\text{mol/L})$	Predicted $IC_{50}/(\mu\text{mol/L})$	Fit value	Error	Experimental scale	Predicted scale
1	0.002	0.001	9.74	-1.4	+++	+++
2	0.047	0.126	7.70	2.7	+++	+++
3	0.051	0.128	7.69	2.5	+++	+++
4	0.059	0.121	7.72	2.0	+++	+++
5	0.082	0.125	7.70	1.5	+++	+++
6	0.110	0.127	7.70	1.2	+++	+++
7	0.114	0.321	7.29	2.8	+++	++
8	0.133	0.125	7.70	-1.1	+++	+++
9	0.190	0.115	7.74	-1.7	+++	+++
10	0.213	0.250	7.40	1.2	++	++
11	0.230	0.114	7.75	-2.0	++	+++

(Continued)

Compd.	Experimental $IC_{50}/(\mu\text{mol/L})$	Predicted $IC_{50}/(\mu\text{mol/L})$	Fit value	Error	Experimental scale	Predicted scale
12	0.240	0.504	7.10	2.1	++	++
13	0.350	0.760	6.92	2.2	++	++
14	0.370	0.203	7.49	-1.8	++	++
15	0.389	0.182	7.54	-2.1	++	+++
16	0.460	0.534	7.07	1.2	++	++
17	0.518	0.379	7.22	-1.4	++	++
18	0.538	0.366	7.24	-1.5	++	++
19	1.6	0.766	6.92	-2.1	++	+
20	1.9	3.350	6.28	1.8	+	+
21	2.9	3.398	6.27	1.2	+	+
22	5.3	2.937	6.33	-1.8	+	+
23	10.5	5.006	6.10	-2.1	+	+
24	13.96	6.073	6.02	-2.3	+	+
25	50.0	33.908	5.27	-1.5	+	+

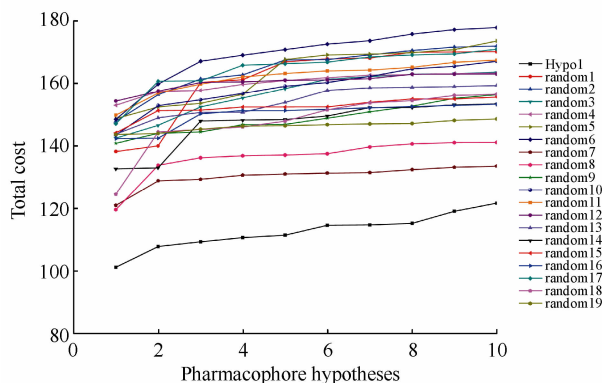


**Figure 4** Correlation curves between experimental and estimated activity values for training set and test set

## 2.2 药效团模型的验证

**2.2.1 测试集验证** 本研究首先用测试集验证方法验证 Hypo1 对测试集分子活性的预测能力。在此之前仍将测试集分子按活性高低划分为 3 组:高活性( $IC_{50} \leq 0.2 \mu\text{mol/L}$ , ++),中等活性( $0.2 < IC_{50} \leq 2.0 \mu\text{mol/L}$ , +),低活性( $IC_{50} > 2.0 \mu\text{mol/L}$ , -)。表 S1 列出了 32 个测试集分子的实验活性以及由 Hypo1 计算得到的预测活性。由图可知,所有 32 个分子的 Error 值均小于 10,初步表明 Hypo1 较高的预测能力。而且,16 个高活性分子均被准确预测;在 12 个中等活性分子中,有 9 个被准确预测,2 个被误认为高活性分子,1 个被误认为低活性分子;4 个低活性分子随被误认为中等活性分子,然而其 Error 值均小于 10,在误差允许范围之内。此外,由测试集分子的实验活性和预测活性的回归分析可知,两者的相关性系数高达 0.76(图 4),进一步表明 Hypo1 有较高的预测 EphB4 抑制剂活性的能力。

