

基于Tomoper CoMFA方法对香豆素类苯扎酰胺类组蛋白去乙酰化酶抑制剂的3D-QSAR研究及分子设计

全建波^{1,2*}, 边帅^{1,2}, 罗钉^{1,2}, 张星^{1,2}, 冯怡^{1,2}

¹陕西科技大学化学与化工学院, 陕西 西安

²陕西省轻工助剂重点实验室, 陕西 西安

Email: *jianbotong@aliyun.com

收稿日期: 2020年11月8日; 录用日期: 2020年11月23日; 发布日期: 2020年11月30日

摘要

组蛋白去乙酰化酶(HDACs)是一类蛋白酶, 其对染色体的结构修饰和基因表达调控发挥着重要的作用, 可以作为治疗癌症和其他疾病的治疗靶点。组蛋白去乙酰化酶抑制剂(HDACIs)可增加细胞内组蛋白的乙酰化程度, 抑制肿瘤细胞的增殖。本文采用易位体比较分子力场(Tomoper CoMFA)方法对一系列以香豆素为基础的苯扎酰胺类化合物作为HDACs抑制剂的HCT116细胞系和A2780细胞系进行三维定量构效关系研究(3D-QSAR), 对生成的模型进行交叉验证和非交叉验证。HCT116细胞系交叉验证系数 $q^2 = 0.517$, 非交叉验证系数 $r^2 = 0.880$, A2780细胞系交叉验证系数 $q^2 = 0.572$, 非交叉验证系数 $r^2 = 0.869$ 。最后采用Topomer search技术在ZINC数据库中进行虚拟筛选, 最终设计出16个具有更高活性的新型HDACIs化合物。

关键词

组蛋白去乙酰化酶, 易位体比较分子力场, 组蛋白去乙酰化酶抑制剂, 三维定量构效关系

3D-QSAR Research and Molecular Design of Coumarin Benzamide Histone Deacetylase Inhibitor Based on Tomoper CoMFA Method

Jianbo Tong^{1,2*}, Shuai Bian^{1,2}, Ding Luo^{1,2}, Xing Zhang^{1,2}, Yi Feng^{1,2}

¹College of Chemistry and Chemical Engineering, Shaanxi University of Science and Technology, Xi'an Shaanxi

²Shaanxi Key Laboratory of Light Chemical Auxiliaries, Xi'an Shaanxi

Email: *jianbotong@aliyun.com

Received: Nov. 8th, 2020; accepted: Nov. 23rd, 2020; published: Nov. 30th, 2020

*通讯作者。

文章引用: 全建波, 边帅, 罗钉, 张星, 冯怡. 基于 Tomoper CoMFA 方法对香豆素类苯扎酰胺类组蛋白去乙酰化酶抑制剂的 3D-QSAR 研究及分子设计[J]. 物理化学进展, 2020, 9(4): 38-50. DOI: 10.12677/JAPC.2020.94005

Abstract

Histone deacetylases (HDACs), a class of proteases, play an important role in the structural modification of chromosomes and the regulation of gene expression, and can be used as therapeutic targets for the treatment of cancer and other diseases. Histone deacetylase inhibitors (HDACIs) can increase intracellular histone acetylation and inhibit tumor cell proliferation. In this paper, a series of coumarin-based benzamide compounds as HDACs inhibitors HCT116 cell lines and A2780 cell lines were studied by using Tomoper CoMFA method for 3D-Quantitative structure-activity Relationship (3D-QSAR), and the generated models were cross-verified and non-cross-verified. HCT116 cell line cross validation coefficient $q^2 = 0.517$, non-cross validation coefficient $r^2 = 0.880$, A2780 cell line cross validation coefficient $q^2 = 0.572$, non-cross validation coefficient $r^2 = 0.869$. Finally, Topomer Search technology was used to conduct virtual screening in ZINC database, and finally 16 new HDACIs compounds with higher activity were designed.

Keywords

HDACs, Tomoper CoMFA, HDACIs, 3D-QSAR

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在肿瘤的表观遗传学研究中,组蛋白去乙酰化酶(histone deacetylases, HDACs)的主要功能是催化组蛋白的去乙酰化,抑制其转录活性,组蛋白的乙酰化修饰对肿瘤的发生起重要作用。如果正常的细胞体出现核内组蛋白乙酰化与去乙酰化失衡,正常细胞周期与细胞代谢就会发生改变从而诱发肿瘤[1]。

大多数 HDACs 活性都具有 Zn^{2+} 依赖性,组蛋白去乙酰化酶抑制剂(histone deacetylase inhibitor, HDACIs)可与 HDACs 活性位点的 Zn^{2+} 离子结合,干扰 HDACs 的活性[2],增加细胞内组蛋白的乙酰化程度,抑制肿瘤细胞的增殖,诱导细胞分化和(或)凋亡。因此,HDACIs 的结构研究对于设计新型抗癌药物的合成具有重要的意义。

本文主要利用 Topomer CoMFA 方法研究对香豆素基苯甲酰胺作为组蛋白去乙酰化酶抑制剂的 HCT116 细胞系和 A2780 细胞系活性进行研究,分析其药物结构与生物活性之间的关系,建立 3D-QSAR 模型,根据主成分分析和偏最小二乘的表征手法对所建立的模型进行验证,然后在 ZINC 数据库中进行构象搜索,设计出更具活性的新型组蛋白去乙酰化酶抑制剂,最后结合分子对接,对新设计的化合物进行进一步分析,研究化学结构与生物活性之间的相互影响关系。

2. 实验材料和方法

2.1. 分子构建与数据来源

本文分子构建中化合物所包含的小分子结构是利用 SYBYL2.0-X 软件中的 sketch molecule 绘图工具进行构建。全部分子加载 Gasteger-Huckel 电荷,采用标准的 Tripos 分子力场和 Powell 能量梯度算法对所有小分子进行能量优化,最大迭代次数 1000 次,能量收敛限定设为 $0.005 \text{ Kcal}\cdot\text{mol}^{-1}$,其余参数均采用系统默认值[3][4]。能量最小化使得化合物都处于能量最低的稳定状态。

