

文章编号 1006-8147(2018)02-0131-04

论著

## 基于虚拟筛选的 3D QSAR 药效团设计 CDC25B 抑制剂

李煜, 晋文燕, 李红莲, 王润玲, 马英

(天津医科大学药学院, 天津市临床药物关键技术重点实验室, 天津 300070)

**摘要** 目的: 用计算机辅助药物设计的方法发现潜在的 CDC25B 抑制剂。方法: 用 Hypogen 方法研究 CDC25B 抑制剂。cost 分析、测试集预测和 Fisher 检验用来验证该模型的可靠性。随后, 运用 hypo-1-CDC25B 对 ZINC 数据库进行筛选, 得到符合成药五规则的 26 个化合物, 26 个化合物进行分子对接得到 6 个对接得分高的化合物。结果: 通过分子对接研究, 发现 6 个化合物有较好的亲和力。结论: 发现 6 个潜在的 CDC25B 抑制剂, 这有助于发现治疗癌症的强有力的先导化合物。

**关键词** CDC25B; Hypogen; 分子对接

中图分类号 R914.2

文献标志码 A

### 3D QSAR pharmacophore based on virtual screening for design of CDC25B inhibitors

LI-Yu, JIN Wen-yan, LI Hong-lian, WANG Run-ling, MA-Ying

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

**Abstract Objective:** To explore potential the cell division cycle 25B(CDC25B) inhibitors by the method of computer-aided drug design.

**Methods:** In this study, 3D QSAR pharmacophore models for CDC25B inhibitors were developed by the module of Hypogen. Three methods (cost analysis, test set prediction, and Fisher's test) were applied to validate whether the models could be used to predict the biological activities of compounds. Subsequently, 26 compounds meeting Lipinski's rules of five were obtained by the virtual screening of the Hypo-1-CDC25B against ZINC databases. **Results:** It was discovered that six identified molecules had satisfying binding affinity.

**Conclusion:** Thus, this study would be helpful to discover potent lead compounds for the treatment for cancers.

**Key words** cell division cycle25B; Hypogen; docking

癌症的特征是细胞周期会发生紊乱并且细胞会恶性增殖。2012 年的癌症患者约有 1 410 万名, 2014 年导致我国人群期望寿命损失的疾病以癌症为首, 占比高达 20.73%, 癌症患者发病率将在 2020 年前增至 1 685 万<sup>[1]</sup>。因此, 发现特异性抗癌药物迫在眉睫。CDC25B 是由细胞周期蛋白依赖性激酶去除磷酸盐和进一步激活蛋白激酶复合物, 从而影响细胞周期的进展。它通过调整 CDK2/cyclin 和 CDK1/CyclinA 在 G2/M 转变中发挥重要作用。CDC25B 在肿瘤组织中的表达水平是正常水平的 3 倍, 这反过来又会促进肿瘤的生长。此外, 研究表明 CDC25B 异常出现在乳腺癌、肺癌、头颈癌等癌症的癌细胞周期的早期阶段<sup>[2]</sup>。细胞增殖水平异常已经涉及到许多人类癌症。因此, CDC25B 成为治疗癌症的潜在靶点。本研究利用计算机辅助药物设计(CADD)的方法来设计 CDC25B 抑制剂。我们建立了 CDC25B 的 3D-

QSAR 药效团模型筛选 ZINC 数据库, 得到先导化合物 Zinc-09340418。随后, 对该化合物进行片段替换, 通过虚拟筛选得到 fitvalue 高于 7.7 的化合物并对这些化合物进行 ADME 预测。利用柔性对接进一步筛选出对接得分高于 ZINC 09340418 的化合物。此研究发现 6 个潜在的 CDC25B 抑制剂。

#### 1 材料和方法

1.1 化合物的选择和数据集的准备 选取 38 个已报道的 CDC25B 抑制剂建立药效团模型。训练集的化合物满足(1)至少有 16 个不同的抑制剂;(2)生物活性跨度至少是 4 个数量级。其中 17 个化合物(图 1)为训练集, 剩下的 21 个化合物(图 2)作为测试集构建药效团模型。

1.2 药效团模型的构建和验证 利用 DS 中 Hypogen 算法构建药效团模型<sup>[3]</sup>。HBA 和 HBD 的数量分别设定为 1-2, HYP 的数量设置为 2-3, Uncert 设为 2。产生的药效团模型用 fixed cost, null cost 和 total cost 来进行评估。用 cost analyses, test sets prediction 和 Fischer randomization 验证药效团的可靠性<sup>[4]</sup>。

基金项目 国家级大学生创新创业训练计划项目;天津市自然科学基金重点项目基金资助(16JCZDJC32500)