**2.2.2 Fischer 随机验证** 为进一步验证所建药效团模型的质量,本研究采用了基于 Fischer 的交叉验证方法。与 19 个随机产生的药效团假设相比,最初建立的药效团模型具有最低的 total cost(图 5),表明 Hypo1 具有高达 95% 的置信水平。



**Figure 5** Difference of total cost between the initial pharmacophore model (Hypo1) spread sheet and 19 random models spread sheets after Fischer's randomization run

## 2.3 基于 Hypo1 和类药性的虚拟筛选

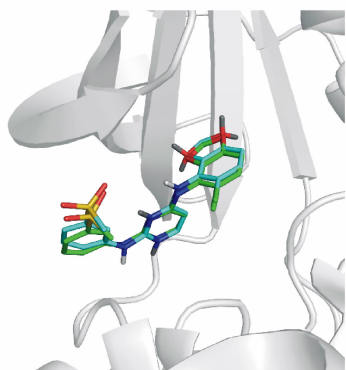
经过药效团验证之后,用 Hypo1 从化合物数据库(970 991 个分子)中筛选新颖的 EphB4 的潜在抑制剂分子。为保证分子的潜在抑制活性,本研究仅保留 Fit value 大于 9(预测  $IC_{50}$  小于  $0.006 \mu\text{mol/L}$ )的分子共 11 235 个。随后,采用 ADMET 预测做进一步筛选,得到 2 633 个分子。再单独对 Ibscreen 及 Asinex 库中通过上述筛选的分子做里宾斯基五规则筛选。然后将两部分化合物合并,最终得到 2 162 个分子用于对接研究。

## 2.4 分子对接

为验证对接参数的合理性及对接结果的可信

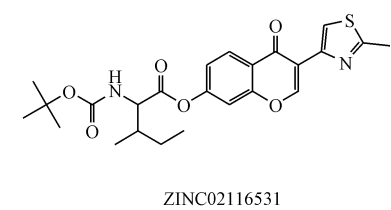
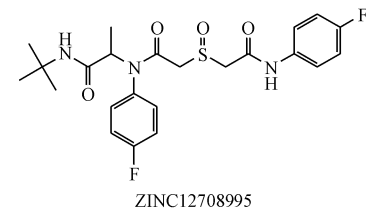
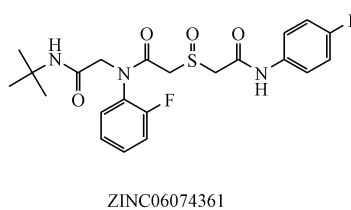
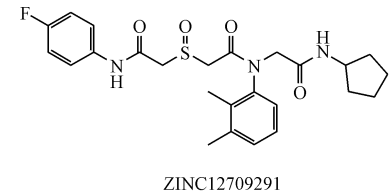
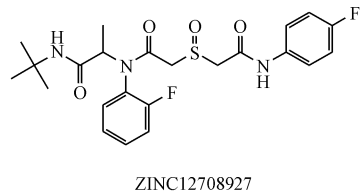
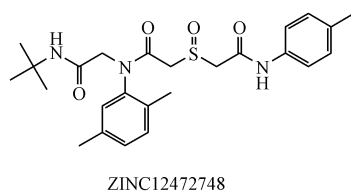
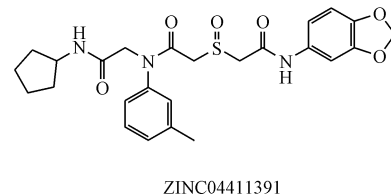
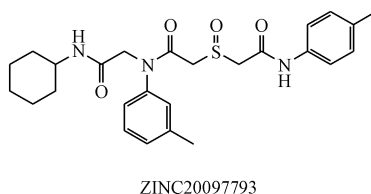
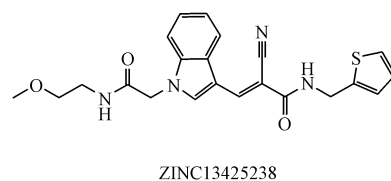
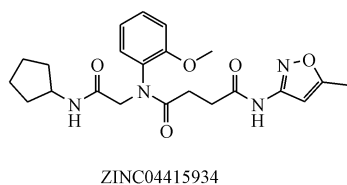
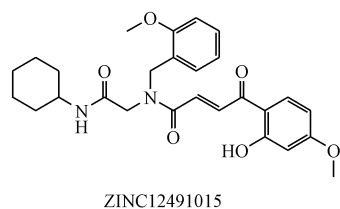