本文对 Abdizadeh [5]组报道的 21 种以香豆素为基础的苯扎酰胺类新一类抗癌药物和 HDACs 抑制剂进行了 QSAR 研究, 研究其降低半数感染细胞有效浓度(IC_{50} 单位 $mol \cdot L^{-1}$)。为了计算方便, 把这些 IC_{50} 值转换为相应的 $pIC_{50}(-\lg IC_{50})$ 值用来表征其生物活性。21 种以香豆素为基础的苯扎酰胺类 HDAC 抑制剂化学结构与活性数据见表 1。将数据集随机分为两个部分, 分别为用于生成 QSAR 模型的训练集(14 个化合物, 占 2/3)和用于模型外部验证的测试集(7 个化合物, 占 1/3)。

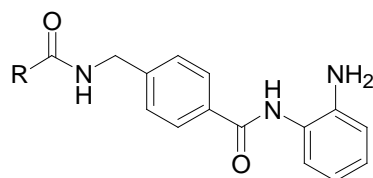
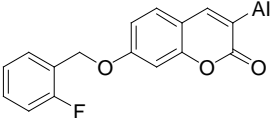
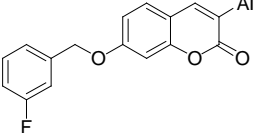
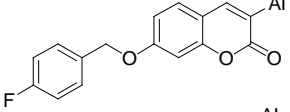
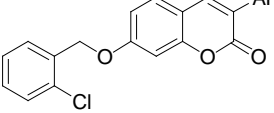
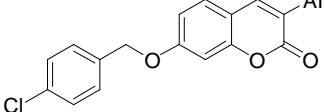
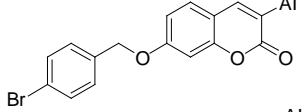
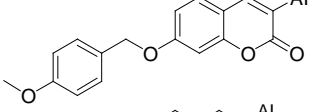
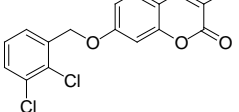
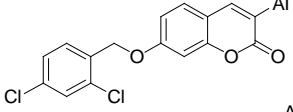
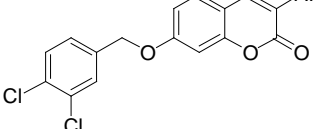


Table 1. Compounds structure and activity value

表 1. 化合物的结构及 pIC_{50} 值

Compd	R	pIC_{50} (HCT116)		pIC_{50} (A2780)	
		Exper	Pred	Exper	Pred
1		4.94	5.16	4.26	4.40
2		5.29	5.16	4.40	4.40
3*		5.69	5.46	4.48	4.40
4		5.51	5.45	4.67	4.70
5		5.64	5.52	4.77	4.76
6*		5.96	5.58	4.83	4.80
7		5.18	5.16	4.56	4.40
8		5.60	5.76	4.89	5.06
9*		5.85	5.59	4.68	4.89
10		5.88	6.01	5.00	5.04
11		6.60	6.51	5.69	5.39

Continued

12*		5.03	5.35	4.76	4.69
13		5.43	5.39	4.80	4.69
14		5.61	5.52	5.01	5.03
15*		5.32	5.32	5.00	4.68
16		5.75	6.02	5.22	5.28
17		6.38	6.01	5.27	5.30
18*		6.00	6.19	5.19	5.29
19		5.21	5.12	4.53	4.40
20		5.37	5.52	4.62	4.84
21*		6.10	5.76	5.31	4.99

*为测试集。

2.2. 三维定量构效关系

定量构效关系(Quantitative Structure-Activity Relationship, QSAR)是一种探索分子结构与其生物活性关系的常用研究手段。通常使用数学模型来研究分子结构和分子生物活性之间的定量关系,以分子结构参数为自变量,分子生物活性为因变量来建立出活性与结构之间的定量关系表达式[6]。三维定量构效关系(3D-Quantitative Structure-Activity Relationship, 3D-QSAR)是根据化合物和生物大分子的三维结构进行 QSAR 的研究。与传统的 QSAR 方法相比,3D-QSAR 增添了生物活性分子的三维构象的性质,在 QSAR 中引入了与生物活性分子三维结构信息有关的参数作为变元,因此能更加精确地反映生物活性分子与受体作用的真实情况,更深刻地说明药物与受体之间的相互作用机制,为新药的研发设计提供一定的理论

基础。

2.3. 比较分子力场分析法

比较分子力场(Comparative molecular field analysis, CoMFA)是研究药物-受体三维定量构效关系(3D-QSAR)最常用的方法之一。CoMFA 能够充分考虑分子的三维结构信息,通过分子结构表征立体场和静电场,再用偏最小二乘法(Partial least squares method, PLS)获取这些三维特征信息与化合物活性之间的联系[7]。通过构建分子结构,获取分子的最低能量构象,选择公共骨架进行分子叠合,根据探针原子计算化合物周围的分子力场,分析药物结构与活性的关系。其中,分子叠合结果对模型预测的精准性有很大的影响[8]。因此要想获取更精确的结果往往需要尝试多种方式进行分子叠合,并且需要花费大量的时间。

2.4. 易位体比较分子力场

易位体比较分子力场(Topomer CoMFA)是在 2004 年,由 Cramer 等人[9]提出,是一种可以预测化合物生物活性或性质的新型 3D-QSAR 方法。不同于 CoMFA, Topomer CoMFA 是同时结合了 Topomer 与 CoMFA 技术。它弥补了传统 CoMFA 方法中的手动叠合分子的缺点,它结合了 3D-QSAR 构建技术和基于配体的虚拟筛选技术,并且可在短时间内构建较为可靠的 3D-QSAR 模型,具有便捷、快速、准确的特点[10]。另外, Topomer CoMFA 还具有重复性高的优势。利用 Topomer CoMFA 可快速建立预测模型并进行分析与评价,为同类小分子抑制剂的结构优化提供理论依据[11]。