作者简介 李煜(1996-),男,学士在读,研究方向:计算机辅助药物设计;通信作者:马英, E-mail: maying@tmu.edu.cn。

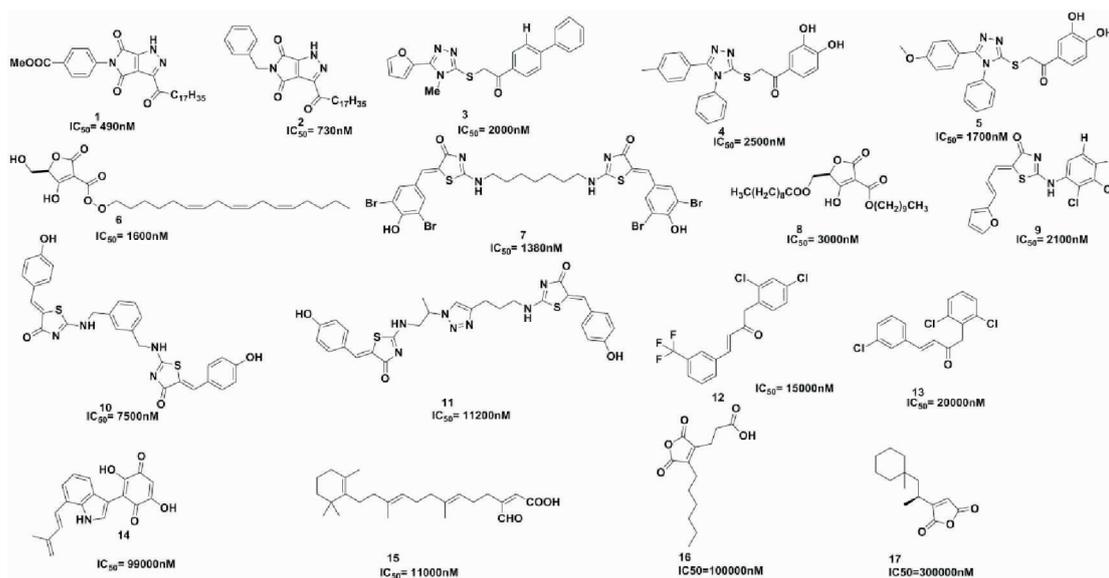


图1 CDC25B 抑制剂的 17 个训练集化合物分子结构

Fig 1 Chemical structures of 17 training set compounds for CDC25B

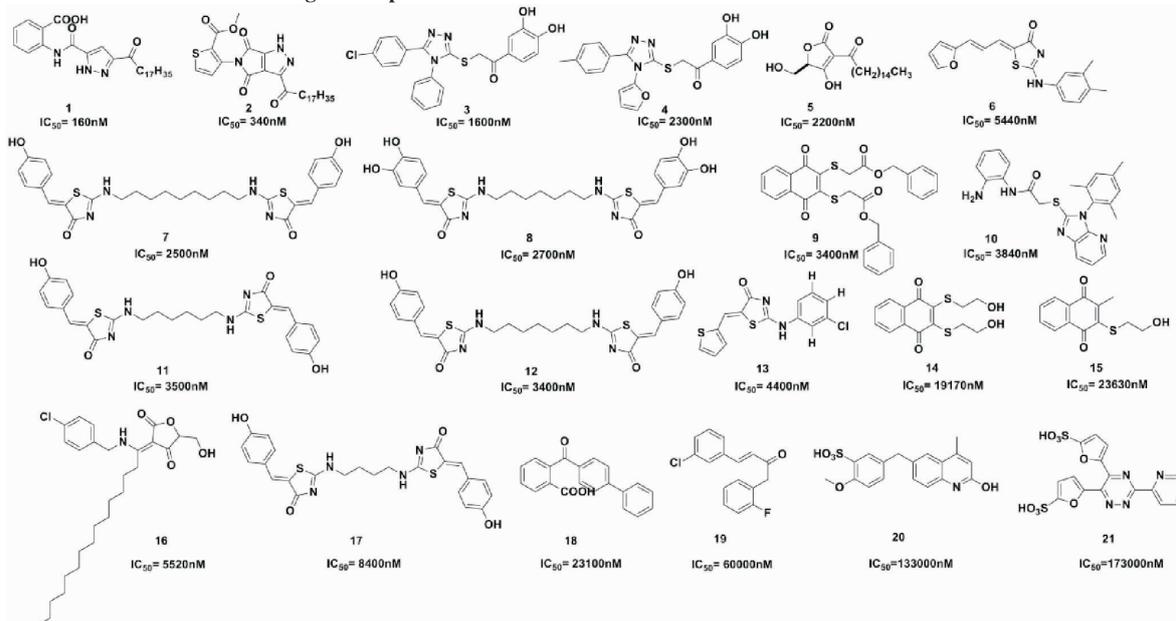


图2 CDC25B 抑制剂的 21 个测试集化合物分子结构

Fig 2 Chemical structures of 21 test set compounds for CDC25B

1.3 虚拟筛选和片段替换 对 ZINC 数据库进行虚拟筛选, 筛选出 fit value 高于 7.7 的化合物, 得到 Zinc-09340418。利用‘Replace Fragment’替换该化合物的片段<sup>[5]</sup>, 对得到的新化合物进行药效团筛选, 得到 fit value 高的化合物。对这些化合物进行 ADME 预测, 筛选出符合利平斯基的成药五规则化合物。

1.4 柔性对接 DS 中的 Flexible Docking 来分析配体和靶蛋白之间的相互作用。第一步, 通过“Clean protein”工具对 CDC25B (PDB ID: 1cwt) 进行蛋白准备<sup>[6]</sup>。第二步是使用 chiflex 计算受体构象能量<sup>[7]</sup>。第三步是配体准备。第四步是定义活性位点, 包含 Leu398, Glu431, Glu446, Cys473, Glu474, Phe475,

Ser476, Ser477, Glu478, Arg 479, Arg482, Cys484<sup>[8-11]</sup>。第五步是使用 ChiRotor<sup>[12]</sup> 细化蛋白质侧链。第六步是用 Flexible Docking 进行分子对接。

## 2 结果

2.1 药效团模型和验证 药效团模型的优劣判断主要是依据 RMS、Correlation coefficient Configuration cost 值以及  $\Delta$ cost 值 (Null cost 与 Total cost 的差值)。利用 DS 中 HypoGen 算法产生 10 个 CDC25B 抑制剂的药效团模型, 所得结果如表 1 所示。Hypo-1-CDC25B 为最好的模型, 包含 1 个 HBA, HBD, 一个 HYP 和 1 个 RA。用 Fischer randomization 方法验证该药效团的可靠性。