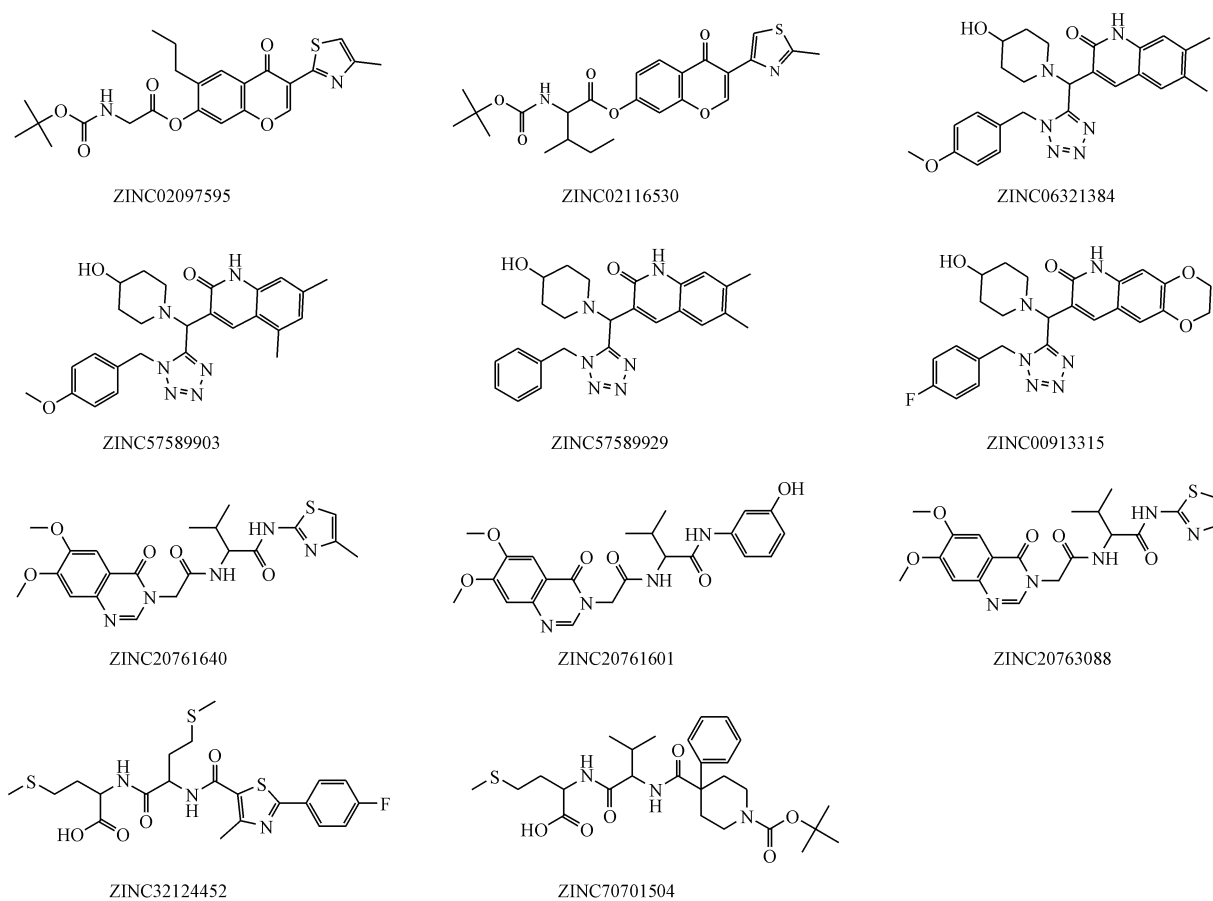
性,分子对接前,本研究将晶体结构照片中的原始配体与蛋白进行了重对接,其配体在重对接前后的叠合模式见图 6, RMSD 值为  $0.460\ 2\ \text{\AA}$ , GOLD-Score 为 67.90。高的叠合程度和低的 RMSD 值表明所选取的对接参数及打分函数的合理性,因此此参数可用于后续分子对接研究。