2.5. 虚拟筛选

本文中的分子设计采用的是虚拟筛选的方法,使用 SYBYL2.0-X 中 Topomer Search 技术进行分子虚拟筛选。此工具是在数据库中基于 R 基团的搜索技术进行搜索,可以找到训练集中没有出现过的更高活性的化学结构片段,进而将找到的高贡献值的片段通过重新排列组合的方式设计出更高活性的化合物[12]。本文中,将具有最高活性的 11 号化合物作为模板分子进行切割以获得分子的结构片段,用获得的模型预测其对活性的贡献值大小[13]。采用 Topomer search 技术在 ZINC 数据库中对基于 R 基团的虚拟筛选,用距离作为打分的依据,去评判搜索化合物小片段与数据库化合物片段之间的相似性[14]。Topomer 的最大距离设置为 185,在数据库中筛选出来的化合物片段 Topomer 距离要尽量接近 185,且片段的贡献值超过模板分子 R 基团的贡献值,筛选出来符合条件的片段,最后将筛选出来的小分子片段重新进行组合排列去替换模板分子中相应的 R 基团片段进行分子设计[15]。

2.6. 分子对接

分子对接(molecular docking)是通过受体的特征以及受体和药物分子之间的相互作用方式来进行药物设计的方法。其主要研究化合物与所选蛋白质活性位点残基之间的相互作用,比较它们之间的亲和性,并根据停靠能量最低来评价最佳构象。通过小分子配体和受体之间的几何、能量匹配与识别,预测其结合模式和亲合力,找到两者之间最佳的结合模式[16]。是计算机辅助药物研究领域的一项重要技术。

本文中采用的是 SYBYL-X2.0 软件中的 Surflex-Dock 模块进行分子对接。本次对接所使用的蛋白酶晶体来源于 PDB 数据库(Protein Data Bank),晶体结构的 ID 为: 3JSX。在进行分子对接前对 3JSX 大分子蛋白进行预处理,将需要的小分子配体从大分子复合物中提取出来,不需要的小分子配体及所有的水分子删除,给蛋白加氢、添加 Castelger 电荷。根据小分子配体确定对接的活性位点,设置对接格子的格

点为 $40 \times 40 \times 40$ ，格点间距为 0.375 \AA 。然后采用拉马克遗传算法对分子对接的配体构象进行搜索，每个配体分子进行 50 次独立对接操作，能量评估的最大值设置 250 万。

先将活性最高的 11 号化合物作为模板与蛋白晶体进行对接，之后再设计的分子全部与之对接，最后观察所有化合物的打分函数 Total-Score、Crash 和 Polar 的值[17] [18]。Total-Score 的打分函数值表示从大分子蛋白中提取出来的小分子配体与受体的亲和力，其值越高亲和力越高；Crash 表示从大分子蛋白中提取出来的配体与受体之间的不舒适度，其值越接近零越好。Polar 为极性函数得分，当结合位点在分子表面时，其值越大越好；在分子内部时，其值越小越好[19]。

2.7. Topomer CoMFA 模型的建立

不同于 CoMFA 和 CoMSIA，Topomer CoMFA 不需要采用基于骨架的三维结构叠合。通常 Topomer CoMFA 的操作流程大致可分为以下两个步骤：

第一、根据训练集分子结构片段建立 Topomer 的三维模型。通过化合物的结构特点确认化合物结构的断裂方式，Topomer CoMFA 可以自动识别结构片段的化学结构和电荷排布，将整个配体分子切割成两个或两个以上的小片段，没有被切割的化合物可以进行手动切割，然后小片段会自行调整，自动生成 Topomer 模型和小片段的三维构像。

第二、采用偏最小二乘法分析(PLS) [20]。先通过抽一法(Leave-One-Out)进行交叉验证的 CoMFA 分析，形成一个预测模型。在生成模型的过程中，CoMFA 方法可处理大量数据客观分析和自动匹配相应的化合物结构，以使 Topomer CoMFA 形成的预测模型更有效[11]。然后根据交叉验证得到的最佳主成分数，进行非交叉验证回归，建立回归模型表达式。

采用 LOO 交叉验证方法作为内部验证，通过公式计算得到交叉验证系数 q^2 ，非交叉验证系数 r^2 。相关公式与说明如下[21] [22] [23] [24]：

$$q^2 = 1 - \frac{\sum_{n=1}^{\infty} (\hat{y}_i - y_i)^2}{\sum_{n=1}^{\infty} (y_i - \bar{y})^2}$$

式中 \hat{y}_i 和 y_i 分别为预测值和实验值， \bar{y} 和 \hat{y} 分别为训练集的实验值和预测值的平均活性值。 $\sum_{n=1}^{\infty} (\hat{y}_i - y_i)^2$ 是预测残差平方和(PRESS)。

良好的内部验证只能说明化合物训练集的 q^2 较高，但并不表明所建立模型的预测能力较高，因此外部验证是必要的，通过计算测试集的化合物的生物活性来进一步验证 3D-QSAR 模型的预测能力[22]。在交叉验证之后，通常采用常规相关系数 r^2 进行非交叉验证分析，计算出的估计标准误差(SEE)和 F 值。

$$r^2 = \frac{\left[\sum (y_i - \bar{y}_i)(\hat{y}_i - \hat{y}) \right]^2}{\sum (y_i - \bar{y}_i)^2 \times \sum (\hat{y}_i - \hat{y})^2}$$

通常，交叉验证系数 q^2 越大，相关系数越大，SEE 越小，表示相关性越好，方程的预测能力越强， q^2 的标准误差(q^2 stderr)和 r^2 的标准误差(r^2 stderr)越小说明模型质量越高[11]。通常，当交叉验证系数 q^2 大于 0.5 且非交叉验证系数 r^2 大于 0.6 时，说明由该方法所建立的模型是比较理想的 Topomer CoMFA 模型[25]。