表 1 HypoGen 算法产生的 10 个 CDC25B 抑制剂的药效团模型的结果

Tab 1 Results obtained from pharmacophore hypotheses generation using the training set molecules of CDC25B

Hypothesis	Total cost	$\Delta$ cost	RMS	Correlation( $r^2$ )	Features
1 <sup>a</sup>	82.893	62.411	1.413	0.908	HBA,HBD,HYA,RA
2	91.914	53.39	1.783	0.748	HBA,HBD,HYA,RA
3	92.302	53.002	1.791	0.846	HBA,HBD,HYA,RA
4	94.146	51.158	1.855	0.834	HBA,HBD,HYA,RA
5	98.274	47.03	1.978	0.808	HBA,HBD,HYA,RA
6	99.533	45.771	1.990	0.807	HBA,HBD,HYA,RA
7	100.393	44.911	2.039	0.795	HBA,HBD,HYA,RA
8	100.662	44.642	2.031	0.797	HBA,HBD,HYA,RA
9	101.216	44.088	2.062	0.789	HBA,HBD,HYA,RA
10	101.886	43.418	2.058	0.791	HBA,HBD,HYA,RA

Null cost=145.304, Fixed cost=64.893 7, Configuration cost=14.416 9; HBA: 氢键受体; HBD: 氢键给体; HYA: 疏水中心; RA: 芳香环中心; a: 最好的药效团有最好的 RMS,最高相关系数和最大 cost 差异

2.2 虚拟筛选 利用 ZINC 数据库的虚拟筛选技术, 得到 fit value 高的化合物 Zinc-09340418。以 Zinc-09340418 为先导化合物进行片段替换(图 3), 发现具有高抑制活性和良好 fit value 值的化合物。这些化合物用利平斯基的五法则进一步筛选, 最后 26 个化合物对接研究。与 Zinc-09340418 相比, 26 个化合物具有较高的 fit value。以 compound1 为代表, Zinc-09340418 和 compound1 在 Hypo-1-CDC25B 模型中的结合模式见图 4。Zinc-09340418 和 compound1 能与药效特征(氢键受体、氢键给体、疏水基团、芳环中心)对应。