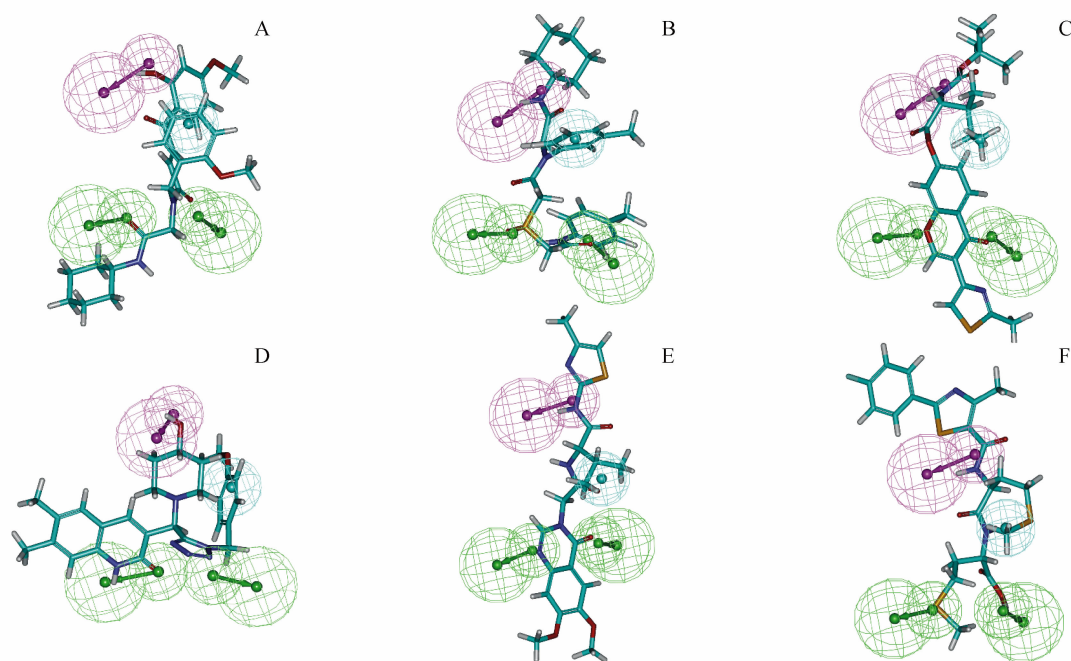
**Figure 6** Superimposition of crystallized and docked conformations in GOLD. Green and cyan sticks represent crystallized and docked conformations respectively

将经过初筛后的 2 162 个分子对接到 EphB4 的活性位点。根据 GOLD 对接结果,以 70 为截断值选取 GOLDScore 大于 70 且结合模式好的分子共 23 个作为 EphB4 的潜在抑制剂分子,其 2D 结构如图 7 所示。由图可知,这 16 个分子属于 6 种骨架。ZINC12491015, ZINC04415934, ZINC13425238, 和 ZINC04083305 属于同一骨骨架; ZINC20097793, ZINC04411391, ZINC12472748, ZINC12708927, ZINC12709291, ZINC06074361 和 ZINC12708995 属于同一骨架; ZINC02116531, ZINC02097595, 和 ZINC02116530 属于同一骨架; ZINC06321384, ZINC57589903, ZINC57589929 和 ZINC00913315 属于同一骨架; ZINC20761640, ZINC20761601, 和 ZINC20763088 属于同一骨架; ZINC32124452 和 ZINC70701504 属于同一骨架。每个骨架中,挑选一个分子分析。6 个不同骨架的分子与 Hypo1 的叠合图如图 8 所示,很明显,这些分子均与 Hypo1 的 4 个药效学元素完美匹配。