2.8. 分子片段划分

以活性最高的化合物 11 作为模板, 将所有化合物按照图 1 所示的分割方式分割, 将分子划分为三个部分, 未自动分割的分子进行手动切割, 切割后的化合物被自动划分为 R1 (红色)、R2 (蓝色) 和公共骨架(绿色)。

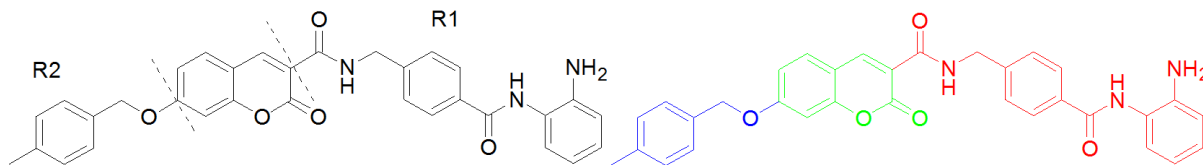


Figure 1. Template molecule segmentation mode (left) and fragment schematic diagram (right)
图 1. 模板分子分割方式(左)和片段示意图(右)

3. 结果与讨论

3.1. Topomer CoMFA 模型分析

采用 Topomer CoMFA 方法分别对以香豆素为基础的苯扎酰胺类化合物作为 HDAC 抑制剂的 21 种化合物的 HCT116 细胞系和 A2780 细胞系研究结果如下:

Table 2. Results of Topomer CoMFA model
表 2. Topomer CoMFA 模型结果

System	Com	q^2	r^2	q^2 stderr	r^2 stderr	SEE
HCT116	3	0.517	0.880	0.36	0.18	0.178
A2780	2	0.572	0.869	0.27	0.15	0.150

如表 2 所示, Topomer CoMFA 模型 HCT116 细胞体系的统计学参数为最佳主成分值为 3, 交叉验证相关系数 q^2 为 0.517, 非交叉验证系数 r^2 为 0.880; A2780 细胞体系的统计学参数为最佳主成分值为 2, 交叉验证相关系数 q^2 为 0.572, 非交叉验证系数 r^2 为 0.869。数据显示, 两者的 q^2 大于 0.5, r^2 大于 0.6, 说明通过该方法建立的模型是比较理想的 Topomer CoMFA 模型, 其统计结果具有很高的预测能力。同时预测值的 SEE 为 0.039, q^2 stderr 为 0.36, r^2 stderr 为 0.18, 进一步说明所建立的模型具有较高的可靠性。

图 2 和图 3 分别是 HCT116 细胞体系与 A2780 细胞体系的训练集和测试集化合物的实验值与预测值之间的线性回归相关图和柱状分析图。从图中可以看出, 线性回归图中所有样本均匀地分布在 45°线附近, 柱状分析图中所有样本基本在同一水平高度。由图可知两种细胞系构建的 QSAR 模型是比较可靠的, 具有较好的拟合力。

3.2. 3D-QSAR 分析

图 4 是以活性最高的 11 号分子为模板建立的 Topomer CoMFA 模型的分子场三维等势图, 根据切割方式, 所有化合物的 R1 基团都是一样的, 所以无法对 R1 基团进行比较分析, 因此本文只对两种细胞系的 R2 基团进行比较分析。

图 4(a)与图(b)分别为 HCT116 细胞系 R_2 基团的立体场等势图和静电场等势图; A2780 细胞系 R_2 基团的立体场等势图和静电场等势图为图 4(c)与(d)。在图 4(a)和(c)中, 绿色的部分表示提高取代基的体积可以增加活性; 黄色的部分则相反, 减少取代基的体积可以提高化合物活性; 在图 4(b)和(d)中, 红色部分表示引入带负电子的基团可以提高化合物活性, 而蓝色部分则表示引入带正电子的基团可以提高化合物活性

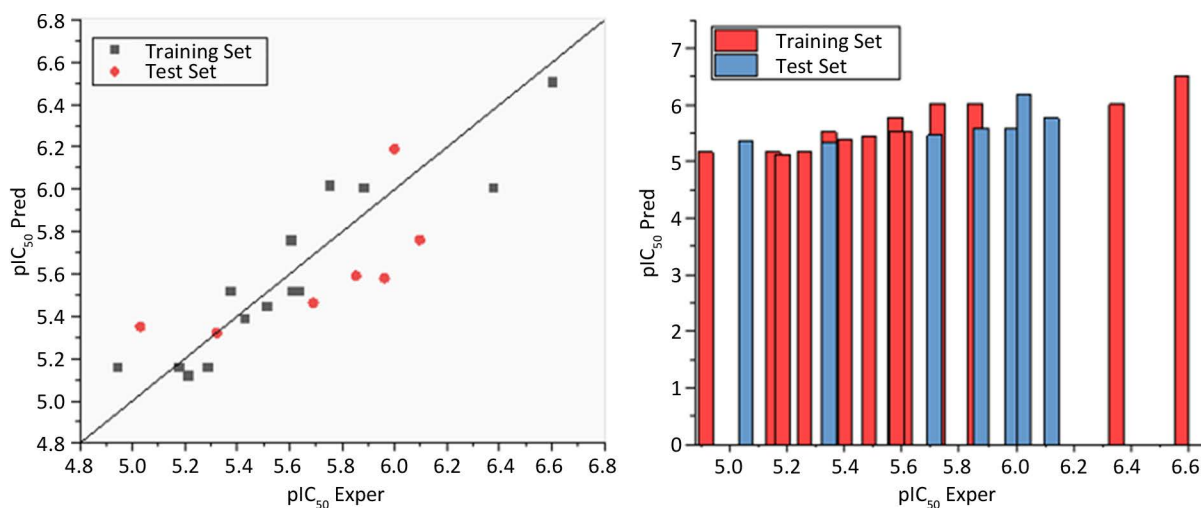


Figure 2. Linear regression diagram (left) and histogram diagram (right) of HCT116 cell system
图 2. HCT116 细胞体系的线性回归图(左)和柱状分析图(右)