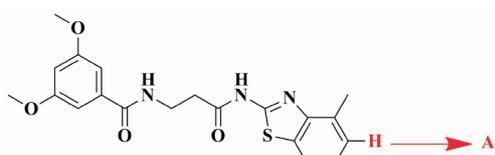
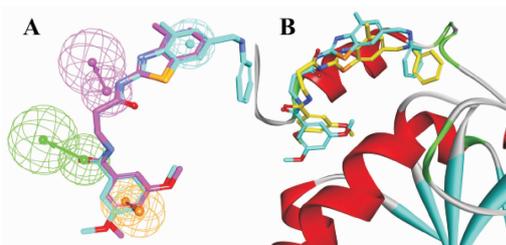


图 3 A 通过数据库筛选进行片段替换

Fig 3 A was replaced by searching the fragment libraries



(A) ZINC 09340418 和化合物 1 在药效团模型 Hypo-1-CDC25B 中的匹配图。绿色球表示氢键受体, 粉红色球表示氢键给体, 蓝色球为疏水基团, 黄色球为芳环中心。化合物 1 和 ZINC 09340418 分别用蓝绿色和红色表示。(B) 化合物 1(黄色)和 ZINC 09340418(蓝绿色)在 Hypo-1-CDC25B 药效团模型中与 CDC25B 的结合构象

图 4 ZINC09340418 和 compound1 与 CDC25B 药效团模型的结合图  
Fig 4 Illustration to show the pharmacophore model of CDC25B generated by Hypogen and the conformation obtained by docking ZINC09340418 and compound 1, respectively, to CDC25B

2.3 柔性对接 把 26 个化合物对接到 CDC25B 蛋白结合口袋。以 Zinc-09340418 与受体复合物的结合能为基准(-CDOCKER ENERGY=38.6714)选出 6 个对接得分高于 Zinc-09340418 的化合物(图 5), 表 2 是 6 个化合物的对接得分以及 ADME 参数值。通过对接模拟得到的受体-配体相互作用的结果表明, 配体与受体结合的关键残基与以前的报告完全一致。图 6A 是 compound1 对接到 CDC25B 活性位点的 2D 图。图 6B 是 Zinc-09340418 对接到 CDC25B 活性位点的 2D 图。而 compound1 不仅产生了相同的范德瓦尔斯与疏水相互作用, 还多产生了一个 Pi 相互作用和疏水作用。Compound1 与 Zinc-09340418 相比多了一个芳香环。

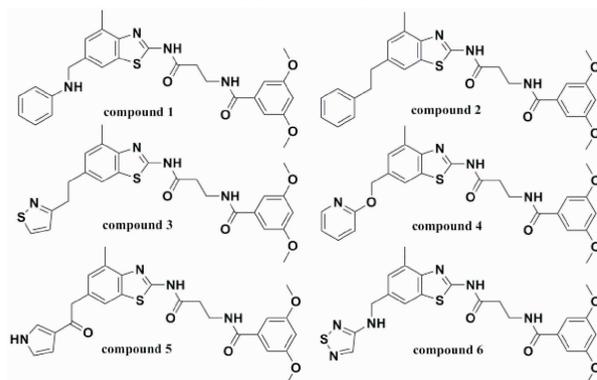
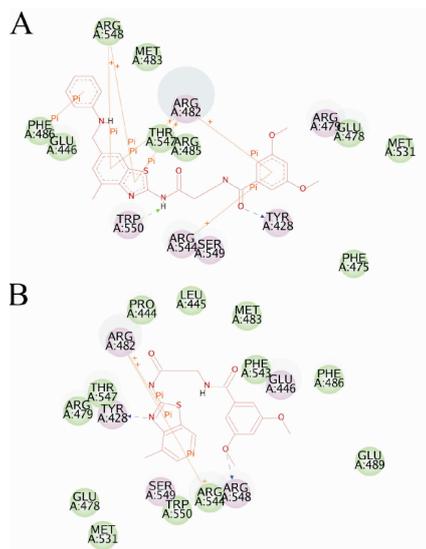


图 5 6 个潜在的 CDC25B 抑制剂的分子结构

Fig 5 Six potential inhibitors against PTP1B and CDC25B



(A) compound1 with CDC25B (PDB ID: 1CWT) 的配体-蛋白相互作用图, 蓝色和绿色的虚线表示 compound 1 和受体的氢键相互作用, 橙色线表示 Pi 相互作用。(B) ZINC 09340418 和 CDC25B (PDB ID: 1CWT) 配体-蛋白相互作用图, 蓝色虚线表示 ZINC 09340418 和受体的氢键相互作用, 橙色线表示 Pi 相互作用。粉红色圆形表示氢键、电荷或极性相互作用; 绿色圆形表示范德华力; 蓝色光环表示氨基酸相互作用的溶剂残留, 直径大小与溶剂残留成正比