**Figure 7** 2D structures of twenty-three potential inhibitors of EphB4

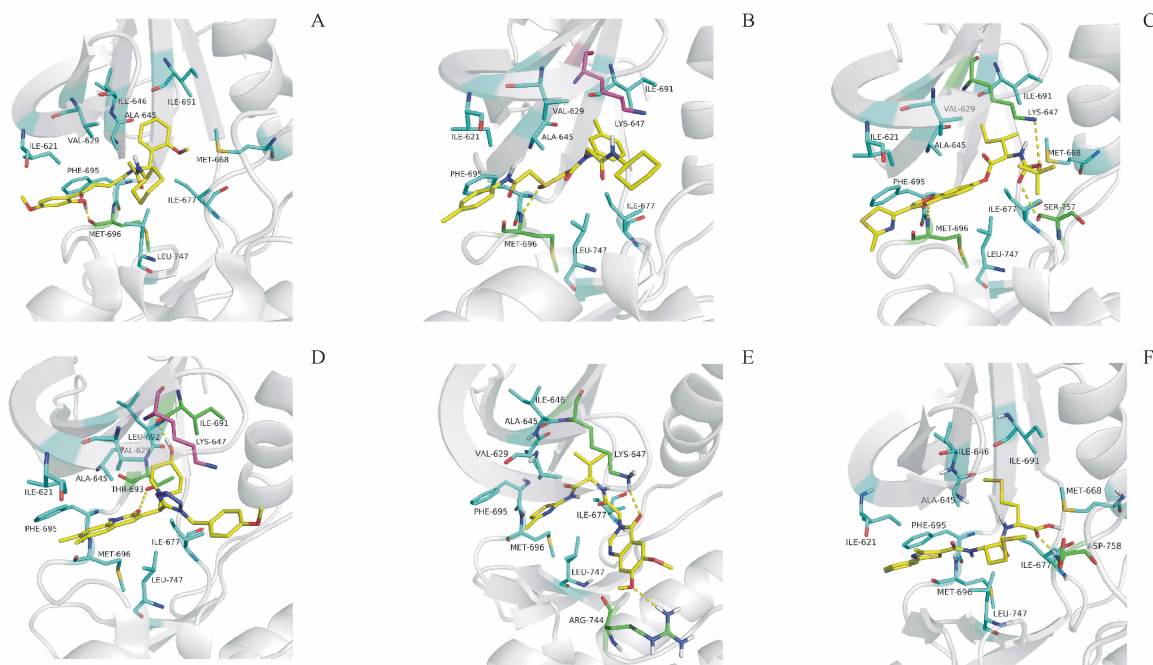


**Figure 8** Superimposition of six potential inhibitors of EphB4 and Hypo1. A-F represent ZINC12491015, ZINC20097793, ZINC02116531, ZINC06321384, ZINC30761640, and ZINC32124452, respectively



图 9 是这 6 个不同骨架的分子与 EphB4 蛋白经 GOLD 对接后的结合模式图。化合物 ZINC12491015 苯环上的羟基氢与 MET696 的羰基氧形成一个氢键,起到配体-蛋白之间很好的固定作用;整个分子与周围疏水性氨基酸 ILE621、VAL629、ALA645、ILE646、MET668、ILE677、ILE691、PHE695 和 LEU747 形成疏水相互作用,结合稳定,结合模式较好。化合物 ZINC20097793 也与 MET696 形成一个氢键,并暴露于由氨基酸残基 ILE621、VAL629、ALA645、ILE677、ILE691、PHE695 和 LEU747 组成的疏水口袋中;另外,该分子位于结合口袋内部的甲基与 LYS647 形成苯-阳离子相互作用,对于配体-蛋白的结合稳定性起到至关重要的作用。化合物 ZINC02116531 也暴露在由疏水氨基酸组成的疏水口袋中;并且除了与