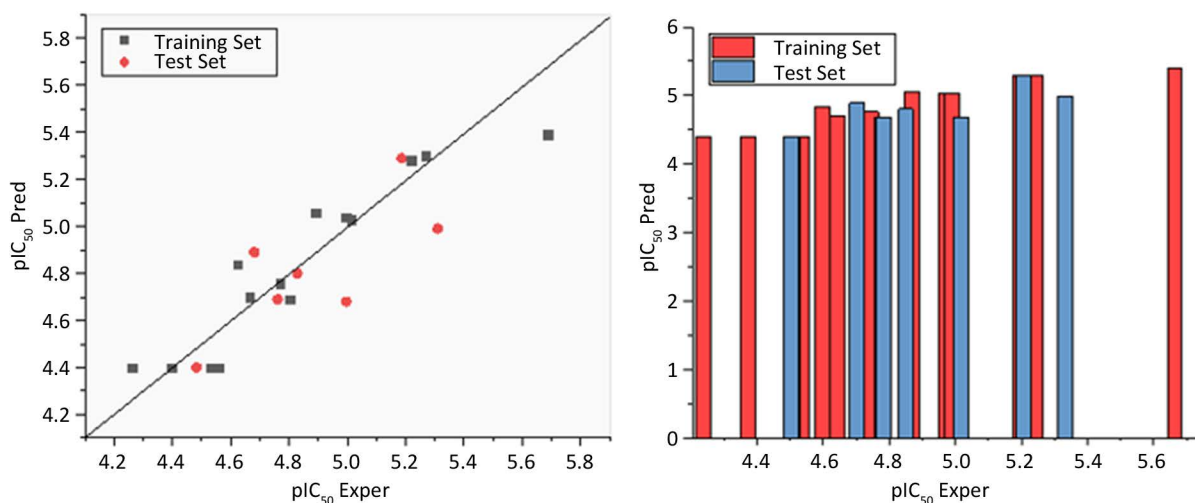


Figure 3. Linear regression diagram (left) and column analysis diagram (right) of A2780 cell system
图 3. A2780 细胞体系的线性回归图(左)和柱状分析图(右)

对于 HCT116 细胞系, 在图 4(a)中的 R_3 、 R_4 和 R_5 取代基的位置上有绿色区域, 若在此区域有大面积基团则活性相对较高。例如, 化合物 10、11、16、17、18 和 21 (pIC_{50} 值分别为 5.88、6.60、5.75、6.38、6.00 和 6.10) 相比化合物 8 (pIC_{50} 值为 5.60) 都是增加了取代基的体积(分别用 $-C$ 、 $-C$ 、 $-Cl$ 、 $-Br$ 、 $-CO$ 和 $-Cl$ 取代了 $-H$), 体积增大, 活性明显升高; 同理, 化合物 R_6 位置有大面积黄色区域, 需要在此位置引入小基团的基团以提高化合物的活性。例如, 化合物 12 和 15 (pIC_{50} 值分别为 5.03 和 5.32) 与化合物 8 (pIC_{50} 值为 5.60) 相比, 在 R_6 取代基的位置上用 $-F$ 和 $-Cl$ 取代了 $-H$ 活性降低。

从图 4(b)中可以看出 R_4 、 R_5 和 R_6 取代基上有蓝色区域, 因此在这些基团上减小取代基的电负性或者引入带正电荷的基团有利于提高活性。例如, R_6 取代基的位置上化合物 8、9 和 15 (pIC_{50} 值分别为 5.60、5.85 和 5.32) 分别在用 $-H$ 、 $-C$ 和 $-Cl$ 取代化合物 12 (pIC_{50} 值为 5.030) 的 $-F$, 电负性增强, 活性明显升高; 在 R_4 取代基的位置上化合物 14 (pIC_{50} 值为 5.61) 用 $-F$ 取代化合物 11、16、17 和 18 (pIC_{50} 值分别为 6.60、5.75、6.38、6.00 和 6.10) 相同位置上的 $-C$ 、 $-Cl$ 、 $-Br$ 和 $-CO$ 基团活性明显降低。另外, 在图 4(b)中 R_1 、 R_2 取代基的位置被红色区域包围, 因此需要增加这些基团上取代基的电负性来提高活性。例如, 化合物 4、5 和 7 (pIC_{50} 值分别为 5.51、

5.64、和 5.18)相比化合物 1 (pIC_{50} 值为 4.94)都是在取代基上引入了高电负性的基团, 活性都有所提高。同理, 对于 A2780 细胞系也是如此。