图 6 化合物 1 和 ZINC 09340418 对接到 CDC25B (1CWT) 活性位点的 2D 图

Fig 6 Illustration to show 2D diagram of compound 1 and ZINC 09340418 docked into the active site of the CDC25B (1CWT)

表 2 按对接得分排名的前 6 个化合物与 CDC25B 的对接得分及其他参数

Tab 2 The 6 compounds ranked roughly according to the strength of their docking scores to the CDC25B

化合物	-CDOCKER / (Kcal/mol)	EstimatedIC <sub>50</sub> / (nmol/L)	ADME 参数					
	CDC25B	CDC25B-1	PSA <sup>a</sup>	logPo/w <sup>b</sup>	logS <sup>c</sup>	Molecular weight	HBD	HBA
Zinc-09340418	38.671 4	69.720 4	89.34	3.33	-4.75	399.46	2	5
化合物 1	41.231 1	47.530 9	102.15	4.42	-5.35	503.59	3	6
化合物 2	39.851 4	58.432 9	89.34	5.46	-5.98	502.61	2	5
化合物 3	39.740 9	59.406 7	100.60	4.60	-5.67	510.63	2	6
化合物 4	39.401 1	62.507	109.53	4.29	-5.26	506.57	2	7
化合物 5	39.268 8	63.757 9	121.70	3.86	-5.17	506.57	3	6
化合物 6	39.110 5	65.285 6	124.68	3.63	-5.51	512.60	3	8

各参数的参考范围:a:7.0-200.0; b:22.0-6.5; c:-6.5-0.5

### 3 讨论

本研究的目的是发现新型有效的抗癌药物。我们使用新技术“Hypogen”构建药效团模型Hypo-1-CDC25B。随后,满足利平斯基成药五规则的 26 个化合物是通过对 ZINC 数据库的虚拟筛选得到。基于药效基团的虚拟筛选和对接分析后,选出得分最高的前 6 个化合物。与 ZINC 09340418 相比,compound1 不仅有相似的功能,在对接中有更好的结合构象。6 个化合物可能成为潜在的候选药物,这一发现提供一种对抗癌症的很有前途的方法。

#### 参考文献:

- [1] 蔡玥,孟群,王才有,等.2015-2020 年我国居民预期寿命测算及影响因素分析[J].中国卫生统计,2016,33(8):2
- [2] Sasaki H, Yukiue H, Kobayashi Y, et al. Expression of the cdc25B gene as a prognosis marker in non-small cell lung cancer[J]. Cancer letters, 2001, 173(2):187
- [3] Garg D, Gandhi T, Mohan C G. Exploring QSTR and toxicophore of hERG K<sup>+</sup> channel blockers using GFA and HypoGen techniques [J]. J Mol Graph Mode, 2008, 26(6):966
- [4] Kumar R, Son M, Bavi R, et al. Novel chemical scaffolds of the tumor marker AKR1B10 inhibitors discovered by 3D QSAR pharmacophore modeling [J]. Acta Pharmacol Sin, 2015, 36(8):998
- [5] Gangwal R P, Das N R, Thanki K, et al. Identification of p38 alpha

MAP kinase inhibitors by pharmacophore based virtual screening[J]. J Mol Graph Mode, 2014, 49: 18

- [6] Ma Y, Jin Y Y, Wang Y L, et al. The discovery of a novel and selective inhibitor of PTP1B over TCPTP: 3D QSAR pharmacophore modeling, virtual screening, synthesis, and biological evaluation[J]. Chem Drug Design, 2014, 83(6):697
- [7] Spassov V Z, Yan L, Flook P K. 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[J]. Protein Sci, 2007, 16, (3): 494
- [8] Martin K R, Narang P, Medina-Franco J L, et al. Integrating virtual and biochemical screening for protein tyrosine phosphatase inhibitor discovery[J]. Methods, 2014, 65 (2):219
- [9] Lavecchia A Di, Giovanni C, Pesapane A, et al. Discovery of new inhibitors of Cdc25B dual specificity phosphatases by structure-based virtual screening[J]. J Med Chem, 2012, 55(9): 4142
- [10] Huang N, Kalyanaraman C, Bernacki K, et al. Molecular mechanics methods for predicting protein-ligand binding[J]. PCCP, 2006, 8(44): 5166
- [11] Braud E, Goddard M L, Kolb S, et al. Novel naphthoquinone and quinolinedione inhibitors of CDC25 phosphatase activity with antiproliferative properties[J]. Bioorg Med Chem, 2008, 16 (19): 9040
- [12] Spassov V Z, Yan L, Flook P K. 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[J]. Protein Sci, 2007, 16(3):494

(2017-09-05 收稿)