MET696 形成氢键外,还与 LYS647 及 SER757 各形成一个氢键,进一步增加了结合稳定性。与前 3 个化合物的作用模式较为不同,化合物 ZINC06321384 与 ILE691 和 THR693 各形成一个氢键;其 4-羟基哌啶与 LYS647 形成苯-阳离子相互作用,结合模式较好。与其他化合物相比,ZINC20761640 作用于 EphB4 活性口袋更深处,与 LYS647 和 ARG744 形成氢键,并与周围的亮氨酸、异亮氨酸、丙氨酸、缬氨酸、甲硫氨酸、苯丙氨酸形成疏水相互作用。化合物 ZINC32124452 与 ASP758 形成一个作用较强的氢键;两条侧链烷烃与周围的 ALA645、ILE646、ILE677、ILE691、MET696 和 LEU747 形成疏水相互作用;其苯环又与 PHE695 及 ILE621 形成疏水相互作用,与蛋白结合稳定。



**Figure 9** Binding modes of the molecules with EphB4. The yellow sticks in A-F represent compound ZINC12491015, ZINC20097793, ZINC02116531, ZINC06321384, ZINC30761640, and ZINC32124452 respectively. Cyan, green and magenta sticks represent key amino acids combined with ligands. Dash yellow lines represent the H-bond between ligand and protein

表 3 列出了这 23 个化合物及晶体结构照片中原始配体匹配药效团的 Fit value 及 GOLD 对接得分值。由表中数据可知,这 23 个分子的 Fit value 均远高于原始配体(8.08),表明它们的预测活性均高于原始配体的 EphB4  $IC_{50} = 0.09 \mu\text{mol/L}$ 。并且,筛选出的分子的 GOLDScore 值都在 70 分以上,均大于原始配体的对接得分值(67.90)。其中化合

物 ZINC12491015、ZINC20097793、ZINC04411391、ZINC02116531 和 ZINC32124452 的 GOLDScore 高达 75 分以上,其相应的 Fit value 也在 9.50 分以上,不仅表明本研究所筛选出的分子与 EphB4 蛋白的结合能力更强、预测活性值更高,同时表明了采用 GOLD 对接的得分值与用 Hypo1 预测的分子的活性值存在很高的相关性,间接表明了本研究所建药效

团及采用的对接方式的合理性和可信性。由此推断,这 23 个化合物具有较高的 EphB4 潜在抑制活性。

**Table 3** Statistical data of 23 potential inhibitors and the native ligand of EphB4

No.	Name	Fit value	GOLDScore
1	ZINC12491015	9.59	79.67
2	ZINC04415934	9.39	72.83
3	ZINC13425238	9.05	72.38
4	ZINC04083305	9.66	71.42
5	ZINC20097793	9.89	75.69
6	ZINC04411391	9.80	75.02
7	ZINC12472748	9.81	72.27
8	ZINC12708927	9.82	71.45
9	ZINC12709291	9.96	71.42
10	ZINC06074361	9.64	71.29
11	ZINC12708995	9.41	71.06
12	ZINC02116531	9.71	75.20
13	ZINC02097595	9.00	73.60
14	ZINC02116530	9.53	71.53
15	ZINC06321384	9.34	73.79
16	ZINC57589903	9.41	73.33
17	ZINC57589929	9.49	72.07
18	ZINC00913315	9.58	72.01
19	ZINC20761640	9.64	74.29
20	ZINC20761601	9.63	72.78
21	ZINC20763088	9.42	70.80
22	ZINC32124452	9.50	76.06
23	ZINC70701504	9.45	74.67
24	Native ligand	8.08	67.90