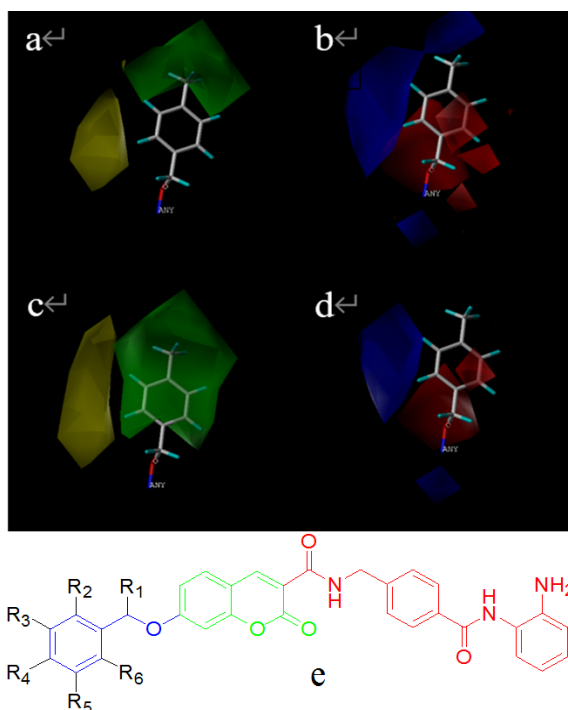


Figure 4. 3D equipotential diagram of Topomer CoMFA model based on template 11

图 4. 基于模板 11 的 Topomer CoMFA 模型的三维等势图

因此, 要想提高化合物的活性应该优先增大整个取代基的体积尤其是 R_4 基团取代基的体积, 并且减少 R_6 基团取代基的体积。同时在 R_4 、 R_5 和 R_6 取代基上减小取代基的电负性或者引入带正电荷的基团, 在苯环的内、外侧和 R_1 基团上增加取代基的电负性或者引入带负电荷的基团。

3.3. 分子设计结果

本文的分子设计是以 3D-QSAR 分析作为参考, 基于化合物 11 的 R_2 基团, 采用 Topomer Search 技术在 ZINC 数据库中的 Drug-like 分子库中进行搜索[25], 以 R_2 基团作为提问结构, 在化合物数据库中进行筛选。在筛选出的这些分子结构片段中挑选出活性贡献值高于 11 号模板分子的结构片段作为新的 R_2 基团。本文中所建立的 Topomer CoMFA 模型中 HCT116 细胞系的 R_2 基团贡献值为 1.36, 挑选的时候选择 Topomer 距离越接近 185 且片段的贡献值超过模板分子 R_2 基团的贡献值的 R 基团。选择出符合条件的 16 个集团, 最终设计出 16 个新的化合物。按照训练集分子的构建方法, 在 SYBYL-X 2.0 中构建出新分子的三维结构, 然后优化、命名, 用化合物 11 作为模板预测其活性, 16 个新设计的化合物结构和预测活性值(pIC_{50})见表 3。

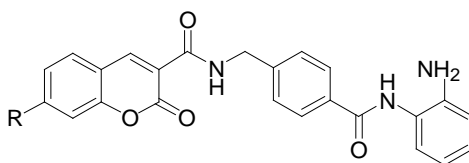
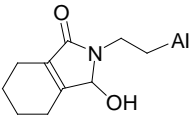
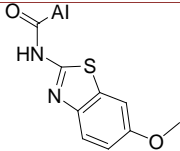
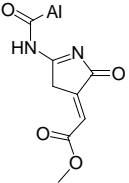
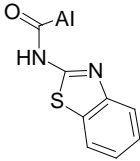
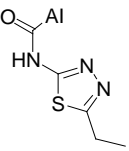
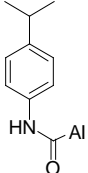
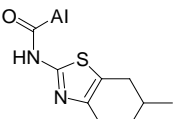
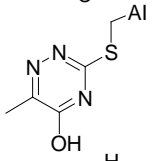
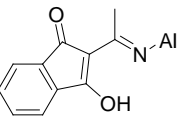
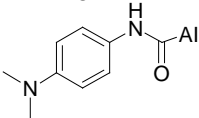
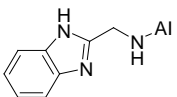
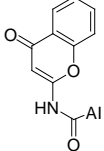
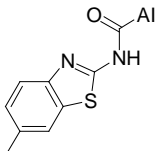
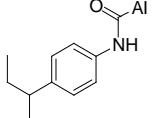
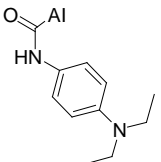
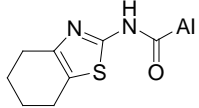


Table 3. Molecular design and predicted activity values of HCT116 cell lines and A2780 cell lines
表 3. HCT116 细胞系与 A2780 细胞系的分子设计和预测活性值

NO	R	pIC_{50} Pred	NO	R	pIC_{50} Pred
1-01		6.85	1-09		6.64
1-02		6.91	1-10		6.61
1-03		6.68	1-11		6.62
1-04		6.72	1-12		6.63
1-05		6.64	1-13		6.62
1-06		6.66	1-14		6.60
1-07		6.66	1-15		6.61
1-08		6.64	1-16		6.72

如表 3 所示, 新设计出的 16 个分子的预测活性值(pIC_{50})都在 6.60 以上, 比模板分子的活性值高。

3.4. 分子对接结果分析

将模板分子和新设计出的 16 个分子都进行分子对接, 模板分子作为对照, 以打分函数 Total-Score、Crash 和 Polar 的值作为评判标准, 结果如表 4 所示。

Table 4. Molecular docking scoring function and number of hydrogen bonds
表 4. 分子对接打分函数和氢键个数

NO	p/C ₅₀ pred	total-score	crash	polar	H-bonds
contrast	6.51	8.2690	-4.2223	1.7664	3
1-01	6.85	8.9593	-3.2400	2.4539	4
1-02	6.91	-0.6480	-8.2273	1.0733	2
1-03	6.68	5.4217	-2.5780	1.2330	3
1-04	6.72	1.2597	-1.0678	0.0670	2
1-05	6.64	6.6386	-2.0465	2.0208	3
1-06	6.66	8.0846	-0.8779	2.2192	3
1-07	6.66	3.8777	-2.3619	0.2803	2
1-08	6.64	6.7794	-1.4545	2.1260	4
1-09	6.64	6.2089	-2.7046	0.7645	3
1-10	6.61	7.1137	-2.9435	1.4274	3
1-11	6.62	8.4407	-2.4090	2.0832	3
1-12	6.63	6.9264	-1.2237	2.3765	2
1-13	6.62	6.5047	-1.1232	2.2471	3
1-14	6.60	8.4058	-1.3742	3.2605	4
1-15	6.61	5.1668	-1.5203	0.9896	2
1-16	6.72	7.3831	-1.2547	2.1163	3

对接分析图见图 5，其中棒状化合物代表分子配体，球棍状化合物代表氨基酸残基，黄色虚线为氢键。从图 5(a)可以看出，11 号模板分子配体与 3JSX 蛋白晶体结构中的主要氨基酸残基共形成了 3 个氢键，与 Tyr128、Tyr126、和 Trp105 形成氢键相互作用，Total-score, crash 和 polar 分别为 8.2690, -4.2223, 1.7664。

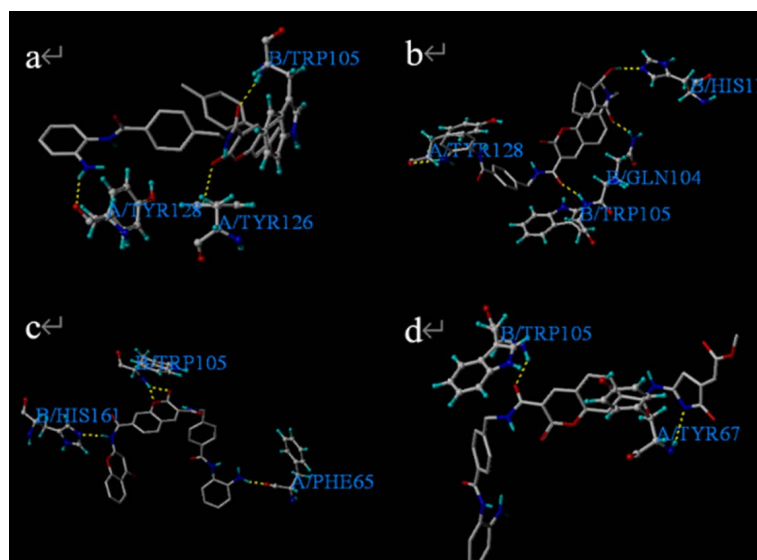


Figure 5. Molecular hydrogen bonding interaction diagram
图 5. 分子对接氢键作用图

如图 5(b)~(d)分别表示设计的新化合物 1-01, 1-14, 1-02 与 3JSX 蛋白晶体进行分子对接的活性位点, 黄线色虚线表示氢键作用力。图 5(b)所示, 化合物 1-01 与蛋白酶晶体的 His11, Tyr128, Gln104, 和 Trp105 氨基酸残基形成氢键; 图 5(c)为化合物 1-14 与 3JSX 蛋白中的 Trp105, His161, 和 Phe65 氨基酸残基形成氢键。结果显示所设计的化合物对接结果是可靠。另外, 从表 4 可以看出, 新设计的化合物 1-02、1-04 和 1-07 均具有较高的预测活性, 但总打分函数相对较低。特别是化合物 1-02, 在设计的 16 个分子中化合物 1-02 的预测活性值(pIC₅₀)最高, 但总打分函数最低。因此, 新分子的设计不仅需要具有较高的活性值, 还需要分子与大蛋白之间的空间匹配和能量匹配。

4. 结论

本文利用 Topomer CoMFA 方法研究了香豆素基苯甲酰胺作为组蛋白去乙酰化酶抑制剂。HCT116 细胞系和 A2780 细胞系的 HDAC 抑制剂活性进行了研究, 建立了稳定且预测能力良好的 3D-QSAR 模型; 结合两种细胞系通过 Topomer search 进行在 ZINC 数据库筛选出活性更好的分子片段, 设计出了 16 个比模板分子活性更高的新化合物; 最后, 使用 Surflex-dock 方法将模板分子和新设计的分子与 3JSX 蛋白进行对接, 对抑制剂和受体蛋白之间的结合机制进行研究。并最终在新设计的化合物中确定了 13 个新化合物, 但这些化合物是否可以应用于实践中, 还有待进一步的研究。本研究对于理解香豆素基苯甲酰胺作为组蛋白去乙酰化酶抑制剂的生物活性, 寻找新型抗癌药物具有一定的指导作用。

基金项目

本课题得到国家自然科学基金(21475081)、陕西省自然科学基金(2019JM -237)、陕西科技大学研究生创新基金的资助。

参考文献

- [1] Minucci, S. and Pelicci, P.G. (2006) Histone Deacetylase Inhibitors and the Promise of Epigenetic (and More) Treatments for Cancer. *Nature Reviews Cancer*, **6**, 38-51. <https://doi.org/10.1038/nrc1779>
- [2] Li, Y. and Seto, E. (2016) HDACs and HDAC Inhibitors in Cancer Development and Therapy. *Cold Spring Harbor Perspectives in Medicine*, **6**, 826-831. <https://doi.org/10.1101/cshperspect.a026831>
- [3] Liang, J., Li, Z., Zhang, T., et al. (2013) 3D-QSAR Studies of 2-Piperazinyl Thiazole Hydroxamic Acids Based on Topomer CoMFA Method. *Computers & Applied Chemistry*, **30**, 781-784.
- [4] 陈玉珍, 范新景, 张浩, 等. 基于 Topomer CoMFA 和 Surflex-dock 的黄酮类醛糖还原酶抑制剂的 3D-QSAR 与作用模式研究[J]. 河南工业大学学报(自然科学版), 2018(39): 91-95.
- [5] Abdizadeh, T., Kalani, M.R., Abnous, K., et al. (2017) Design, Synthesis and Biological Evaluation of Novel Coumarin-Based Benzamides as Potent Histone Deacetylase Inhibitors and Anticancer Agents. *European Journal of Medicinal Chemistry*, **132**, 42-62.
- [6] Marshall, G.R. and Cramer, R.D. (1988) Three-Dimensional Structure-Activity Relationships. *Trends in Pharmacological Sciences*, **9**, 285-289. [https://doi.org/10.1016/0165-6147\(88\)90012-0](https://doi.org/10.1016/0165-6147(88)90012-0)
- [7] Tenenhaus, M., Vinzi, V.E., Chatelinc, Y.M., et al. (2005) PLS Path Modeling. *Computational Statistics & Data Analysis*, **48**, 159-205. <https://doi.org/10.1016/j.csda.2004.03.005>
- [8] Shirota, Y., Luo, H., Qin, W., et al. (2002) Hepatitis C Virus (HCV) NS5A Binds RNA-Dependent RNA Polymerase (RdRP) NS5B and Modulates RNA-Dependent RNA Polymerase Activity. *Journal of Biological Chemistry*, **277**, 11149-11155. <https://doi.org/10.1074/jbc.M111392200>
- [9] Cramer, R.D., Cruz, P., Stahl, G., et al. (2008) ChemInform Abstract: Virtual Screening for R-Groups, Including Predicted pIC₅₀ Contributions, within Large Structural Databases, Using Topomer CoMFA. *Journal of Chemical Information & Modeling*, **48**, 2180-2195. <https://doi.org/10.1021/ci8001556>
- [10] Jilek, R.J. and Cramer, R.D. (2004) Topomers: A Validated Protocol for Their Self-Consistent Generation. *Journal of Chemical Information and Computer Sciences*, **35**, 1221-1227. <https://doi.org/10.1002/chin.200439196>
- [11] 刘永香, 施海枫. 基于 TopomerCoMFA 方法的姜黄素类化合物的三维定量构效关系研究[J]. 化工技术与开发,