### 3 结 论

本研究首先通过 25 个结构、活性多样化的训练集分子构建 3D 药效团模型,其中 Hypo1 具有 4 个药效学元素(2 个 HBA,1 个 HBD,1 个 HY),以及最高的 *Correl* 值(0.96),最低的 RMS 值(0.89),与 fixed cost(89.20)最接近的 total cost(101.26),和最高的  $\Delta$ cost 值(89.14),被认为是最好的药效团模型。随后,Hypo1 经过测试集方法及 Fischer 随机方法用以验证可信性。然后用基于药效团筛选和类药性筛选方法筛选包含 970 991 个分子的数据库,共得到 2 162 个分子。再将这些分子利用 GOLD 软件对接到 EphB4 蛋白的活性位点,通过对接打分值分析及结合模式分析,最终得到 6 种全新骨架共 23 个化合物作为 EphB4 的潜在抑制剂分子。综上所述,Hypo1 模型有能力从数据库中筛选得到具有 EphB4 潜在抑制活性的分子;本研究筛

选得到的这 23 个分子可作为抑制 EphB4 蛋白的候选化合物用于后续研究。

### 参 考 文 献

- [1] Zhang XG,Yao LB. The signal transduction pathway and function of Eph receptors and its ligands[J]. *Prog Biochem Biophys*(生物化学和生物物理学进展),2001,**28**(4):498-501.
- [2] Sun JF,Liu HJ,Li C. Relationship between expressions of EphB4 with angiogenesis in human colorectal carcinoma[J]. *J Frontier Med*(医药前沿),2013,**5**(21):250-255.
- [3] Bardelle C,Cross D,Davenport S,*et al.* Inhibitors of the tyrosine kinase EphB4. Part 1:structure-based design and optimization of a series of 2,4-bis-anilinopyrimidines[J]. *Bioorg Med Chem Lett*,2008,**18**(2):2776-2780.
- [4] Bardelle C,Coleman T,Cross D,*et al.* Inhibitors of the tyrosine kinase EphB4. Part 2:structure-based discovery and optimisation of 3,5-bis substituted anilinopyrimidines[J]. *Bioorg Med Chem Lett*,2008,**18**(5):5717-5721.
- [5] Barlaam B,Ducray R,Read J,*et al.* Inhibitors of the tyrosine kinase EphB4. Part 4:discovery and optimization of a benzylic alcohol series[J]. *Bioorg Med Chem Lett*,2011,**21**(22):2207-2211.
- [6] Lafleur K,Huang D,Zhou T,*et al.* Structure-based optimization of potent and selective inhibitors of the tyrosine kinase erythropoietin producing human hepatocellular carcinoma receptor B4 (EphB4)[J]. *J Med Chem*,2009,**52**(6):6433-6446.
- [7] Lafleur K,Dong J,Huang D,*et al.* Optimization of inhibitors of the tyrosine kinase EphB4. 2. Cellular potency improvement and binding mode validation by X-ray crystallography[J]. *J Med Chem*,2013,**56**(1):84-96.
- [8] Miyazaki Y,Nakano M,Sato M,*et al.* Design and effective synthesis of novel templates,3,7-diphenyl-4-amino-thieno-furo-[3,2-c]pyridines as protein kinase inhibitors and in vitro evaluation targeting angiogenetic kinases[J]. *Bioorg Med Chem Lett*,2007,**17**(2):250-254.
- [9] Mitchell SA,Danca MD,Blomgren PA,*et al.* Imidazo[1,2-a]pyrazine diaryl ureas;inhibitors of the receptor tyrosine kinase EphB4[J]. *Bioorg Med Chem Lett*,2009,**19**(6):6991-6995.
- [10] Fischer R. *The Principle of Experimentation,Illustrated by a Psycho-Physical Experiment,in the Design of Experiments*[M]. New York:Hafner Publishing Company,1966:129-153.
- [11] Walters WP,Murcko MA. Prediction of 'drug-likeness'[J]. *Adv Drug Delivery Rev*,2002,**54**(2):255-271.
- [12] Cheng A,Merz KM Jr. Prediction of aqueous solubility of a diverse set of compounds using quantitative structure-property relationships[J]. *J Med Chem*,2003,**46**(3):3572-3580.
- [13] Vermunt JE,Magidson J. *Latent GOLD 5.0 Upgrade Manual*[M]. Belmont:Statistical Innovations Inc.,2013:540-555.