- 2014, 43(2): 36-38.
- [12] 刘永澜, 李月婷, 史博智, 等. 基于 Topomer CoMFA 和 Surflex-dock 的 GSK-3 β 抑制剂的 3D-QSAR 与作用模式研究[J]. 中国科学: 化学, 2013(2): 198-208.
- [13] Xiang, Y.H., Song, J. and Zhang, Z.Y. (2014) Topomer-CoMFA and Virtual Screening Studies of Azaindole Class Renin Inhibitors. *Combinatorial Chemistry & High Throughput Screening*, **17**, 458-472. <https://doi.org/10.2174/1386207317666140107094708>
- [14] Tong, J.B., Zhan, P., Bai, M., *et al.* (2016) Molecular Modeling Studies of Human Immunodeficiency Virus Type 1 Protease Inhibitors Using Three-Dimensional Quantitative Structure Activity Relationship, Virtual Screening, and Docking Simulations. *Journal of Chemometrics*, **30**, 523-536. <https://doi.org/10.1002/cem.2809>
- [15] Irwin, J.J., Sterling, T., Mysinger, M.M., *et al.* (2012) ZINC: A Free Tool to Discover Chemistry for Biology. *Journal of Chemical Information & Modeling*, **52**, 1757-1768. <https://doi.org/10.1021/ci3001277>
- [16] Duan, A.X., Chen, J., Liu, H.D., *et al.* (2009) Application and Development of Molecular Docking Method. *Journal of Analytical Science*, **25**, 473-477.
- [17] Ghorab, M.M., Ragab, F.A. and Heiba, H.I. (2016) Design and Synthesis of Some Novel 4-Chloro-N-(4-(1-(2-(2-cyanoacetyl)hydrazono)ethyl)phenyl)benzenesulfonamide Derivatives as Anticancer and Radiosensitizing Agents. *European Journal of Medicinal Chemistry*, **117**, 8-18. <https://doi.org/10.1016/j.ejmech.2016.04.009>
- [18] Xu, C. and Ren, Y.J. (2015) Molecular Modeling Studies of [6,6,5]tricyclic Fused Oxazolidinones as FXa Inhibitors Using 3D-QSAR, Topomer CoMFA, Molecular Docking and Molecular Dynamics Simulations. *Bioorganic & Medicinal Chemistry Letters*, **25**, 4522-4528. <https://doi.org/10.1016/j.bmcl.2015.08.070>
- [19] 全建波, 雷珊, 秦尚尚, 等. 基于 Topomer CoMFA 方法的喹诺酮羧酸类衍生物的 QSAR 研究及分子设计[J]. 分析测试学报, 2018(5): 517-524.
- [20] 蔡伟平, 韦星船, 郑成, 等. 姜黄素类似物抑制酪氨酸酶活性 3D-QSAR 模型的研究[J]. 现代食品科技, 2017(8): 41-50.
- [21] 汪斌, 刘蒙蒙, 周朋朋, 等. 嘧啶衍生物类 PI3K/mTOR 双重抑制剂的 3D-QSAR 及分子对接研究[J]. 中国科学: 化学, 2017, 47(7): 865-875.
- [22] Abdizadeha, R., Hadizadehb, F. and Abdizadeh, T. (2019) QSAR Analysis of Coumarin-Based Benzamides as Histone Deacetylase Inhibitors Using CoMFA, CoMSIA and HQSAR Methods. *Journal of Molecular Structure*, **1199**, Article ID: 126961. <https://doi.org/10.1016/j.molstruc.2019.126961>
- [23] Waller, C.L. (2004) A Comparative QSAR Study Using CoMFA, HQSAR and FRED/SKEYS Paradigms for Estrogen Receptor Binding Affinities of Structurally Diverse Compounds. *The Journal for Chemical Information and Computer Scientists*, **44**, 758-765. <https://doi.org/10.1021/ci0342526>
- [24] Perez-Areales, F.J., Betari, N., Viayna, A., *et al.* (2017) Design, Synthesis and Multitarget Biological Profiling of Second-Generation Anti-Alzheimer Rheinehuprine Hybrids. *Future Medicinal Chemistry*, **9**, 965-981. <https://doi.org/10.4155/fmc-2017-0049>
- [25] Ji, Y., Shu, M., Lin, Y., *et al.* (2013) Combined 3D-QSAR Modeling and Molecular Docking Study on Azacycles CCR5 Antagonists. *Journal of Molecular Structure*, **1045**, 35-41. <https://doi.org/10.1016/j.molstruc.2013.03.062>